

Study protocol

Open Access

UK Back pain Exercise And Manipulation (UK BEAM) trial – national randomised trial of physical treatments for back pain in primary care: objectives, design and interventions [ISRCTN32683578]

UK Back pain Exercise And Manipulation (UK BEAM) Trial Team*

Address: Institute of Community Health Sciences, Barts and the London, Queen Mary's School of Medicine and Dentistry, Mile End, London, UK

* Corresponding author

Published: 01 August 2003

Received: 10 April 2003

BMC Health Services Research 2003, 3:16

Accepted: 01 August 2003

This article is available from: <http://www.biomedcentral.com/1472-6963/3/16>

© 2003 UK Back pain Exercise And Manipulation (UK BEAM) Trial Team; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Low back pain has major health and social implications. Although there have been many randomised controlled trials of manipulation and exercise for the management of low back pain, the role of these two treatments in its routine management remains unclear. A previous trial comparing private chiropractic treatment with National Health Service (NHS) outpatient treatment, which found a benefit from chiropractic treatment, has been criticised because it did not take treatment location into account. There are data to suggest that general exercise programmes may have beneficial effects on low back pain. The UK Medical Research Council (MRC) has funded this major trial of physical treatments for back pain, based in primary care. It aims to establish if, when added to best care in general practice, a defined package of spinal manipulation and a defined programme of exercise classes (*Back to Fitness*) improve participant-assessed outcomes. Additionally the trial compares outcomes between participants receiving the spinal manipulation in NHS premises and in private premises.

Design: Randomised controlled trial using a 3 × 2 factorial design.

Methods: We sought to randomise 1350 participants with simple low back pain of at least one month's duration. These came from 14 locations across the UK, each with a cluster of 10–15 general practices that were members of the MRC General Practice Research Framework (GPRF). All practices were trained in the active management of low back pain. Participants were randomised to this form of general practice care only, or this general practice care plus manipulation, or this general practice care plus exercise, or this general practice care plus manipulation followed by exercise. Those randomised to manipulation were further randomised to receive treatment in either NHS or private premises. Follow up was by postal questionnaire one, three and 12 months after randomisation. The primary analysis will consider the main treatment effects before interactions between the two treatment packages. Economic analysis will estimate the cost per unit of health utility gained by adding either or both of the treatment packages to general practice care.

Background

Back pain is a widespread and costly health problem in many countries including the UK[1]. In the UK the

number of days of Invalidity Benefit attributable to spinal disorders rose threefold over the 1980s [2]. There is little evidence that the rise in reported disability reflects

changes in pathology, prevalence or even morbidity. Since the early 1990s there have been many systematic reviews of the effectiveness of physical interventions for low back pain. There appears to be a consensus, reflected in national guidelines in the UK [3,4], USA [5], the Netherlands [6] and elsewhere [7], that patients with acute back pain should be encouraged to return to normal activity as soon as possible and have early access to physical therapy. It is less clear what form such physical therapy should take. Two commonly suggested physical treatments for low back pain are manipulation and physiotherapist led exercise programmes. The evidence to support their use is far from conclusive.

Evidence for manipulation

There have been many systematic reviews of the effectiveness of manipulation. Koes *et al.* [8] reviewed 38 trials and concluded that, although some results were encouraging, further trials were needed to establish the effectiveness of manipulation. In contrast, Shekelle *et al.* [9] did a meta-analysis combining data from nine trials and concluded that manipulation could increase the rate of recovery from acute uncomplicated low back pain, but that there were insufficient data to provide evidence for the effectiveness of manipulation in patients with chronic pain. The US Agency for Health Care Policy and Research (AHCPR) [5] reviewed four meta-analyses and 12 additional randomised trials and also concluded that manipulation could speed the recovery of patients with acute back pain and that the evidence to support the use of manipulation for radiculopathies or longer standing back pain was inconclusive. The systematic review of reviews by Assendelft *et al.* [10] was highly critical of the general standard of reviews. Nevertheless, nine of their best ten reviews, as judged by methodological criteria, reported positive effects of manipulation.

Subsequently Assendelft *et al.* in the most recent systematic review of which we are aware [11] concluded that spinal manipulative therapy has no statistically or clinically significant advantage over general practice care, analgesics, physical therapy, exercise or back school for acute or chronic back pain. However, they did find some advantages for spinal manipulation when compared to sham manipulation or therapies judged to be ineffective or harmful. This systematic review, that was published after we had collected our data, went on to recommend that any future trials of manipulation should concentrate on cost-effectiveness rather than effectiveness. This trial includes a cost-effectiveness analysis.

Evidence for exercise

Fewer systematic reviews have examined the effectiveness of exercise for back pain. Koes *et al.* [12] examined 16 randomised trials and found most to be of poor quality with

inconclusive results. The AHCPR [5] extended this to 20 randomised trials, of which six related to acute back pain. With the exception of one well-designed study [13], however, the studies included other interventions that made it difficult to evaluate the effects of exercise. Faas [14] extended the data set to 28 randomised trials but concluded that there was insufficient evidence that *specific* back exercises produce clinically significant improvement in acute low back pain. While endorsing these conclusions, Waddell *et al.* [3] cited evidence that general exercise programmes can improve pain and functional levels in those with chronic low back pain. More recently, a Cochrane review identified 39 studies and concluded that the data did not indicate that specific exercises are effective for the treatment of acute low back pain. However, (general) exercises may be helpful for chronic low back pain patients to increase return to normal daily activities and work [15].

Waddell *et al.* [16] carried out a systematic review of advice about staying active with back pain and concluded that advice to continue normal activities leads to less chronic disability and time off work than the traditional advice to rest and 'let pain be your guide'. Subsequent Cochrane reviews of treatments for acute low back pain and sciatica concluded that 'advice to stay active alone has little beneficial effect for patients' [17] and that, compared to bed rest, advice to stay active alone will have limited beneficial effects [18].

On the basis of the limited evidence available the UK Clinical Standards Advisory Group [2] and the authors of the national acute low back pain guidelines, produced by the Royal College of General Practitioners (RCGP) [3,4], advised widespread access to physical treatment for back pain. However, because the evidence underpinning this approach is weak, and the health and social impact of back pain is large, it became a NHS research priority. Consequently, the UK MRC formed an international working party, including back pain clinicians and researchers from different disciplines, to design a national trial to evaluate the effectiveness of different physical treatments for back pain in primary care: the UK Back Pain Exercise and Manipulation Trial (UK BEAM).

Methods

Trial design

The trial had a two-dimensional 3 × 2 factorial design [19]. Each participant was randomised:

1. between the 'Back to Fitness' progressive exercise package and general practice management; and
2. between the spinal manipulation package and general practice management.

Those randomised to the manipulation package were further randomised to be treated in NHS or private premises, to allow the effect of treatment location on outcome to be measured. Each participant had an equal chance of receiving general practice management, the manipulation package only, the exercise package only or both manipulation and exercise. Participants randomised to manipulation each had an equal chance of receiving treatment in NHS or private premises. (Figure 1).

Hypotheses tested

Essentially explanatory hypotheses

The trial tested a total of four 'null' hypotheses. The three primary null hypotheses were, that in people consulting their general practitioners with back pain there is no difference in clinical outcome between:

- a. Those receiving an additional defined package of spinal manipulation or general practice management only;
- b. Those receiving this package of spinal manipulation in NHS premises or those receiving it in the private premises of chiropractors, osteopaths or physiotherapists; and
- c. Those receiving a defined package of exercises (*Back to Fitness*) or general practice management only.

The secondary null hypothesis was that:

- d. In those people receiving both spinal manipulation and the *Back to Fitness* exercise package, the improvement in clinical outcome is merely the cumulative effect of those treatments, i.e. that there is no interaction.

For hypothesis b, the feasibility study confirmed that we could recruit therapists to perform spinal manipulation from those working in their own premises who also worked, or were willing to work, in NHS premises. For hypothesis d, the feasibility study confirmed that those performing spinal manipulation and those providing *Back to Fitness* exercise classes could deliver their own package to participants who also received the other package.

Essentially pragmatic estimates

The trial will estimate (to within a confidence interval) the cost per unit of health utility gained by adding to general practice management:

- a. The spinal manipulation package in NHS premises;
- b. The spinal manipulation package in the private premises;
- c. The exercise package; and

- d. The spinal manipulation package (NHS or private) plus the exercise package.

Interventions

General practice management (the 'control' treatment)

All participants in the trial continued to be managed by their general practitioners even if randomised to spinal manipulation or the exercise package. Members of the MRC Working Party and the trial team developed an active management package for use by general practices within the trial [20]. This included training in active management for all clinical and support staff from all participating practices.

The AHCPR [5], the Clinical Standards Advisory Group (CSAG) [2], and the RCGP [3,4] have all advocated active management. The recommended advice is also available through the evidence-based patient resource '*The Back Book*' [21]. An early evaluation showed that this booklet can produce positive changes in beliefs about back pain [22]. Open-ended questions showed that almost all respondents found it easy to understand, interesting, and helpful, and assimilated its main messages. A more recent trial showed that it improved beliefs about back pain and had a positive effect on outcomes [23].

Thus we based the active management package used in the trial on the RCGP guidelines [3,4] and *The Back Book* [21]. With the support of general practitioners and their staff (practice nurses and non-clinical support staff) it aimed to achieve early activity and allay fears about future consequences of back trouble. The message concentrated on minimising bed rest, promoting ordinary physical activity, remaining or returning to work, and encouraging positive attitudes to pain.

We actively sought the support of all relevant staff for this package since changes in clinical management may be successful when the entire clinical team delivers the innovation in a consistent manner [24]. To this end we issued all general practitioners and nursing staff in trial practices with copies of the RCGP evidence base [4] and clinical guidelines. To reinforce the theory, principles, and practice of active management we invited all members of participating practice teams to a training session. We also trained them in distributing and emphasising the messages in *The Back Book* to patients with back pain. *The Back Book* complements the clinical guidelines and thus reinforces the information and advice given by practitioners and other primary care staff.

We encouraged practitioners to limit prescriptions to a small range of drugs including analgesics (paracetamol with or without codeine) and non-steroidal anti-inflammatory agents (diclofenac or ibuprofen or naproxen).

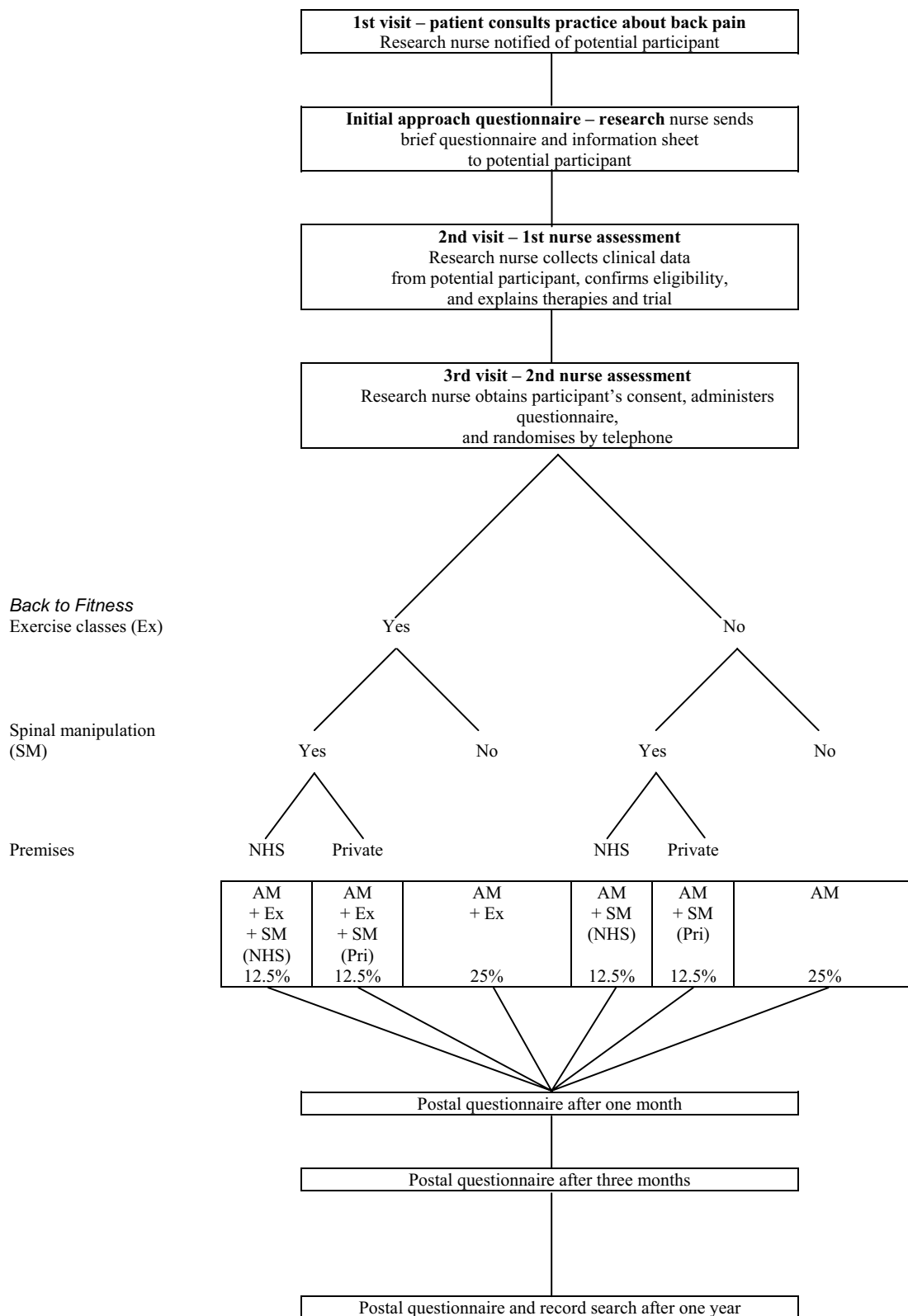


Figure 1
Trial design

However, adherence to this guidance is not crucial since there is little evidence that analgesics differ in their effectiveness in back pain. Adherence was assessed by collecting data on other treatments used by participants from participant questionnaires and examination of the medical record.

We recruited general practitioners to the trial only if they agreed not to perform spinal manipulation or manual therapy on trial participants. We discouraged them from referring participants to other physical therapists or other specialists especially during the initial three months of treatment although not during the remainder of the follow up period.

We originally planned to compare the effect of training practice teams in the active management of low back pain with conventional care in general practice as an additional layer of our factorial design. Thus in the feasibility study we randomised 26 general practices in two centres (West Yorkshire and Stockport) between an active management training programme *and* conventional general practice management. For the active management comparison we randomised by practice rather than individual because of evidence from the guideline literature that the chances of successful implementation are greater when an entire clinical team (or practice) delivers an innovation coherently and consistently [24]. Practice-based randomisation also circumvents the possibility of 'contamination' between the active intervention and usual management.

However, the active management practices recruited more than twice as many participants (165) as usual management practices (66). Furthermore these additional participants reported a milder form of back pain. Thus continuation with this comparison would have made the other trial comparisons difficult to interpret and could have adversely affected recruitment to the main trial. With the agreement of the MRC Trial Steering Committee, we modified the study design of the main trial. For the main study we provided the active management training package to staff from all practices.

In light of the existing evidence, the wide distribution of the acute back pain treatment guidelines, and the increasing popularity of the active management philosophy, we judged that active management represented 'best care in general practice'. Additionally this would improve recruitment and standardise the 'control' treatment. We therefore chose this approach to the management of low back pain as our 'control' treatment.

Manipulative package

Different practitioners have practised spinal manipulation for many years. In the UK, Acts of Parliament in the

mid-1990s formally recognised chiropractic and osteopathy. At the same time, the physiotherapy profession has been increasingly utilising manual therapy. In the UK, similarities between chiropractors, osteopaths and some physiotherapists seem to be much greater than their differences. The MRC Working Party therefore agreed that to compare the effectiveness of the three disciplines was unnecessary, and probably impossible, in view of the large sample size needed.

There is evidence that people with back pain may derive modest benefit from spinal manipulation. In the UK, at the time the trial was designed, the only substantial trial of manipulation was in secondary care, suggesting there may be some longer term benefit [25,26]. However, the design of the trial was criticised because it compared NHS outpatient care with treatment in private chiropractic premises [27] raising the possibility that the location and style of treatment might explain the differences that were found.

Members of the MRC Working Party developed a package in collaboration with the three physical therapy professions involved (chiropractic, osteopathy and physiotherapy) who agreed to use it in this trial. It defines a common core of manipulative practice while permitting enough flexibility in both assessment and treatment to be representative of all three professions. The package, described in detail elsewhere [28], comprised the commonly used range of manual treatment techniques together with non-manual elements usually used by physical therapists (for example exercises and hot/cold packs); oral advice was to be consistent with *The Back Book* [21].

For individual participants the package comprised a series of scheduled sessions with the same therapist. The maximum of eight sessions could be spread across the intervention period of 12 weeks (six weeks for participants randomised to both manipulation and exercise) at the discretion of the practitioner. We provided a detailed manual to all therapists delivering manipulation treatments as part of the trial. They also attended familiarisation sessions and were asked to commit to the 'active' philosophy of the trial.

Exercise package

Physiotherapy is widely and increasingly available in UK primary care. The evidence to support the use of any individual physiotherapy technique in those with back pain is scant. There is some weak evidence to support the use of intensive extensor strengthening exercises [29], [30] and for group education or 'back schools' [31,32]. There is better evidence for the effectiveness of a progressive exercise programme based on cognitive behavioural principles [14,33-35] and for the benefits of exercise programmes

aimed at endurance, muscular strength and normal use of the back [5,33,29,34–37]. Building on previous small UK studies [34,35], members of the MRC working party developed the *Back to Fitness* exercise programme. The aims of the programme were to:

- encourage normal movement;
- increase participants' confidence in their spines; and
- help participants take control of their problem.

This trial will help to confirm whether:

1. This type of exercise programme is effective in the UK NHS, as much of the evidence is from studies in Scandinavia and the US;
2. An exercise programme is effective among patients recruited from primary care, as most studies have either been in hospital or in the work place; and
3. The '*Back to Fitness*' programme is effective when implemented nationally.

The *Back to Fitness* programme is based on the principles of circuit training and incorporates a cognitive-behavioural approach. The programme is described in detail elsewhere [38]. The exercise classes took place in community settings accessible to patients from participating general practices, usually in the early evening. Each class had up to ten participants. If necessary to maintain group size, additional non-trial participants were recruited to the classes. The classes lasted one hour and took place twice a week over a four-week period. Those who found this schedule difficult could spread their classes over a period of, at most, 12 weeks (six weeks for participants randomised to both manipulation and exercise). Before joining, each participant in the exercise class had an individual assessment by a physiotherapist. We asked the physiotherapists to provide advice consistent with messages in *The Back Book* [21]. Participating physiotherapists attended an intensive study day led by a multidisciplinary team including a behavioural scientist specialising in this field and therapists with experience of similar programmes.

Combination of therapies

If the trial showed that the manipulative package [28] and the *Back to Fitness* programme [38] were both effective, those who received a combination of these two might experience an average health gain greater or less than the sum of the average health gains attributable to each individual therapy. For example, theoretically the combination of manipulation and exercise could be especially

beneficial if the exercise maintained any new range of movement gained through manipulation [39]. Thus it is important to test whether there are any positive or negative interactions between the two treatments. Those participants randomised to receive both spinal manipulation and exercise attended the manipulator exclusively for the first six weeks and the exercise programme for up to another six weeks. Thus all treatment was complete within the specified 12-week period. This arrangement meant that participants were not under the clinical management of different therapists at the same time. However, both treatments occupied a shorter period in the combined arm than in the other two intervention groups. Analysis will test whether participants receiving the combined package experience an average health gain significantly less or significantly greater than the sum of the average health gains attributable to each individual therapy.

Inclusion and exclusion criteria

The target population was people aged between 18 and 65 years presenting in general practice with non-specific back pain with or without referred leg pain. We defined back pain as pain of musculoskeletal origin in the area bounded by the lowest palpable ribs, the gluteal folds, and the posterior axillary lines. We included those with pain referred into the legs provided it was predominantly above the knee. To avoid carry-over effects, participants should not have had physical therapy in the previous three months. To facilitate participation in the exercise class and assessment they had to be fluent in English and able to read and write.

To avoid including participants who recovered rapidly from an acute episode without specific treatment, we included those whose current episode of back pain had lasted at least four weeks from their initial consultation. They did not have to have had pain every day of this episode. They were eligible if they had had back pain either for 28 consecutive days, or, for at least 21 of these days and at least 21 of the 28 days before that. No upper limit to the duration of pain was set. Exclusion criteria are summarised in Figure 2.

Serious adverse events

We defined a serious adverse event as one leading to hospital admission or death within one week of treatment within the trial. To ensure that we recognised any relevant events the manuals describing spinal manipulation and exercise classes each asked participating health professionals to report and follow up any potential events. Additionally, the follow-up questionnaire after three months asked participants whether they had been admitted to hospital, and practice research nurses searched participants' records for hospital discharge summaries.

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[2] [5].
- There was a possibility of serious spinal disorder, including malignancy, osteoporosis, ankylosing spondylitis, cauda equina compression and infection [2].
- 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 [3].
- They had another musculo-skeletal 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, for example cardiovascular disease, which could interfere with therapy.
- They had moderate to severe hypertension (systolic blood pressure greater than 180 mm Hg, or diastolic blood pressure greater than 105 mm Hg, on at least two separate occasions).
- They were taking anti-coagulant therapy.
- They were taking long-term steroids, which may lead to osteoporosis.
- They could not walk 100 metres when free of back pain, because exercise would be difficult.
- They could not get up and down off the floor unaided.
- They had received physical therapy (including acupuncture) in the previous three months.
- They had an RDQ score [41] of three or less on the day of randomisation.
- They could not read and write fluently in English.

Figure 2
Exclusion criteria

Participant recruitment and allocation to treatments

General practitioners and other practice team members notified the practice research nurse of patients consulting for back pain. In practices that computerised their consultation data the nurses also did regular computer searches for back pain consultations. Potential participants were sent an invitation letter, an information sheet, and a brief

questionnaire covering the main inclusion and exclusion criteria.

Those interested in the trial made an initial appointment with the nurse. At this appointment the nurse confirmed eligibility by collecting further data. She asked about employment status and how the current back pain was affecting them and explained the treatment packages

available through the trial. Data on those who did not join the study will allow us to generalise our findings.

Those who appeared eligible and interested in participating in the trial saw the nurse again at least one week later. This qualifying week gave them time to reflect on the implications of taking part in the trial [40]. It also allowed the general practitioner to confirm that they appeared suitable for the trial. Randomisation took place at least four weeks after the initial consultation with the practice, to exclude anyone who made a rapid spontaneous recovery. Those who no longer met the severity entry criterion, a Roland Morris Disability Questionnaire (RDQ) [41] score of four or more could have become eligible if their condition subsequently deteriorated without needing to go through the waiting period.

Following consent, participants completed the main baseline questionnaire. The research nurse then telephoned the York randomisation service to obtain the participant's random treatment allocation. Randomisation was stratified by practice.

Outcome measures

The main outcome measures were self report questionnaires asking about participants' health, beliefs about back pain, and psychological profile. After recruitment, they received similar follow-up questionnaires by post after one month, three months (when all treatment within the trial has finished) and one year. Non-responders received written reminders after two and four weeks, the second by recorded delivery. Practice research nurses also prompted non-responders by phone at the time of the second reminder.

Health outcomes

The internationally agreed minimum data set for back pain research informed our selection of outcome measures [42]. We used two generic health outcome measures: the Short Form 36-item Health Survey (SF-36) [43], a valid and reliable measure that is responsive to changes in the health of people presenting to primary care with low back pain [40,44]; and the EuroQol (EQ5D), which produces a single index based on health state valuations [45]. The primary outcome measure was the Roland Morris Disability Questionnaire (RDQ) [41], an established outcome measure for community based back pain studies [46]. We shall use it in its original form for our main analyses. However, we added four alternative items, that Patrick *et al.* [47] suggested, improved the performance of the RDQ for assessing the health status of patients with sciatica.

The two generic and one specific measure were complemented by:

1. a general measure of pain and disability, derived from the chronic pain grade proposed by Von Korff *et al.* [48]. This has been shown to be acceptable, valid, and reliable when modified to measure back problems over a one month period in a postal questionnaire in the United Kingdom [49];

2. a general question about the perceived global effect of the condition [50] (by analogy with the 'generic health transition' question of the SF-36 health survey [43], we describe it as 'specific health transition').

Back pain beliefs

We used the Fear Avoidance Beliefs Questionnaire (FABQ) [51] and the Back Beliefs Questionnaire (BBQ) [52] to measure participants' beliefs about back pain at recruitment and follow-up. The FABQ comprises 16 items that measure beliefs about the effect of physical activity and work on low back pain. The BBQ comprises 14 items that measure beliefs about the consequences of low back pain. There is evidence for the validity and reliability of both instruments [51,52]. The BBQ is also responsive to changes in beliefs about back pain [53]. However, we excluded items from both measures. Low response rates in the feasibility study led us to drop the work component of the FABQ and base our analyses on the FABQ-physical scale only [51]. To reduce respondent burden we used only the nine items forming the 'inevitability' sub-scale of the BBQ [52].

Psychological instruments

Psychological instruments may predict the outcome of episodes of back pain in primary care [54,55]. Psychological distress at recruitment was assessed by the Distress and Risk Assessment Method (DRAM) [56]. This consists of 55 items to categorise people into four types – normal, at risk, distressed depressive, and distressed somatic. However, only 36 of these items contribute to the scale scores for its components, the Modified Somatic Perception Questionnaire [57] and Modified Zung [58]. To reduce respondent burden we included only those items that are scored. It is valid for use by those with low back pain.

Table 1 summarises the resulting portfolio of measures.

Health economic outcomes

The economic design for this trial is that of an incremental cost-utility analysis from the perspective of society as a whole. The main focus of the economic evaluation is the EuroQol – a generic measure of health utility [59]. However we shall also analyse the other outcomes in Table 1 to check whether choice of outcome measure affects the economic comparison of treatment packages.

Table 1: Measures of health outcome, back pain beliefs and psychological profile

Measure	Authors	Items	Initial	I/12	3/12	12/12
SF-36 health survey	Ware & Sherbourne [43]	36	Yes	Yes	Yes	Yes
EuroQol	EuroQol Group [59]	6	Yes	Yes	Yes	Yes
Roland Morris Disability Questionnaire	Roland & Morris [41] Patrick <i>et al.</i> [47]	24 + 4	Yes	Yes	Yes	Yes
Fear Avoidance Beliefs Questionnaire	Waddell <i>et al.</i> [51]	4	Yes	Yes	Yes	Yes
Back Beliefs Questionnaire	Symonds <i>et al.</i> [64]	9	Yes	Yes	Yes	Yes
Modified Von Korff	Von Korff <i>et al.</i> [48] Underwood <i>et al.</i> [49]	6	Yes	Yes	Yes	Yes
'Troublesomeness' etc	Deyo [65]	6	Yes	Yes	Yes	Yes
Specific health transition	Beurskens <i>et al.</i> [50]	1	Yes	Yes	Yes	Yes
DRAM	Main <i>et al.</i> [56]	36	Yes	No	No	No

The other main economic benefit is the extent to which each of the (combinations of) treatment packages reduces subsequent use of health care over 12 months. Given the difficulty of disentangling back-specific from other costs, it will estimate generic health care costs including hospital admissions and referrals, and consultations in primary care. Practice research nurses abstracted data from their records at the end of data collection.

We estimated the costs of receiving different packages in the form of time off work and normal activities and the costs of travel in both time and money. The feasibility study tested the most reliable means of collecting these data through an embedded randomised trial comparing simple prospective diaries and retrospective questions within the main outcome questionnaires. As diaries achieved lower response rates, we used retrospective questions in the main trial. These data from participants serve to validate, and elaborate upon, the estimates of health care use derived from practice records.

We collected data on the cost of each package including all fixed and variable costs incurred in delivering them. Each therapist and practitioner recorded direct costs on a simple form with an estimate of the time associated with each consultation including non-contact time. We valued these costs in terms of both opportunity costs to individual practices, and national averages, for greater generalisability. We shall estimate the cost of training for the *Back to Fitness* exercise package and include it in our evaluation of the widespread implementation of the packages in the trial. We shall also estimate the marginal cost of extending each package to individual practice populations.

Trial centres

The main trial recruited participants from practices that were part of the Medical Research Council General Prac-

tice Research Framework (GPRF) <http://www.mrc-gprf.ac.uk/> in 12 centres across the UK – Belfast, Edinburgh & Tayside, Exeter, Harrow, Northampton, Norwich, Nottingham, Plymouth, Reading, Sheffield, Teesside and Wrexham & Chester. The aim was for each area to provide an average population of 83,000 patients registered with ten to 15 general practices. The practices and patients needed to be within easy travelling distance of the set of (at least) three premises where physical therapy was available – one in the community for the *Back to Fitness* exercise package and at least two for spinal manipulation, one within and one outside the NHS. In each centre a part-time administrative 'local co-ordinator' supported by a senior 'clinical adviser', had responsibility for organising the delivery of physical therapies.

Sample size

The primary outcome measure for this trial is the RDQ [41] used by many community-based back pain researchers in the UK and elsewhere. It is therefore the appropriate criterion by which to estimate the target sample size for the trial.

The MRC Working Party, originally responsible for the trial design, considered a difference of 2.5 points in the RDQ to be clinically important when comparing the manipulation or exercise packages with general practice care. A previous study [35], and baseline data from our feasibility study, found the standard deviation for the RDQ for participants recruited from primary care to be around 4.0.

However the crucial comparison in estimating sample size is between manipulation in NHS and private premises (hypothesis 'b'). There would be much less value in testing hypothesis 'b' if a prior test had already established that there was no clinically important difference between

Table 2: Power of trial to detect differences across main comparisons

Power to detect difference of 2.5 RDQ points at 1% significance level	No interaction present between manipulation and exercise	Interaction present between manipulation and exercise	Analysing only completers ^a (no interactions present)
Exercise vs. AM only	>99%	>99%	>99%
Manipulation vs. AM only	>99%	>99%	>99%
Site of manipulation ^b	80%	42%	64%
Site of manipulation ^c	91%	66%	83%

^a Based on estimated completion rate of 75% ^b Power to detect difference of 1.67 RDQ points at 1% significance level ^c Power to detect difference of 1.67 RDQ points at 5% significance level

spinal manipulation in general and general practice care. Thus hypothesis 'b' is most important when the test of hypothesis 'a' has established that spinal manipulation participants have average RDQ scores that are at least 2.5 points lower than participants just receiving general practice care. Furthermore the test of hypothesis b is most in need of statistical power when this difference is exactly 2.5 points. It follows that the alternative hypothesis to hypothesis 'b' is one in which spinal manipulation participants in one location have average RDQ scores that are 3.33 lower than general practice participants, while those in the other location have average RDQ scores that are 1.67 lower than general practice participants. Thus for hypothesis 'b' we defined a clinically important difference as 1.67 points on the RDQ or 0.417 of the standard deviation of four points. To yield a power of 80 per cent of detecting such a difference using a *one* per cent significance level would need one year follow up data on two simple random sub-samples of 130. Since half the participants did not receive the manipulation package, a sample size of 520 was needed.

Embedded within this population there are three types of clusters – clusters of participants from the same general practice, clusters of participants attending the same exercise class, and clusters of participants receiving spinal manipulation from the same therapist. It follows that the power of each of the three dimensions of this trial would depend on the relevant cluster sizes and intra-cluster correlation coefficients [60]. Using existing published estimates [35,61–63] and data from our feasibility study we estimated the sample size inflation required to account for clustering effects at practice, exercise class, and manipulator levels. The largest inflation factor is required to account for intra-manipulator effects. In our feasibility study the intra-manipulator correlation coefficients were 0.04. In each centre in the main study there were two manipulators, each seeing patients in NHS and private premises, and thus each treating 25% of the sample in that centre. If the average cluster size were 19 (that is one quarter of the mean sample size of 75 per centre), we needed

to inflate the sample size by a factor of 1.72 {viz. $1 + (19 - 1) \times 0.04$ } [60].

We planned to have data from 900 participants after one year, slightly more than 520×1.72 . These sample sizes yield a power of at least 99% of detecting a difference of 2.5 RDQ points between either exercise only or manipulation only and active management using a *one* per cent significance level. This would allow us to estimate the main effects of exercise or manipulation, even if we were to find significant interactions, or if the active management (best care) training package were to reduce the baseline scores of participants and give less scope for therapists to improve RDQ scores through manipulation or exercise (Table 2).

Allowing for one third loss to follow up at one year we sought to recruit 1,350 participants. In our feasibility study active management practices recruited 1.62 participants per 1000 registered patients. Assuming recruitment rate of 1.5/1,000 in the main study, in which all practice received the active management training package, we required practices with 900,000 registered patients for the main study.

Statistical analysis

The analysis of the main trial will take account of the 3 × 2 factorial design. The primary analysis will be by intention to treat. It will include all participants properly randomised, even if they did not subsequently receive the package allocated, or if they received elements of other packages. Proposed secondary analyses will focus on participants who received the essentials of the package to which they were allocated. All significance tests will be two-sided. The analysis will take into account the existence of three natural hierarchies – participants *within* practices, *or* therapists (for spinal manipulation if allocated), *or* community settings (for progressive exercise if allocated) *within* centres.

The primary outcome measure is the RDQ score at three and 12 months. We shall calculate the change from baseline for each participant and use this as the dependent variable. The baseline RDQ score is a potential covariate in this analysis. The secondary outcome measures are: the RDQ at one month; and, at one, three and 12 months, the SF-36; EuroQol; FABQ; BBQ; Modified Von Korff, Deyo's 'Bothersomeness' scores (anglicised to 'Troublesomeness'); and a specific health transition question (Table 1).

If interactions are not statistically significant, the analysis will present the main effects of treatments, namely NHS manipulation, private manipulation and exercise. As significant interactions are possible however, we may need to estimate the effects of the six distinct combinations of treatments, for example, NHS manipulation and exercise. Thus we plan to build a model by first estimating main effects and then testing for interactions. This differs from a traditional factorial analysis that assumes that interactions are absent or tests this immediately.

First, we shall estimate the main effect of exercise by comparing the exercise only group with the AM group. Second we shall estimate the main effect of manipulation by comparing the manipulation only group with the AM group. Third we shall estimate the differences between NHS and private manipulation within the manipulation only group. Then we shall investigate interactions between exercise and manipulation by examining the group allocated to this combination. Multi-level modelling will take account of variation between practices. It can also estimate the effects of centres, manipulators, and exercise physiotherapists, albeit with wide confidence intervals.

Missing data will be treated in two ways. Missing items within individual outcome measures will be treated according to the instructions for that measure. Then we shall calculate response rates for each treatment group. If these differ we shall compare responses to three distinct mailings (namely first questionnaires, and first and second reminders) and adjust accordingly.

If some or all treatments are effective, secondary analysis will examine whether variations in the process of care can explain variations in outcome. For example, we shall test whether variation in the numbers of exercise and manipulation sessions, and in the manipulation techniques used, affect primary and secondary outcomes. Further secondary analyses will investigate potential prognostic variables collected at the first and second appointments with the practice research nurse.

Conclusion

The UK BEAM trial is a major trial of physical treatments for low back pain. Obtaining participation by members of

the three physical therapy professions in the UK (chiropractic, osteopathy and physiotherapy) to work to an agreed treatment was an important achievement. Whatever the outcome of the trial, the results will inform the future management of low back pain both within the UK and internationally. Participant recruitment and follow up is now complete. We recruited 1334 participants from 168 practices. With agreement from the Trial Steering Committee we included data from participants recruited by the 13 practices within the 'active management' arm of the feasibility study, thus adding two centres and making 14 in all. We obtained 12 month follow up data on 995 (75%) of all participants these. This provides ample data to test all our main hypotheses.

Authors' contributions

Full details of all contributors to UK BEAM are available at: <http://www.york.ac.uk/healthsciences/centres/trials/ukbeam/contrib.htm>.

Acknowledgments

Sponsors: Research Costs, Medical Research Council. Treatment costs, NHS Research and Development. Underwood is supported by an NHS research and development primary care career scientist award.

Trial Steering Committee: Assendelft, W, Croft P, Hughes R (chair), Klaber Moffett J, Lamb S, Russell I, Thompson S, Underwood M, Vickers M, Watt G, Vickers M, Watt G,

MRC working party: Beurskens A, Breen A, Burton K, Frost H, Garratt A, Godfrey C, van der Heijden G, Klaber Moffett J, Little P, Main C, Meade T, Roland M (chair), Russell I, Smith I, Waddell G, Underwood M, Vickers M.

References

1. Maniadas N and Gray A: **The economic burden of back pain in the UK.** *Pain* 2000, **84**:95-103.
2. CSAG: **Back Pain. Report of a CSAG committee on Back Pain.** London, HMSO 1994.
3. Waddell G, Feder G, McIntosh G, Lewis M and Hutchinson A.: **Low Back Pain Evidence Review.** London, Royal College of General Practitioners 1996.
4. Waddell G, Feder G, McIntosh G, Lewis M and Hutchinson A.: **Low Back Pain Evidence Review.** London, Royal College of General Practitioners 1999.
5. Agency for Health Care Policy and Research [AHCPR]: **Acute low back problems in adults.** *Clinical Practice Guideline 14.* Rockville, Maryland, US Department of Health and Human Services 1994.
6. Faas A, Chavannes AW and Koes BW *et al.*: **NHG-standaard lage-rugpin.** *huisarts wet* 1996, **39**:18-31.
7. Koes BW, van Tulder MW, Ostelo R, Burton AK and Waddell G: **Clinical Guidelines for the Management of Low Back Pain in Primary Care.** *Spine* 2001, **26**:2504-2514.
8. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM and Knipschild PG: **Spinal manipulation and mobilisation for back and neck pain: a blinded review.** *BMJ* 1991, **303**:1298-1303.
9. Shekelle PG, Adams AH, Chassin MR, Hurwitz EL and Brook RH: **Spinal manipulation for low-back pain.** *Ann Intern Med* 1992, **117**:590-598.
10. Assendelft WJ, Koes BW, Knipschild PG and Bouter LM: **The relationship between methodological quality and conclusions in reviews of spinal manipulation.** *JAMA* 1995, **274**:1942-1948.
11. Assendelft WJ, Morton SC, Yu EI, Suttrop MJ and Shekelle PG: **Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies.** *Ann Intern Med* 2003, **138**:871-881.

12. Koes BW, Bouter LM, Beckerman H, van der Heijden GJ and Knipschild PG: **Physiotherapy exercises and back pain: a blinded review.** *BMJ* 1991, **302**:1572-1576.
13. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE and Fordyce WE et al.: **The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach.** *Phys Ther* 1992, **72**:279-290.
14. Faas A: **Exercises: which ones are worth trying, for which patients, and when?** *Spine* 1996, **21**:2874-2878.
15. van Tulder MW, Malmivaara A, Esmail R and Koes BW: **Exercise therapy for low back pain (Cochrane Review).** In: *The Cochrane Library Issue 1 Oxford: Update Software*; 2003.
16. Waddell G, Feder G and Lewis M: **Systematic reviews of bed rest and advice to stay active for acute low back pain.** *Br J Gen Pract* 1997, **47**:647-652.
17. Hagen KB, Hilde G, Jamtvedt G and Winnem MF: **The cochrane review of advice to stay active as a single treatment for low back pain and sciatica.** *Spine* 2002, **27**:1736-1741.
18. Hagen KB, Hilde G, Jamtvedt G and Winnem MF: **The Cochrane review of bed rest for acute low back pain and sciatica.** *Spine* 2000, **25**:2932-2939.
19. Altman DG: **Practical statistics for medical research.** London: Chapman and Hall 1991.
20. Underwood M, O'Meara S and Harvey E: **The acceptability to primary care staff of a multidisciplinary training package on acute back pain guidelines.** *Fam Pract* 2002, **19**:511-515.
21. Roland M, Waddell G, Klüber Moffett J, Burton AK, Main CJ and Cantrell T: **The Back Book.** Norwich: The Stationery Office 1996.
22. Burton AK, Waddell G, Burt R and Blair S: **Patient educational material in the management of low back pain in primary care.** *Bulletin - Hospital for Joint Diseases* 1996, **55**:138-141.
23. Burton AK, Waddell G, Tillotson KM and Summerton N: **Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care.** *Spine* 1999, **24**:2484-2491.
24. Grimshaw JM and Russell IT: **Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations.** *Lancet* 1993, **342**:1317-1322.
25. Meade TW, Dyer S, Browne W, Townsend J and Frank AO: **Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment.** *BMJ* 1990, **300**:1431-1437.
26. Meade TW, Dyer S, Browne W and Frank AO: **Randomised comparison of chiropractic and hospital outpatient management for low back pain: results from extended follow up.** *BMJ* 1995, **311**:349-351.
27. Assendelft WJ, Bouter LM and Kessels AG: **Effectiveness of chiropractic and physiotherapy in the treatment of low back pain: a critical discussion of the British Randomized Clinical Trial.** *Journal of Manipulative & Physiological Therapeutics* 1991, **14**:281-286.
28. Harvey E, Burton AK, Moffett JK and Breen A: **Spinal manipulation for low-back pain: a treatment package agreed to by the UK chiropractic, osteopathy and physiotherapy professional associations.** *Man Ther* 2003, **8**:46-51.
29. Manniche C, Lundberg E, Christensen I, Bentzen L and Hesselsoe G: **Intensive dynamic back exercises for chronic low back pain: a clinical trial.** *Pain* 1991, **47**:53-63.
30. Evans G and Richards SH: **Low back pain: An evaluation of therapeutic interventions.** Bristol, Health Care Evaluation Unit 1996.
31. Cohen JE, Goel V, Frank JW, Bombardier C, Peloso P and Guillemin F: **Group education interventions for people with low back pain. An overview of the literature.** *Spine* 1994, **19**:1214-1222.
32. Koes BW, van Tulder MW, van der Windt WM and Bouter LM: **The efficacy of back schools: a review of randomized clinical trials.** *J Clin Epidemiol* 1994, **47**:851-862.
33. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE and Nachemson A: **Mobility, strength, and fitness after a graded activity program for patients with subacute low back pain. A randomized prospective clinical study with a behavioral therapy approach.** *Spine* 1992, **17**:641-652.
34. Frost H, Klüber Moffett J, Moser JS and Fairbank JC: **Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain.** *BMJ* 1995, **310**:151-154.
35. Klüber Moffett J, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H and Farrin A et al.: **Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences.** *BMJ* 1999, **319**:279-283.
36. Kellett KM, Kellett DA and Nordholm LA: **Effects of an exercise program on sick leave due to back pain.** *Phys Ther* 1991, **71**:283-291.
37. Gundewall B, Liljeqvist M and Hansson T: **Primary prevention of back symptoms and absence from work. A prospective randomized study among hospital employees.** *Spine* 1993, **18**:587-594.
38. Klüber Moffett J and Frost H: **Back to Fitness Programme.** *Physiotherapy* 2000, **86**:295-305.
39. Erhard RE, Delitto A and Cibulka MT: **Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome.** *Phys Ther* 1994, **74**:1093-1100.
40. Knipschild P, Leffers P and Feinstein AR: **The qualification period.** *J Clin Epidemiol* 1991, **44**:461-464.
41. Roland M and Morris R: **A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain.** *Spine* 1983, **8**:141-144.
42. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P and Koes B et al.: **Outcome measures for low back pain research. A proposal for standardized use.** *Spine* 1998, **23**:2003-2013.
43. Ware JE and Sherbourne CD: **The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.** *Med Care* 1992, **30**:473-483.
44. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK and Russell IT: **The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS?** *BMJ* 1993, **306**:1440-1444.
45. Ruta DA, Garratt AM, Wardlaw D and Russell IT: **Developing a valid and reliable measure of health outcome for patients with low back pain.** *Spine* 1994, **19**:1887-1896.
46. Roland M and Fairbank J: **The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire.** *Spine* 2000, **25**:3115-3124.
47. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A and Keller RB: **Assessing health-related quality of life in patients with sciatica.** *Spine* 1995, **20**:1899-1908.
48. Von Korff M, Ormel J, Keefe FJ and Dworkin SF: **Grading the severity of chronic pain.** *Pain* 1992, **50**:133-149.
49. Underwood MR, Barnett AG and Vickers MR: **Evaluation of two time-specific back pain outcome measures.** *Spine* 1999, **24**:1104-1112.
50. Beurskens AJ, de Vet HC and Koke AJ: **Responsiveness of functional status in low back pain: a comparison of different instruments.** *Pain* 1996, **65**:71-76.
51. Waddell G, Newton M, Henderson I, Somerville D and 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**:157-168.
52. Symonds TL, Burton AK, Tillotson KM and Main CJ: **Do attitudes and beliefs influence work loss due to low back trouble?** *Occup Med (Oxf)* 1996, **46**:25-32.
53. Buchbinder R, Jolley D and Wyatt M: **2001 Volvo Award Winner in Clinical Studies: Effects of a media campaign on back pain beliefs and its potential influence on management of low back pain in general practice.** *Spine* 2001, **26**:2535-2542.
54. Klenerman L, Slade PD, Stanley IM, Pennie B, Reilly JP and Atchison LE et al.: **The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting.** *Spine* 1995, **20**:478-484.
55. Burton AK, Tillotson KM, Main CJ and Hollis S: **Psychosocial predictors of outcome in acute and subchronic low back trouble.** *Spine* 1995, **20**:722-728.
56. Main CJ, Wood PL, Hollis S, Spanswick CC and Waddell G: **The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome.** *Spine* 1992, **17**:42-52.
57. Main CJ: **The Modified Somatic Perception Questionnaire (MSPQ).** *J Psychosom Res* 1983, **27**:503-514.
58. Main CJ and Waddell G: **The detection of psychological abnormality in chronic low-back pain using four simple scales.** *Curr Concepts Pain* 1984, **2**:10-15.
59. The EuroQol Group: **EuroQol - a new facility for the measurement of health-related quality of life.** *Health Policy* 1990, **16**:199-208.

60. Donner A, Birkett N and Buck C: **Randomization by cluster. Sample size requirements and analysis.** *Am J Epidemiol* 1981, **114**:906-914.
61. Medical audit in general practice. II: **Effects on health of patients with common childhood conditions. North of England Study of Standards and Performance in General Practice.** *BMJ* 1992, **304**:1484-1488.
62. Wilkinson CE, Peters TJ, Stott NC and Harvey IM: **Prospective evaluation of a risk scoring system for cervical neoplasia in primary care.** *Br J Gen Pract* 1994, **44**:341-344.
63. McKinley RK, Cragg DK, Hastings AM, French DP, Manku-Scott TK and Campbell SM et al.: **Comparison of out of hours care provided by patients' own general practitioners and commercial deputising services: a randomised controlled trial. II: The outcome of care.** *BMJ* 1997, **314**:190-193.
64. Symonds TL, Burton AK, Tillotson KM and Main CJ: **Absence resulting from low back trouble can be reduced by psychosocial intervention at the work place.** *Spine* 1995, **20**:2738-2745.
65. Deyo RA and Centor RM: **Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance.** *J Chronic Dis* 1986, **39**:897-906.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1472-6963/3/16/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

