Research Report

Implementation of Clinical Guidelines on Physical Therapy for Patients With Low Back Pain: Randomized Trial Comparing Patient Outcomes After a Standard and Active Implementation Strategy

Background and Purpose. An active strategy was developed for the implementation of the clinical guidelines on physical therapy for patients with low back pain. The effect of this strategy on patients' physical functioning, coping strategy, and beliefs regarding their low back pain was studied. Subjects. One hundred thirteen primary care physical therapists treated a total of 500 patients. Methods. The physical therapists were randomly assigned to 1 of 2 groups. The control group received the guidelines by mail (standard passive method of dissemination). The intervention group, in contrast, received an additional active training strategy consisting of 2 sessions with education, group discussion, role playing, feedback, and reminders. Patients with low back pain, treated by the participating therapists, completed questionnaires on physical functioning, pain, sick leave, coping, and beliefs. Results. Physical functioning and pain in the 2 groups improved substantially in the first 12 weeks. Multilevel longitudinal analysis showed no differences between the 2 groups on any outcome measure during follow-up. Discussion and Conclusion. The authors found no additional benefit to applying an active strategy to implement the physical therapy guidelines for patients with low back pain. Active implementation strategies are not recommended if patient outcomes are to be improved. [Bekkering GE, van Tulder MW, Hendriks HJM, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. *Phys Ther.* 2005;85:544–555.]

Bone and Joint DECADE AMERICA 2000-2010

APTA is a sponsor of the Decade, an international, multidisciplinary initiative to improve health-related quality of life for people with musculoskeletal disorders.

Key Words: *Cluster-randomized controlled trial, Implementation, Low back pain, Patient outcomes, Physical therapy, Practice guidelines.*

Geertruida E Bekkering, Maurits W van Tulder, Erik JM Hendriks, Marc A Koopmanschap, Dirk L Knol, Lex M Bouter, Rob AB Oostendorp

Low back pain is experienced by at least 80% of adults[1] and is associated with substantial socioeconomic and health care costs. There are a variety of approaches to the management of patients with low back pain, of which physical therapy is one option. Within physical therapy, there is a growing body of scientific evidence with regard to the effectiveness and efficacy of commonly used interventions.[2] For the management of patients with low back pain, commonly used physical therapy interventions are exercise, advice, manual techniques, and electrotherapeutic or thermal modalities.[3–5] Only some of these interventions have been shown to be effective. This fact suggests that the dissemination of findings regarding the effectiveness of specific physical therapy interventions for patients with low back pain needs to be improved.

To assist physical therapists in the Netherlands in using evidence-based interventions in practice, the KNGF (Royal Dutch Society for Physical Therapy) has issued physical therapy guidelines for patients with low back pain.[6] Worldwide, these are the first guidelines on low back pain specifically developed for physical therapists.[7]

The Dutch physical therapy guidelines were developed by a group of physical therapists and researchers in the field of low back pain. Wherever possible, the recommendations were based on scientific evidence. The concept of low back pain in these guidelines refers to *nonspecific low back pain*, defined as low back pain without a specified physical cause (eg, nerve root compression, trauma, infection, tumor).[8] In about 85% of patients with low back pain, no specific medical diagnosis is

GE Bekkering, PhD, is Researcher, Dutch Institute of Allied Health Care, Amersfoort, the Netherlands, and Researcher, Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, the Netherlands. Address all correspondence to Dr Bekkering at Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol BS8 2PR United Kingdom (G.E.Bekkering@bristol.ac.uk).

MW van Tulder, PhD, is Associate Professor, Health Technology Assessment, Institute for Research in Extramural Medicine, VU University Medical Center, and Associate Professor, Health Technology Assessment, Institute for Health Sciences, Faculty of Earth and Life Sciences, Free University, Amsterdam, the Netherlands.

EJM Hendriks, PhD, is Physical Therapist, Health Scientist, and Epidemiologist, Dutch Institute of Allied Health Care, Amersfoort, the Netherlands; Senior Researcher, Department of Epidemiology, Maastricht University; and Co-director, Center of Evidence Based Physiotherapy (CEBP), Maastricht, the Netherlands.

MA Koopmanschap, PhD, is Senior Health Economist, Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, the Netherlands.

DL Knol, PhD, is Statistician, Institute for Research in Extramural Medicine, VU University Medical Center, and Statistician, Department of Clinical Epidemiology and Biostatistics, VU University Medical Center.

LM Bouter, PhD, is Professor of Epidemiology and Scientific Director, Institute for Research in Extramural Medicine, VU University Medical Center.

RAB Oostendorp, PhD, is Physical Therapist, Manual Therapist, and Research Director, Dutch Institute of Allied Health Care; Professor of Allied Health Care, Department of Quality of Care, Research Centre for Allied Health Care, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands; and Professor of Manual Therapy, Free University of Brussels, Faculty of Medicine and Pharmacology, Postgraduate Education Manual Therapy, Brussels, Belgium.

Dr van Tulder, Dr Hendriks, Dr Koopmanschap, Dr Bouter, and Dr Oostendorp provided concept/idea/research design. Dr Bekkering, Dr van Tulder, Dr Hendriks, Dr Koopmanschap, and Dr Bouter provided writing. Dr Bekkering provided data collection, and Dr Bekkering, Dr van Tulder, Dr Hendriks, and Dr Knol provided data analysis. Dr van Tulder, Dr Bouter, and Dr Hendriks provided fund procurement. Dr Bekkering provided subjects and clerical support. Dr Oostendorp provided facilities/equipment. Dr Hendriks provided institutional liaisons. Dr Bekkering, Dr Hendriks, Dr Koopmanschap, Dr Knol, Dr Oostendorp, and Dr Bouter provided consultation (including review of manuscript before submission). The authors acknowledge Arno J Engers' contribution to the development of the active strategy and provision of training sessions and Michel Wensing's advice on the content of the active strategy.

The Medical Ethics Committee of VU University Medical Center approved the study.

The results were presented at the National Physical Therapy Congress, November 2003, The Hague, the Netherlands.

This study was supported by a grant from the Ministry of Health, Welfare and Sports.

An accompanying article on process of care outcomes was published in the April 2005 issue of *Journal of Quality and Safety of Health Care.*

This article was received October 31, 2004, and was accepted March 20, 2005.

made; thus, the majority of patients with low back pain could be considered to have nonspecific low back pain.[9]

The guidelines distinguish patients with a normal course of back pain from those with an abnormal course. Patients with a "normal course" of back pain are described as having complaints with a maximum duration of 3 weeks or showing an improvement in their physical functioning over the last 3 weeks. Patients with an "abnormal course" of back pain are described as having complaints for more than 3 weeks with no signs of improvement in physical functioning.[6] The course reflects the patient's potential prognosis and the intensity and content of treatment recommended by the guidelines. For all patients, the guidelines suggest that the intervention should consist of an active approach in which the patient learns to take control over his or her back pain. The physical therapy intervention should focus on restoring physical functioning and improving participation as soon as possible, including retaining or returning to work. For patients with a normal course, a limited number of sessions are recommended. The most important interventions are reassurance, adequate information, and the advice to stay active. For patients with an abnormal course, the guidelines suggest adding exercise therapy with a behavioral approach.[6]

The main benefit of clinical guidelines is to improve the quality of care.[10] Therapists, patients, and insurance companies may have different perspectives about the definition of "high-quality care." From the point of view of the therapist, guidelines should improve the process of care, whereas patients expect improved outcomes and insurance companies look for improved cost-effectiveness. Although guidelines on physical therapy in the Netherlands ultimately intend to improve patient outcomes, the main reason for their formulation was to decrease the variation in care provided.

Publication of guidelines does not automatically result in their use in practice.[11] Therefore, some form of implementation is needed. Systematic reviews on the effectiveness of implementation of interventions show that information transfer is an essential part of the implementation process, but that multiple interventions are usually needed to achieve changes in practice.[12] Another review showed that effective interventions for implementation include reminders, multifaceted interventions, and interactive educational meetings. Interventions such as educational materials and didactic educational meetings are unlikely to be effective.[13] In addition, strategies that are closely linked to the level of the clinical decision-making process are more likely to have good results.[14,15] In summary, reviews on the effect of implementation of interventions suggest that an active, multifaceted intervention may have the highest chance of success in changing the practice of professionals.

The KNGF standard introduces guidelines on physical therapy by disseminating them among their members by mail. Although the guidelines include some educational tools, there is a lack of active and structured implementation. Because it is known that passive approaches are unlikely to achieve substantial changes in practice,[12,16] a new active strategy was developed. This strategy was built on the results of a survey among 100 physical therapist practices to identify perceived barriers for implementation of the Dutch physical therapy guidelines. The active strategy was further developed using a model for changing behavior of professionals and systematic reviews on the effectiveness of implementation interventions.[17]

In this article, we describe the effects of this active implementation strategy for the Dutch guidelines for patients with low back pain, with specific attention to patient outcomes. This article accompanies an article published in the *Journal of Quality and Safety of Health Care*[18] that describes positive effects of the active implementation strategy on adherence of the physical therapists to the guidelines. Physical therapists in the intervention group more often gave 3 or fewer sessions to patients with a normal course of back pain (27% of patients in the intervention group and 13% of patients in the control group), set functional treatment goals (79% versus 71%), used mainly active interventions (77% versus 60%), and gave adequate patient education (96% versus 87%).[18]

Based on these results, we expected that the active strategy would have beneficial effects on patient outcomes. First, we hypothesized that active implementation of the guidelines would improve physical functioning. Second, no improvements in intensity of pain were expected because the guidelines recommend aiming the treatment on improving functioning rather than on reducing pain. Finally, the patients' coping strategy and beliefs about their back pain were expected to be improved with a potential beneficial effect on sick leave, because the guidelines put more emphasis on adequate information and education.

Method

From May 2001 until December 2002, a cluster-randomized controlled trial was carried out among physical therapists in the Netherlands. This trial aimed to evaluate the effect of an active implementation strategy given to therapists on guideline adherence and patient outcomes. We used the practice as a unit of randomization and analyses to minimize contamination among physical therapists.

Recruitment of Physical Therapists

The KNGF randomly selected, from all of their member practices (N=6,261), 325 practices located in or around the cities of Utrecht, Amersfoort, and Hilversum in the center of the Netherlands. All selected practices received a letter from the primary investigator (GEB) explaining the purpose and methods of the study and asking them whether they were willing to participate. Physical therapists were eligible for participation if they worked in primary care and if they expected to treat at least 5 patients with low back pain during the enrollment period. A total of 112 practices responded to the letter, of which 39 practices were interested in participation and 73 were not interested. Two hundred thirteen practices did not respond to our letter, and these practices received a telephone follow-up. An information meeting was organized for those practices that were interested in participation. Eventually, 68 practices (113 physical therapists) participated in our study. In 38 practices, only 1 physical therapist participated, whereas 2 to 5 therapists participated in the other 30 practices.

Randomization

For each practice, all physical therapists were randomly allocated to either an intervention group or a control group. A statistician who was not involved in this study drew up an allocation schedule using a computerized random number generator. Block randomization (in blocks of 4 practices) was carried out after prestratification for the work setting (solo and duo practices versus group practices). The primary investigator, without any knowledge of the practices, listed them alphabetically by name of their street and subsequently assigned them to the intervention group or the control group, according to the allocation schedule.

Control group. The control group consisted of physical therapists who received the guidelines by the standard passive method of dissemination that is used by the KNGF. *Standard dissemination* implies that all physical therapists who are a member of the KNGF receive the guidelines by mail, along with 4 forms to facilitate use of the guidelines. These forms consisted of a self-evaluation form to assess whether the current management is consistent with the guidelines, 2 forms facilitating discussion with other physical therapists or general practitioners, and a copy of the Quebec Back Pain Disability Scale (QBPDS).[19,20] A summary of the guidelines was included in mailing. An article was published in a Dutch professional journal for physical therapists about the development of the guidelines.[6] In order to simulate the standard method of implementation, the physical therapists in the control group were instructed "to act as usual," that is, to read the guidelines on low back pain if they have read previous published guidelines and not read these guidelines if they have not read any other guidelines.

Intervention group. The intervention group consisted of physical therapists who received an additional active strategy to implement the guidelines. The strategy consisted of 2 training sessions, each lasting 2.5 hours with groups of 8 to 12 physical therapists.[17] For each session, a preparation time of 2 hours was recommended. The aim of the strategy was to improve the knowledge and skills of physical therapists with regard to evidence-based physical therapy for patients with low back pain. During the sessions, we used interventions that have been shown to be effective, such as interactive education and discussion, feedback, and reminders.[12–15,21] The content of the strategy was determined on the basis of information about the expected barriers for implementation that was gathered during the development of the guidelines. The first session contained a didactic overview of the diagnostic and treatment process, questions and discussion, and 2 role-plays with an actor.

There was a 4-week time interval between the 2 training sessions. This period aimed to give physical therapists the opportunity to try out the lessons learned in practice and gain experience in using the guidelines. Their experiences were discussed at the second training session, in which participants also received feedback on current management and reminders about correct patient education. Two experts gave advice on the content of the strategy. The primary investigator and 1 of 2 additional trainers with ample experience in the management of low back pain delivered the training sessions. The therapists in the intervention group also received all interventions of the standard passive method of dissemination.

Recruitment of Patients

Patients were included following the final physical therapist training sessions. All participating physical therapists were asked to include consecutive patients who had been referred by their general practitioner or medical specialist because of a new episode of nonspecific low back pain (ie, low back pain with no specific cause, such as tumor, infection, fracture, herniated disk, and so on). To prevent therapists from including a disproportional number of patients, a maximum of 10 patients per therapist was set. Patients could be included only if the diagnosis by the physical therapist was nonspecific low back pain and if the patient was able to complete questionnaires in the Dutch language. Patients who were pregnant were excluded. All patients gave written informed consent.

Outcome Measures

A battery of self-report questionnaires was used to measure patient outcomes. The patients completed questionnaires at baseline and again at 6, 12, 26, and 52 weeks after baseline. The baseline questionnaire also contained some demographic questions and general questions about back pain. The primary outcomes were physical functioning (QBPDS),[19,20] pain (11-point numeric rating scale [NRS]),[22,23] and sick leave (number of days off work in the last 6 weeks). The QBPDS consists of 20 activities of daily living. Each activity is scored on a 6-point scale ranging from 0 ("no trouble") to 5 ("unable to"), and the total score ranges from 0 ("no dysfunction") to 100 ("maximum dysfunction"). The QBPDS has been shown to have good psychometric qualities (Pearson correlation coefficients and intraclass correlation coefficients [ICCs] for test-retest reliability=.90).[20] The cross-sectional construct validity coefficients were .80 using the Roland-Morris Disability Questionnaire.[20] Pain was rated on an 11-point NRS ranging from 0 ("no pain") to 10 ("very severe pain"). An effect size of 0.86 for responsiveness of the NRS has been reported.[22] Sick leave was measured by asking the patients how many days they were off work due to low back pain in the previous 6 weeks.

At baseline and after 12 and 52 weeks, pain coping (Pain Coping Inventory [PCI])[24] and pain beliefs (Back Beliefs Questionnaire [BBQ])[25] were measured as secondary outcomes. The PCI assesses specific cognitive and behavioral pain coping strategies and consists of 33 statements on what a patient might think or may do in case of back pain (eg, "When I have back pain, I take a bath or shower [relaxation]," "When I have back pain, I focus on the pain all the time [catastrophizing]"). Each item is scored on a 4-point scale ranging from 1 ("seldom or never") to 4 ("very often"). Reversed items were recoded and separate sum scores were calculated for 3 active coping strategy subscales (transforming [4–16], relaxation [5–20], and lowering demands [3–12]) and for 3 passive coping strategy subscales (withdrawing [7–28], catastrophizing [9–36], and resting [5–20]). A higher score refers to the use of more active and passive coping strategies for active and passive subscales, respectively. The validity of data for the PCI has been investigated in several groups of patients; alpha coefficients for internal consistency within the 6 subscales for outpatients with chronic pain ranged between .62 and .77. Test-retest reliability in patients with rheumatoid arthritis ranged between .43 and .82 for the 6 subscales.[24]

The BBQ measures beliefs about the inevitable consequences of future life with low back pain and consists of 14 items. Patients indicate their degree of agreement on a 5-point scale (1="totally disagree" to 5="totally agree"). Examples of items are: "There is no real treatment for back pain" and "People with back pain should do exercises." Reserved items were recoded, and a sum score was calculated. The scores range from 9 to 45, and a higher score represents more positive beliefs, suggesting better ability to cope with low back pain. Symonds et al[25] reported a Cronbach alpha of .7 for internal consistency of the BBQ and an ICC of .87.

Patients who agreed to participate received the first questionnaire from the physical therapist with a stamped addressed envelope in which to return it. If the questionnaire had not been sent back within 10 days, the patient received a reminder in the form of a postcard. If, subsequently, the questionnaire was not sent back within the next 11 days, the patient was contacted by telephone. Incoming questionnaires were verified for missing data, and, if necessary, the patient was telephoned and asked to provide the missing information. The people who made the telephone calls were not blinded to group allocation of the patients because usually only a few questions were missing. If a whole questionnaire was left blank, a paper copy was mailed to the patient with the request to complete and return it. The follow-up questionnaires and stamped addressed envelopes were mailed to the patients. The same reminders (postcard and telephone call) were used to increase the response level, and again telephone calls were made in case of missing values.

Data Analysis

The baseline characteristics of the patients in the 2 study groups were compared using chi-square tests, unpaired Student *t* tests, or Mann-Whitney *U* tests. Due to the skewed distribution of the data, unadjusted results were presented using the median and the interquartile range (range between the first and third quartiles).

To determine the effectiveness of the intervention, multilevel regression analysis was performed.[26] Regression analysis can assess the strength of the relationship between the implementation of intervention and patient outcome. However, an assumption of regression analysis is that the observations are independent. Due to the clustered design of this study, the outcome of each patient cannot be assumed to be fully independent of that for any other patient. Multilevel regression analysis was performed to analyze the patient outcomes adjusting for this clustering of data. The data of this study are clustered at 4 levels: practice, physical therapist, patient, and follow-up moment (Fig. 1).

The analysis results in regression coefficients, which can be interpreted as the difference on patient outcomes between the 2 groups at a certain time point. Wald chi-square tests were used to determine whether the regression coefficients were statistically significant ($P<.05$).

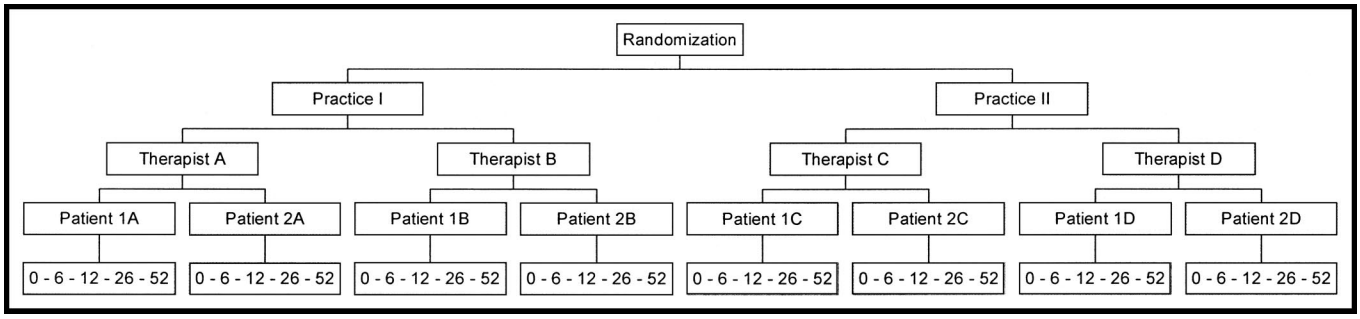

Figure 1.
Clustering of data in the trial. 0=baseline moment; 6, 12, 26, and 52=numbers of weeks after baseline.

Main analysis. Multilevel modeling also has the ability to adjust for confounders on several levels (eg, it allows adjustment for variables on both practice level and patient level).[27] The analyses were adjusted for baseline values and for health insurance (which differed between the 2 groups at baseline).[28] The analyses of pain and physical functioning also were adjusted for potential prognostic variables based on the literature (ie, sex,[29,30] duration of the current episode of back pain,[31] previous back pain episode[s],[32] and education[33]). All multilevel analyses were performed with MLwiN (version 1.10).[34,*]

In the primary analysis, data for all patients of physical therapists who were originally randomly allocated to the intervention group were included in the analysis, regardless of whether their therapists had attended the training sessions. This is according to the intention-to-treat principle that refers to the practice of attributing all participants to the randomized group regardless of what subsequently occurred.[35] In the secondary analysis, data for the patients of the physical therapists in the intervention group who had not attended both training sessions were excluded from the analyses.

Subgroup analysis. Separate additional analyses were performed to examine potential effect modification for the subgroups of patients with acute (<6 weeks), subacute (6–12 weeks), and chronic (>12 weeks) low back pain and for the subgroups of patients with and without previous back problems.

Sample size calculation. The a priori calculation of sample size was based on the ability to detect a standardized difference of 0.30 for physical functioning (QBPDS) and pain (NRS), which we considered as clinically relevant. This was based on the assumption of finding a small difference and a large standard deviation because patients with acute, subacute, and chronic low back pain could be included. The difference was estimated to be 5 (SD=15). A sample size of 349 patients was required (2-sided α=.05, β=.20). To achieve the equivalent power of a patient randomized trial, standard sample size calculations for clustered trials need to be inflated using a measure for intracluster dependence, the ICC, and the size of the clusters. The ICC takes a value of between 0 and 1 and would be high if, for example, the outcome of patients within practices is very consistent but there is wide variation across different practices.[27] The sample size was corrected by an ICC of .05, which is an estimate for the correlation of outcome variables in primary care.[27,36] We estimated the cluster size to be 5 patients per physical therapist. Therefore, a total of 418 patients and a total of 84 physical therapists were needed.

Results

Study Population

During a period of 8 months (May–December 2001), 515 patients with nonspecific low back pain were included. Data of 15 patients were excluded (2 patients did not meet the inclusion criteria, and 13 patients did not give informed consent). The flow diagram in Figure 2 shows the number of patients throughout the trial.

Dropouts (n=72) were younger than patients who completed questionnaires at the 52-week follow-up (*P*=.001); there were no other differences between the 2 groups with regard to baseline characteristics. The mean number of patients recruited by one physical therapist (cluster size) was 5 (median=9, interquartile range=5–10). The median number of treatment sessions was 8.0 in the control group and 6.5 in the intervention group (*P*=.001).

Baseline Characteristics

Table 1 presents the baseline characteristics of the patients in both groups. More patients in the intervention group had public health insurance. Table 2 shows the unadjusted outcomes for both groups with regard to physical functioning, pain, and sick leave. The unadjusted outcomes for both groups with regard to beliefs and coping are presented in Table 3. At baseline,

* Centre for Multilevel Modelling, Institute of Education, 20 Bedford Way, London WC1H 0AL, United Kingdom.

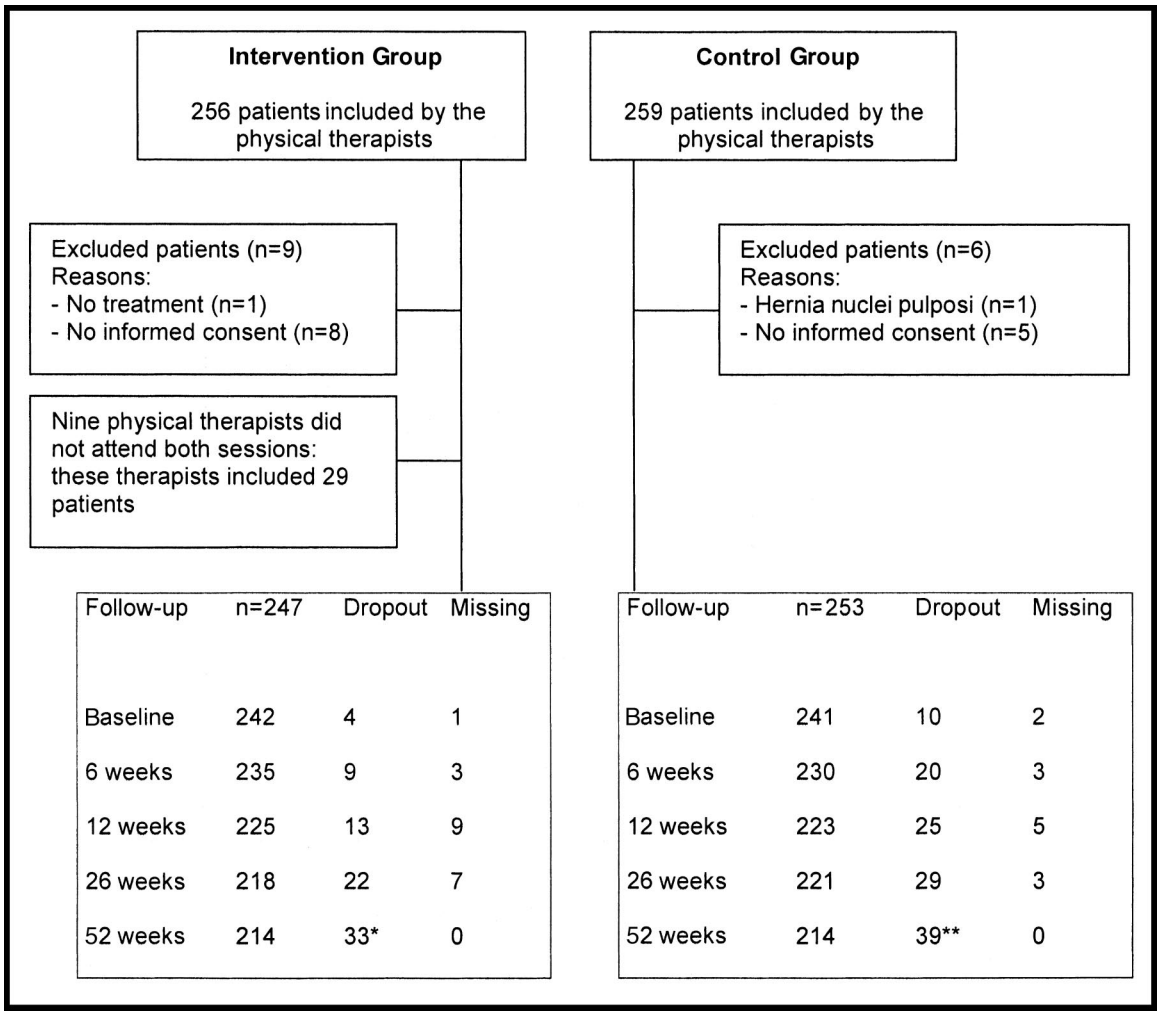

Figure 2.
Flow diagram of patients throughout the trial. Dropout=cumulative number of patients who did not return the questionnaire from a certain time forward; missing=number of patients who did not return a questionnaire but contributed to later follow-up. Single asterisk (*) indicates reasons for dropout (n=33): did not return questionnaires (n=18), not wanting to participate anymore because the patient had recovered from his or her back pain (n=3), the treatment had no effect (n=1), there were too many irrelevant questions (n=2), the patient did not have enough privacy (n=2), the patient could not be contacted (n=2), the patient was not registered by the therapist (n=1), the patient lost interest (n=2), and private issues (n=2). Double asterisk indicates reasons for dropout (n=39): did not return questionnaires (n=20), not wanting to participate anymore because the patient had recovered from his or her back pain (n=2), the treatment had no effect (n=3), there were too many irrelevant questions (n=5), the patient could not be contacted (n=3), the patient was not registered by the therapist (n=2), the patient lost interest (n=3), and private issues (n=1).

patients in the intervention group scored higher on the relaxation subscale of the PCI (P=.009). There were no other differences between the 2 groups with regard to any of the other characteristics and baseline variables.

Effect of Intervention

The multilevel analysis showed that there was no need to adjust for health insurance, but all other potential prognostic variables and the baseline PCI subscale relaxation were included as adjustments. Variation among physical therapists' practices on physical functioning of the patients was 0%, so this level was deleted from the model. The total percentages of variance at the remaining levels for physical functioning were 3% for physical therapists, 47% for patients, and 50% for measurement time. The total percentages of variance at the remaining levels for pain were 1% for physical therapists, 16% for patients, and 83% for measurement time.

At 12 weeks, the difference in physical functioning between the patients of the 2 groups was 2.83 points on the QBPDS. The 95% confidence interval (CI) included zero (95% CI=−0.66, 6.31), which indicated that there was no difference in the physical functioning of the patients in the 2 groups. The difference in pain intensity between the 2 groups was 0.34 (95% CI=−0.19, 0.88) (Tab. 4). No multilevel analyses were performed for sick leave because only 7% of all patients were on sick leave at 12 months.

Over the total follow-up period of 12 months, there was no difference between the 2 groups on physical func-

Table 1.
Patient Characteristics of Intervention and Control Groups

	Intervention Group (n=247)	Control Group (n=253)
Age (y)		
$\bar{X}$	46.2	44.4
SD	14.8	13.3
Sex (% female)	53.4	50.2
Paid job (% yes)	69.1	77.0
Health insurance (% public)	69.5[a]	60.9
Duration of current episode (%)		
<6 wk (acute)	51.9	49.6
6–12 wk (subacute)	15.5	20.6
>12 wk (chronic)	32.6	29.8
Previous episode of low back pain (% yes)	73.7	72.4
Education (%)		
Low	23.4	29.2
Medium	44.0	31.7
High	32.4	39.1

[a] P=.045.

Table 2.
Unadjusted Median Scores and Interquartile Ranges (in Parentheses) of the Primary Patient Outcome Measurements at Baseline and at 6, 12, 26, and 52 Weeks Follow-up[a]

	Intervention Group (n=247)	Control group (n=253)
Functioning (QBPDS: 0–100)		
Baseline	38.0 (26.5–50.5)	40.5 (26.3–55.8)
6 wk	24.0 (13.0–40.0)	23.5 (11.0–37.8)
12 wk	20.0 (7.0–32.8)	17.5 (6.0–30.8)
26 wk	16.0 (5.0–32.0)	11.0 (4.0–29.0)
52 wk	17.0 (4.6–32.0)	13.0 (4.8–29.0)
Pain (NRS: 0–10)		
Baseline	7.0 (5.0–8.0)	7.0 (5.0–8.0)
6 wk	3.0 (2.0–5.0)	3.0 (2.0–5.0)
12 wk	2.0 (1.0–4.0)	2.0 (1.0–4.0)
26 wk	2.0 (1.0–4.0)	1.0 (0.0–4.0)
52 wk	2.0 (0.0–4.0)	1.0 (0.3–3.0)
Sick leave: % yes (median days during previous 6 wk)		
Baseline	46.4 (5)	42.5 (5)
6 wk	25.5 (10)	22.8 (10)
12 wk	9.6 (10)	9.8 (8)
26 wk	8.8 (9)	6.8 (14)
52 wk	9.8 (30)	5.0 (15)

[a] QBPDS=Quebec Back Pain Disability Scale, NRS=numeric rating scale. A higher score means more pain or more disabilities.

tioning (χ^2=4.88, df=4, P>.05) or on pain (χ^2=6.05, df=4, P>.05). There was also no overall difference between the 2 groups for any of the coping subscales (transforming: χ^2=1.41, df=2; relaxation: χ^2=1.63, df=2; lowering demands: χ^2=1.05, df=2; withdrawing: χ^2=0.34, df=2; catastrophizing: χ^2=0.11, df=2; resting: χ^2=1.39, df=2; P>.05 for all coping subscales) or beliefs (χ^2=0.20, df=2, P=.09).

Per Protocol and Subgroup Analysis

Twenty-nine patients, included by 9 physical therapists who did not attend both training sessions, were excluded from the secondary analysis. The results also showed no difference between the 2 groups on physical functioning or pain (data not shown). Subgroup analysis showed no differences between the 2 groups on physical functioning or pain in either patients with acute low back pain, subacute low back pain, or chronic low back pain or in patients with or without a previous episode of back pain (data not shown).

Discussion

This study evaluated the effect of an active strategy for implementation of the Dutch physical therapy guidelines on patient outcomes. It showed that, although patients in both groups showed a substantial improvement in physical functioning and pain over the first 12 weeks, an active strategy did not improve patient outcomes compared with the standard method of dissemination of the guidelines. These results were similar for all outcomes: physical functioning, pain, sick leave, coping, and beliefs about back pain.

The strength of this study is the cluster-randomized design in which groups of people rather than individuals are randomly allocated to groups. Random allocation to groups by individuals would be inappropriate for evaluating implementation interventions because a physical therapist cannot be expected to use knowledge about the guidelines for the first patient and ignore this knowledge for a second patient. Cluster randomization by practice minimizes the potential for contamination between treatments if different therapists within the same setting

Table 3.
Unadjusted Median Scores and Interquartile Ranges (In Parentheses) of the Secondary Patient Outcome Measurements at Baseline and at 12 and 52 Weeks Follow-up[a]

	Intervention Group (n=247)	Control Group (n=253)
Beliefs (BBQ:9–45)		
Baseline	31.0 (26.0–36.0)	31.0 (26.5–35.0)
12 wk	31.0 (27.0–36.0)	31.0 (27.0–36.0)
52 wk	32.0 (26.0–37.0)	32.0 (27.0–36.0)
Coping (PCI) subscales (active coping)		
Transforming (4–16)		
Baseline	8.0 (7.0–10.0)	8.0 (7.0–9.0)
12 wk	8.0 (7.0–9.0)	8.0 (6.0–9.0)
52 wk	7.0 (6.0–9.0)	7.0 (5.0–9.0)
Relaxation (5–20)		
Baseline	12.0 (10.0–14.0)[b]	11.0 (9.0–14.0)
12 wk	12.0 (10.0–14.0)	11.0 (8.5–14.0)
52 wk	10.0 (8.0–13.0)	10.0 (7.0–13.0)
Lowering demands (3–12)		
Baseline	7.0 (6.0–8.0)	6.5 (6.0–8.0)
12 wk	7.0 (5.0–8.0)	7.0 (5.0–8.0)
52 wk	6.0 (5.0–7.0)	6.0 (5.0–7.0)
Coping (PCI) subscales (passive coping)		
Withdrawing (7–28)		
Baseline	10.0 (8.0–12.0)	10.0 (8.0–12.8)
12 wk	10.0 (8.0–12.0)	9.0 (8.0–12.0)
52 wk	9.0 (7.0–12.0)	8.0 (7.0–12.0)
Catastrophizing (9–36)		
Baseline	16.0 (13.0–20.0)	16.0 (13.0–20.0)
12 wk	15.0 (12.3–18.8)	16.0 (13.0–19.0)
52 wk	14.0 (11.0–17.0)	14.0 (11.0–17.0)
Resting (5–20)		
Baseline	11.0 (9.0–13.0)	11.0 (9.0–13.0)
12 wk	11.0 (9.0–13.0)	11.0 (9.0–13.0)
52 wk	10.0 (8.0–12.0)	10.0 (8.0–12.0)

[a] BBQ=Back Beliefs Questionnaire, PCI=Pain Coping Inventory. A higher score on the BBQ and PCI subscales for active coping indicates more positive beliefs and adequate coping (more active coping); a higher score on the PCI subscales for passive coping indicates inadequate coping (more passive coping).
[b] Difference between groups (P=.009).

manage trial patients. This design, however, also explains why the patient characteristics between the 2 groups were slightly different. Random allocation to groups at the level of the patient would generate 2 groups with similar patient characteristics. Random allocation to groups at the level of the practice thus generates 2 groups with similar practice characteristics and does not guarantee the patients included by these therapists having similar characteristics. However, the analysis showed that the differences in patients characteristics did not influence the results.

Limitations

A limitation of this study is that the participating physical therapists were self-selected. They probably volunteered because they were interested in low back pain, and they may have already been familiar with the latest evidence in this field. This limitation results in a control group that may have read the guidelines or that may already adhere to important recommendations of the guidelines. This limitation potentially decreased the contrast in guideline adherence between both groups. This is expected to be the most important reason for the lack of difference in patient outcomes between both groups. It is also possible that the control group read the guidelines better than they would have done if they did not participate in this study.

Another restriction is that selection may have taken place because mainly motivated patients will have completed questionnaires. Although there is no reason to expect that the proportion of motivated patients is not equally distributed among the 2 groups, this restriction may limit the generalizability of the results. In addition, telephoning patients and asking them to provide missing data without blinding to group allocation could have introduced bias. However, it also is not expected that this restriction importantly biased our results because usually only a few questions had been left blank.

Finally, this trial evaluated the effect of the implementation of guidelines on patient outcomes. Implementation of guidelines will result in improved patient outcomes only if patients managed according to the recommendations in the guidelines have better outcomes. However, we expected patient outcomes to be improved because the guidelines are evidence-based. Thus, if a systematic review shows that exercise therapy improves functioning, we expect that using guidelines recommending exercise therapy will improve patient outcomes. Therefore, we argue that it is unlikely that this limitation contributed to the lack of effect on patient outcomes.

Although guidelines are expected to improve patient outcomes, the evidence is sparse and not yet convincing.[37] Worrall et al[37] suggested that failure to find positive results may be related to the finding that most of the investigated guidelines were relatively old. The studied guidelines concerned chronic conditions such as

Table 4.
Adjusted Multilevel Regression Coefficients[a] and 95% Confidence Intervals (in Parentheses) for the Outcomes Physical Functioning and Pain[b]

Effect of intervention at:	Functioning (QBPDS)	Pain (NRS)
6 wk	1.96 (−1.44, 5.37)	0.16 (−0.35, 0.69)
12 wk	2.83 (−0.66, 6.31)	0.34 (−0.19, 0.88)
26 wk	4.00 (0.68, 7.33)	0.62 (0.06, 1.18)
52 wk	3.55 (−0.25, 7.35)	0.55 (−0.02, 1.11)
Overall effect: χ^2 (df=4)	4.88 (P>.05)	6.05 (P>.05)

[a] Adjusted for baseline, sex, previous episode of back pain, duration of the current episode of back pain, education, and Pain Coping Inventory relaxation subscale. The regression coefficient represents the mean difference on the outcome variable between the intervention and control groups for a certain measurement, adjusted for these variables. The overall effect represents the effect of the active strategy across all measurements.
[b] QBPDS=Quebec Back Pain Disability Scale, NRS=numeric rating scale.

hypertension, asthma, obesity, and smoking cessation. Gross et al[11] argued that failure to implement clinical guidelines for such conditions emphasizes our need to change the structure and process of care.

With regard to patients with low back pain, the need to change the structure or process of care is supported by a study by Rossignol et al.[38] They implemented a program for coordination of primary health care treatment of patients who had been off work for 4 to 8 weeks due to low back pain. Even though this intervention had only a small effect on the latency to return to work, it resulted in improved outcomes with regard to pain and functional status after 6 months. McGuirk et al[39] also reported favorable outcomes of introducing evidence-based guidelines in patients with acute low back pain (defined as <12 weeks) in a nonrandomized trial. In contrast, educating a professional did not seem sufficient to improve patient outcomes.[40] Curtis et al[41] concluded that the lack of results may be related to the failure of some participants to undertake the required sequence of maneuvers. Only 43% of the patients who underwent manual therapy actually received the complete planned sequence of maneuvers, which could be related to a lack of skills among the care providers. This lack of specific skills may have played a role in the present results, because our guidelines require a certain level of knowledge and skills from the therapists. It is possible that the strategy was not successful enough for the physical therapists to gain sufficient new knowledge and skills to improve patient outcomes. A central element in the guidelines is, for example, the use of behavioral principles. These skills are difficult to teach in only 2 training sessions because they need to be integrated in the entire diagnostic and treatment process. This lack of specific skills also may explain why we did not find a difference in the patient's beliefs, as we hypothesized. The therapists may need more training to be able to change patients' beliefs.

It has been shown that a change in management is not automatically accompanied by a better health outcome.[40,41] It may be possible that the intensity of our active strategy (eg, number of contact hours, number of sessions or frequency of sessions) was insufficient to result in improved patient outcomes. Our strategy included 5 contact hours. This is more than the 2-hour program of Cherkin et al,[42] but less than the 18-hour program of Curtis et al.[41] A 192-hour program for physical therapists, which was given over a period of 12 months, improved outcomes in patients with chronic low back pain.[43] Although it may be expected that a more intense strategy automatically improves the results, this may not necessarily be true. Specific barriers for implementation, such as culture within settings or organizations or financial barriers, need specific interventions and not just a more intense strategy.

Finally, it is also possible that no further benefit is to be expected for patients with low back pain. The outcomes in both groups showed an important improvement in the first 12 weeks, and it may be difficult to further improve these results.

Physical Therapy Guidelines on Patients With Low Back Pain

Treatment according to the Dutch guidelines on physical therapy does not imply identical interventions for all patients. Sackett and colleagues[44] defined *evidence-based medicine* as the integration of best research with clinical expertise and patient values. The guidelines make broad recommendations based on evidence for the "average" patient. The physical therapist needs to translate this evidence onto the individual patient with his or her clinical presentation, preferences, and experiences. Despite the above, one of the problems of current guidelines on low back pain is the lack of detail on specific physical therapy diagnoses and interventions.[7] Based on current evidence, it is not possible to give very specific recommendations for the management of patients with low back pain. Consequently, there will still be a lot of variation among the interventions. Furthermore, although all interventions could all well be within the recommendations of the guidelines, this may decrease the effect on patient outcomes. New and methodologically sound studies are needed to provide evidence for more detailed physical therapy interventions and diagnoses. The guidelines were published in 2001, and an update is necessary to include recent studies. A recent study[45] suggests that classification-based physical therapy for patients with occupational low back pain results in better patient outcomes if compared with treatment according to guidelines. Another study[46] sup-

ports predictive validity for the pain pattern procedures. Although these findings need to be confirmed by other studies, these type of studies may provide valuable information for future updates of the guidelines.

Future Research

Future research should evaluate whether more intensive implementation strategies are needed to improve patient outcomes. It also should evaluate what characteristics of clinical guidelines make them more likely to result in improvement of patient outcomes, such as more detailed recommendations that take the characteristics of the individual patient into account.

Conclusions

The active implementation strategy designed for this study did not result in additional beneficial effects on patient outcomes. The benefit in physical functioning by the implementation intervention, as we hypothesized, may not have been achieved because the physical therapists of the control group already could have managed their patients according to the guidelines, thus leading to a lack of contrast between the 2 groups.

Although there were no additional benefits on patient outcomes, there still can be other good reasons for using an active strategy to implement guidelines. In case of similar outcomes, a more transparent health care process or reduction in costs can be reasons to recommend this strategy broadly.

The active strategy did improve the adherence of the physical therapists to the guidelines. Therefore, it is useful to assist physical therapists in practicing evidence-based care for patients with low back pain, and it helps to decrease variation in care. Ultimately, this makes the health care process of physical therapy explicit and clear, which improves collaboration with other health care professionals and may contribute to improving patient outcomes. The results of this study also would apply to the implementation of clinical guidelines for similar health problems (ie, other musculoskeletal problems), other physical therapists working in a similar health care system (ie, getting patients referred for intervention), and having autonomy in determining which intervention is best for an individual patient.

References

1 Frymoyer JW. Back pain and sciatica. *N Engl J Med.* 1988;318:291–300.

2 Moseley AM, Herbert RD, Sherrington C, Maher CG. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). *Aust J Physiother.* 2002;48:43–49.

3 Foster NE, Thompson KA, Baxter GD, Allen JM. Management of nonspecific low back pain by physiotherapists in Britain and Ireland: a descriptive questionnaire of current clinical practice. *Spine.* 1999;24:1332–1342.

4 Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. *Phys Ther.* 2001;81:1018–1028.

5 Gracey JH, McDonough SM, Baxter GD. Physiotherapy management of low back pain: a survey of current practice in Northern Ireland. *Spine.* 2002;27:406–11.

6 Bekkering GE, Hendriks HJM, Oostendorp RAB. Clinical guidelines in physiotherapy for low back pain are published [in Dutch: Richtlijn Lage-rugpijn gepubliceerd]. *FysioPraxis.* 2001;4:28–31.

7 Grimmer K, Milanese S, Bialocerkowski A. Clinical guidelines for low back pain: physiotherapy perspective. *Physiother Can.* 2003;55:185–194.

8 Waddell G. *The Back Pain Revolution.* Philadelphia, Pa: Churchill Livingstone Inc; 1998.

9 Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA.* 1992;268:760–765.

10 Woolf SH, Grol R, Hutchinson A, et al. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ.* 1999;318:527–530.

11 Gross PA, Greenfield S, Cretin S, et al. Optimal methods for guideline implementation: conclusions from Leeds Castle meeting. *Med Care.* 2001;39:II85–II92.

12 Wensing M, Van der Weijden T, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective? *Br J Gen Pract.* 1998;48:991–997.

13 Bero LA, Grilli R, Grimshaw JM, et al (The Cochrane Effective Practice and Organization of Care Review Group). Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ.* 1998;317:465–468.

14 Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance, a systematic review of the effect of continuing medical education strategies. *JAMA.* 1995;274:700–705.

15 Grimshaw J, Freemantle N, Wallace S, et al. Developing and implementing clinical practice guidelines. *Qual Health Care.* 1995;4:55–64.

16 Grimshaw JM, Shirran L, Thomas R, et al. Changing provider behavior: an overview of systematic reviews of interventions. *Med Care.* 2001;39:II2–II45.

17 Bekkering GE, Engers AJ, Wensing M, et al. Development of an implementation strategy for the physiotherapy guideline on low back pain. *Aust J Physiother.* 2003;49:208–214.

18 Bekkering GE, Hendriks HJM, van Tulder MW, et al. The effect on the process of care of an active strategy to implement the clinical guidelines on physiotherapy for low back pain: a cluster-randomised controlled trial. *Qual Saf Health Care.* 2005;14:107–112.

19 Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec Back Pain Disability Scale: measurement properties. *Spine.* 1995;20:341–352.

20 Schoppink LE, van Tulder MW, Koes BW, et al. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther.* 1996;76:268–275.

21 Davis DA, Taylor-Vaisey A. Translating guidelines into practice: a systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *CMAJ.* 1997;157:408–416.

22 Bolton JE, Wilkinson RC. Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients. *J Manipulative Physiol Ther.* 1998;21:1–7.

23 Lundeberg T, Lund I, Dahlin L, et al. Reliability and responsiveness of three different pain assessments. *J Rehabil Med.* 2001;33:279–283.

24 Kraaimaat FW, Evers AWM. Pain-coping in chronic pain patients: Psychometric characteristics of the Pain-Coping Inventory (PCI). *Int J Behav Med.* 2003;10:343–63.

25 Symonds TL, Burton AK, Tillotson KM, Main CJ. Do attitudes and beliefs influence work loss due to low back trouble? *Occup Med (Lond).* 1996;46:25–32.

26 Twisk JWR. *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide.* Cambridge, United Kingdom: University Press; 2003.

27 Campbell MK, Mollison J, Steen N, et al. Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract.* 2000;17: 192–196.

28 Altman DG. *Practical Statistics for Medical Research.* London, United Kingdom: Chapman & Hall; 1996.

29 Von Korff M, Deyo RA, Cherkin D, Barlow W. Back pain in primary care: outcomes at 1 year. *Spine.* 1993;18:855–862.

30 Gatchel RJ, Polatin PB, Mayer TG. The dominant role of psychosocial risk factors in the development of chronic low back pain disability. *Spine.* 1995;20:2702–2709.

31 van den Hoogen HJ, Koes BW, Deville W, et al. The prognosis of low back pain in general practice. *Spine.* 1997;22:1515–1521.

32 Thomas E, Silman AJ, Croft PR, et al. Predicting who develops chronic low back pain in primary care: a prospective study. *BMJ.* 1999;318:1662–1667.

33 Cats-Baril WL, Frymoyer JW. Identifying patients at risk of becoming disabled because of low-back pain: the Vermont Rehabilitation Engineering Center predictive model. *Spine.* 1991;16:605–607.

34 Rasbash J, Browne W, Goldstein H, et al. *A User's Guide to MlwiN: Version 2.1d for Use With MLwiN 1.10.* London, United Kingdom: Centre for Multilevel Modeling, Institute of Education, University of London; 2000.

35 Guyatt GH, Sackett DL, Cook DJ (Evidence-Based Medicine Working Group). Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? *JAMA.* 1993;270:2598–2601.

36 *Changing Professional Practice: Theory and Practice of Clinical Guidelines Implementation.* Copenhagen, Denmark: Danish Institute for Health Services Research and Development; 1999.

37 Worrall G, Chaulk P, Freake D. The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review. *CMAJ.* 1997;156:1705–1712.

38 Rossignol M, Abenhaim L, Seguin P, et al. Coordination of primary health care for back pain: a randomized controlled trial. *Spine.* 2000;25:251–258.

39 McGuirk B, King W, Govind J, et al. Safety, efficacy, and cost effectiveness of evidence-based guidelines for the management of acute low back pain in primary care. *Spine.* 2001;26:2615–2622.

40 Cherkin D, Deyo RA, Berg AO. Evaluation of a physician education intervention to improve primary care for low-back pain, II: impact on patients. *Spine.* 1991;16:1173–1178.

41 Curtis P, Carey TS, Evans P, et al. Training primary care physicians to give limited manual therapy for low back pain: patient outcomes. *Spine.* 2000;25:2954–2960.

42 Cherkin D, Deyo RA, Berg AO, et al. Evaluation of a physician education intervention to improve primary care for low-back pain, I: impact on physicians. *Spine.* 1991;16:1168–1172.

43 Levsen M, Hansen ML, Kent AD, et al. Effects of physical therapist training on outcomes of patients with chronic low back pain or chronic shoulder pain. *J Manual Manipulative Ther.* 2001;9:84–91.

44 Sackett DL, Strauss SE, Richardson WS, et al. *Evidence-Based Medicine.* 2nd ed. Edinburgh, United Kingdom: Churchill Livingstone; 2000.

45 Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain. *Spine.* 2003;28:1363–1372.

46 Werneke MW, Hart DL. Categorizing patients with occupational low back pain by use of the Quebec Task Force Classification System versus pain pattern classification procedures: discriminant and predictive validity. *Phys Ther.* 2004;84:243–254.