

Does Strengthening the Abdominal Muscles Prevent Low Back Pain — A Randomized Controlled Trial

ANTOINE HELEWA, CHARLES H. GOLDSMITH, PETER LEE, HUGH A. SMYTHE, and LORIE FORWELL

ABSTRACT. *Objective.* To investigate the effects of abdominal muscle strengthening exercise on low back pain (LBP) risk reduction.

Methods. We compared the effects of abdominal muscle strength (AMS) exercise and back education with back education only on 402 asymptomatic subjects with weak AMS. The primary outcome was the percentage of subjects with at least one episode of LBP at 24 months. A diary was used to monitor compliance.

Results. There were no statistically significant differences between experimental (E) and control (C) subjects at 24 months for LBP episodes (E: 34.7%, C: 30.4%; C - E = -4.2%, P2 = 0.481; 95% CI -16.1%, 7.6%). The LBP episodes were also not statistically significantly different at 6 months (E: 13.2%, C: 16.1%; C - E: 2.9%, P2 = 0.493; 95% CI -5.3%, 11.0%) or at 12 months (E: 24.8%, C: 23.6%; C - E = -1.2%, P2 = 0.818; 95% CI -11.6%, 9.2%). Adjusting the data for strata and baseline AMS did not alter the findings. Imputed results for LBP episodes at 6 months (C - E: 4.8%, P2 = 0.191; 95% CI -2.4%, 12.0%), 12 months (C - E: -1.0%, P2 = 0.821; 95% CI -9.5%, 7.6%), and 24 months (C - E: -3.3%, P2 = 0.483; 95% CI -12.6%, 5.9%) were also not statistically significantly different.

Conclusion. Back education and abdominal exercise instructions are similar to back education alone. There were no group differences in LBP episodes, possibly due to noncompliance with the exercise program. While the estimated benefit of 2.9% at 6 months from the complete data and 4.8% from the imputed data were not statistically significant, a larger study or future metaanalyses may be needed to confirm whether there is a clinical benefit or whether these results should be considered a chance finding. (J Rheumatol 1999;26:1808-15)

Key Indexing Terms:

LOW BACK PAIN
RANDOMIZED CONTROLLED TRIAL

PRIMARY PREVENTION
COMPLIANCE
EXERCISE

Low back pain (LBP) is a common source of disability, with major consequences for health care resources. LBP can affect up to 80% of adults during their lives: from 30 to 50% suffer from LBP during the course of a year¹⁻³. The prognosis for acute LBP is generally very good, with the majority resolving completely within months. Up to 10% of cases, however, become chronic, resulting in a high cost to society^{1,4-6}. To date, a substantial portion of health care funds are spent on the medical and rehabilitative management of LBP,

and very little on its prevention^{7,8}. To prevent LBP among those at risk, 2 commonly used interventions are exercise and back education⁸.

In a recent review of 16 studies on the prevention of LBP using exercise⁸, 4 were randomized trials, using aerobic exercise or exercises to strengthen the trunk flexors or extensors⁹⁻¹². Subjects in the intervention groups had fewer days of work lost because of LBP; none was followed up for longer than 18 months; and the sample sizes (a total of 350 patients in all 4 trials) were relatively small. Altogether, the studies suggest that exercise intervention is mildly protective against LBP; it is not clear whether trunk extensor, trunk flexor, or aerobic exercise is preferable.

Of particular interest to us were the results of the trunk flexor exercise study by Donchin, *et al*¹². Subjects who performed trunk flexor exercises reported significantly less pain, compared to a back education group and a no treatment control group. This in part supports the notion that exercise to strengthen the abdominal muscles may be protective. Stronger abdominal muscles may produce higher intraabdominal pressures when performing physically demanding tasks, such as lifting heavy objects. It is assumed that this increased pressure results in a hydraulic lift effect, and safeguards the lower back against injury¹³. No experimental or clinical evidence exists to support this hypothesis.

From the School of Physical Therapy and Department of Epidemiology, University of Western Ontario, London; The Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University and Father Sean O'Sullivan Research Centre, St. Joseph's Hospital, Hamilton; and the Department of Medicine, University of Toronto, Toronto, Canada.

Supported by grants from the Ontario Ministry of Labour (grant 255/R) and the Conn Smythe Foundation, Toronto, Canada.

A. Helewa, MSc (Epid), School of Physical Therapy, University of Western Ontario; C.H. Goldsmith, PhD, Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University; P. Lee, FRCPC; H.A. Smythe, FRCPC, Department of Medicine, University of Toronto; L. Forwell, MSc, School of Kinesiology, University of Western Ontario.

Address reprint requests to Prof. A. Helewa, Department of Physical Therapy, Elborn College, University of Western Ontario, London, Ontario, Canada N6G 1H1.

Submitted April 8, 1998 revision accepted January 21, 1999.

Back education is the second most commonly used prevention strategy. Although educational programs are popular, tend to be attractively packaged, and can be provided at low cost, their effect on the prevention of LBP appears to be minimal. One of 5 randomized trials showed positive clinical outcomes, but it is difficult to attribute the benefit to education alone^{8,9}. More recently, a large scale, randomized controlled trial of an educational program to prevent work associated LBP found no longterm benefit associated with training¹⁴.

To determine whether an increase in abdominal muscle strength and back education reduces the rate of LBP among individuals at risk, we designed a primary prevention randomized controlled trial, comparing the effects of abdominal muscle exercise and back education to back education alone among asymptomatic individuals with reduced abdominal muscle strength.

MATERIALS AND METHODS

Design specifications. A randomized controlled trial was conducted to examine the association between abdominal muscle strength and LBP over a 2 year period. Subjects were reviewed for eligibility by specially trained, non-medical independent assessors using a questionnaire constructed by our group. Eight strata were created according to age (≤ 40 or > 40 years), sex (male or female), and regular exercise (yes or no). Eligible subjects were randomly allocated equally to one of 2 groups in blocks of 4: a control group that received a standardized back education program, or an experimental group that received the same back education program plus exercise to strengthen the abdominal muscles. Random number lists were prepared with the criteria, and the code was held in confidence by research staff. Independent assessors were masked to the random number lists. One of us (AH) made all the decisions about patient eligibility for the trial. Once the patient was eligible, a research assistant opened the allocation envelopes, hence concealing the randomization plan from the person making the eligibility decision.

Study population. Subjects were University of Western Ontario employees and PhD students, University Hospital staff, and London area residents. Prior to the study, the cooperation of university employee groups, budget unit heads, and the Department of Occupational Health and Safety was sought. Participants were sent a letter inviting them to participate and, for select employee groups, brief seminars were given about the study.

Subjects were excluded from the study if they were < 23 or > 57 years of age; had LBP at the time; had LBP in the previous 5 years that required surgery; missed work in the previous 2 years due to LBP; refused to perform the abdominal muscle strength test; had abdominal muscle strength in the middle or upper third strength for that population; had experienced major illness or undergone major surgery in the previous 6 months; or had a problem that precluded them from performing a sit-up.

Interventions. Eligible subjects were stratified, then randomized equally in blocks of 4 to either an experimental or control group. Both groups of subjects then attended back education classes at baseline and at one and 2 years. A 90 min informal back education class was given by a physical therapist that covered spinal anatomy, pathophysiology, posture, lifting techniques, and general fitness. Additional resource information was displayed for greater understanding of back care.

For experimental subjects, a physical therapist obtained a history, conducted a physical assessment of the lower back, and measured mobility, posture and abdominal muscle strength¹⁵. Subjects were then classified into 3 levels of abdominal muscle strength: (1) unable to perform a sit-up, or an abdominal muscle strength score of < 120 mm Hg; (2) could do a sit-up and scored 121 to 160 mm Hg; and (3) scored above 160 mm Hg.

Subjects at level 1 were given extended one-on-one management by a physical therapist and taught special abdominal muscle strength exercises, using manually assisted/resisted diagonal spiral movement patterns or isometric contractions. These patterns also enhanced group action of muscles, rather than causing single muscles to work in isolation¹⁶. When abdominal muscle strength exceeded 120 mm Hg, subjects progressed to level 2.

Subjects at level 2, or those who progressed to level 2, were instructed to perform 3 simple, active graduated sit-up exercises, supervised by the physical therapist. A straight sit-up, and sit-up with rotation to the right and left were used. In each case, the exercises were performed starting with arms outstretched to arms crossed, then graduating to hands on opposite elbows, and finally to arms crossed and hands on opposite shoulders. Subjects were instructed to perform the exercises regularly at home as part of a short term maintenance program. When abdominal muscle strength exceeded 160 mm Hg, subjects were progressed to level 3.

Those at level 3, or those who had progressed to that level, were instructed to perform a single sit-up with flexion and rotation to one side, then the other, hand pressing against opposite knee maximally, for 5 s. They were to perform 6 repetitions on each side alternately, once a day. At 6, 12, and 24 months, experimental subjects were seen by the physical therapist to review the exercises and to enhance their compliance with the prescribed program.

The exercises were designed for simplicity and ease of performance. They required no aids, little time (about 5 min), and were attainable at all fitness levels. Each exercise was performed slowly and rhythmically. Subjects were instructed to report pain or other difficulties resulting from the exercises to the physical therapist, and to perform the exercises once daily, about 30 min after rising in the morning, preferably at the same time of day to establish a daily routine.

To enhance compliance, patients were asked to maintain a daily diary. The diaries focused on general health, the presence of LBP episodes, and confinement to the house due to LBP. Experimental group diaries also covered problems with the exercise program. Experimental subjects were given an abdominal muscle strength testing device¹⁵ as a compliance enhancing strategy, and were asked to report their abdominal muscle strength once a week on their diary sheet.

Control subjects received the back education classes only. Unlike experimental subjects, they were not asked to measure their abdominal muscle strength.

Outcome measures. Measurements were taken at baseline, 6, 12, and 24 months by specially trained independent assessors. The primary outcome was the percentage of subjects with at least one episode of LBP reported during the 24 month experimental period. Secondary outcome measures were abdominal muscle strength and grip strength.

LBP was defined as acute pain in the lower back (below the belt line and above the gluteal sulcus) that was not due to a vertebral fracture, cancer, infection, or pain referred from other pelvic structures. An episode of LBP was defined as continuous or intermittent pain resulting in moderate to severe limitation of function lasting more than 2 days.

The initial questionnaire included 14 questions that focused on information regarding present LBP or LBP in the previous 5 years. Other questions asked about general health, lifestyle, type of work, physical effort required at work, tobacco and alcohol use, and regular exercise. At 6, 12, and 24 months, a similar followup questionnaire was administered, seeking information on LBP episodes in the intervening period.

The independent assessors had non-medical backgrounds, were generally blind to the objectives and results of the study, and were selected on the basis of their interpersonal skills, previous work record, and accuracy of reporting. They measured tender points, leg length inequality, grip strength, abdominal muscle strength, height (cm) and weight (kg). Any adverse effects resulting from abdominal muscle strength testing, such as soreness in the abdominal muscles, pain in the lower back or other regions, were to be reported the following day. All physical measures, with the exception of the tender point count and leg length inequality, were repeated at 6, 12, and 24 months.

For the tender point assessment a subset of 8 sites (4 points bilaterally) included in the 1990 American College of Rheumatology criteria were assessed for tenderness and graded from 0 to 4¹⁷.

Leg length inequality was reported in cm, and the limb measured from the anterior superior iliac spine to the tip of the medial malleolus with the subject supine, and hip and knee joints in a neutral position.

For grip strength, subjects were assessed while sitting on a standard height chair with the back supported, feet flat on the floor, and heels resting against the legs of the chair. A specially modified grip strength manometer previously tested by our group for its validity and reliability (unpublished data) was used. The device incorporated an air pressure gauge (Bourdon Gauge MB631/4, Sarnia Piping Specialties Ltd., London, ON, Canada) with a scale that provides a minimum reading of 10 mm Hg and a maximum reading of 760 mm Hg, to avoid the ceiling effect of the modified sphygmomanometer. The gauge was attached to an adult size bladder of a standard sphygmomanometer that was folded into 3 equal parts, and secured in a sewn cotton bag. The system was inflated to 80 mm Hg to reduce creases in the bladder, and then deflated to a baseline pressure reading of 20 mm Hg. Subjects were asked to gradually squeeze the bag as hard as possible for 3 seconds, holding it lengthwise with the rubber tubes protruding upwards. Right and left hand grip strengths were measured, with the contralateral arm resting by the side.

Abdominal muscle strength was measured using the same device as the grip strength measurement but, instead of a standard bladder, a pediatric bladder (Tycos 5089-25) enclosed in a 13 × 8 cm bag was used. The bag was mounted on an oak board 1.25 cm thick, and 13 cm × 8 cm in dimension. A semicircular wooden handle was attached to the other side for use by the observer to transmit the force to the inflated bag. The air pressure gauge was mounted at a 45° angle on the back of the wooden board, facing away from the assessor. To prepare for the test, the valve was closed tightly and the system was inflated to 20 mm Hg, the standard required to register the changes in pressure when force is applied through the handle. Subjects were lying on their back and supported on a wedge pillow 30° from horizontal, hips at 30° and knees at 90°, and feet restrained by an assessor. Subjects were then assisted by a second assessor to a 45° angle while rotating the upper trunk to the right, using one composite diagonal spiral pattern¹⁶. Upper limbs were stretched to the side of the right knee, left hand gripping the extended right wrist. The inflated cuff was placed diagonally on the front of the left shoulder and overlying the sternal fibers of the pectoralis major. An assessor applied maximal pressure gradually for 5 s, without breaking the subject's ability to hold the rotated half sit-up position. The maneuver was then repeated to the left, and the mean of the 2 measurements was the composite abdominal muscle strength value.

Height (cm) and weight (kg) measurements were taken, using a standard scale (Detecto Physician's Scale 8525-M063-01, Detecto Scale Company, Webb City, MO, USA).

Subject safeguards. Informed consent was obtained in all instances prior to inclusion in the trial. Attending physicians of experimental subjects were asked whether abdominal exercises would be harmful. Subject and data confidentiality were maintained by using both a respondent number, given to all subjects who completed Phase I of the study, and a subject number, given to those subjects who were admitted to the trial.

Sample size. Of our baseline survey of 1236 faculty and staff at University of Western Ontario, 69% stated they had experienced an episode of LBP in the previous 5 years. From our estimated 5 year incidence of LBP of 69%, we would expect the 2 year incidence to be 28%. Using power sample size programs¹⁸, a sample size estimate for a 2 tailed chi-squared test (at $\alpha = 0.05$ and $\beta = 0.10$) required to detect a 15% reduction in risk of LBP or a 50% relative risk reduction (from 30% control to 15% experimental), with an estimated 15% 2 year dropout rate and accommodating for stratification factors, we set a goal of 400 subjects for the study.

Statistical technique. A one month window was employed for missed assessments. If the subject was unable to attend within that time period, the assessment was considered missed and entered into the computer as such. Any losses to followup assessment data due to moving, death, or voluntary

withdrawal were documented and included in the analysis. Subjects who withdrew during the 2 year assessment period or who deviated from the protocol were asked if they would agree to participate in the assessments only, at the intervals set out in the protocol, for the purposes of the intent-to-treat analysis. Losses to followup and protocol deviations were excluded from the efficacy analysis.

All data were entered, verified, and coded using FoxPro¹⁹ and analyzed using SAS²⁰. The questionnaires were reviewed by independent assessors for completeness. If abdominal muscle strength data were missing due to inability to perform the test, a value of 10 mm Hg was substituted. If the abdominal muscle strength test was not attempted, the subject was excluded from subsequent analyses. Probability values were obtained from chi-squared, logistic regression, Student's t test, and analysis of variance (ANOVA). The logistic model and ANOVA were adjusted for the 3 stratification factors and baseline covariates of risk of LBP in the last 5 years, abdominal muscle strength, and grip strength.

When outcome data were missing, imputation at 6, 12, and 24 months was modelled with a logistic regression using the design strata as well as treatment group, and stepwise selected those baseline factors that predicted whether the LBP outcome was available at the 5% level of significance. Imputed results were combined with complete results for the 3 time points^{21,22}.

RESULTS

From August 1988 to March 1993, 1890 subjects were screened for eligibility criteria (Figure 1). Of the 509 eligible subjects, 402 were allocated at random: 203 to abdominal exercise and back education, and 199 to back education alone. Fifty-two subjects (13%) were excluded from the efficacy analysis due to protocol deviations. Of the 52 subjects excluded due to protocol deviations, one had upper third muscle strength, 2 had middle third muscle strength, 5 had pain and limitation of shoulder, neck, rib cage, upper limb and lower limb function, 14 had LBP at the time of randomization, 5 were over the age limit, one missed the education program, 4 had heart and lung disease, 2 were unable to work in past 2 years due to LBP, and 18 were assigned to an incorrect stratum. A further 101 subjects did not complete the 24 month assessment (Figure 1). Among losses to followup, 99 subjects refused to continue with the study and 9 others moved from the study site (Table 1). There were more than twice as many refusals to continue with the study among experimental subjects than control subjects ($\chi^2_1 = 21.460$, $p < 0.0001$). There was no death of any randomized subjects.

The 2 groups had similar baseline characteristics (Table 2), with the exception of a 12.7% higher 5 year prevalence of LBP among experimental versus control subjects. We considered this a clinically important difference, which was also statistically significant ($\chi^2 = 4.638$, $P_2 = 0.03$).

At 6, 12, and 24 months, there were no statistically significant differences in the percentage of subjects with at least one LBP episode between experimental and control subjects (Table 3A and Figure 2A). There was a 2.9% risk difference in favor of experimental subjects at 6 months, however, this benefit was reversed at 12 (-1.2%) and 24 (-4.2%) months. There were no significant changes in abdominal muscle strength and grip strength at 12 and 24

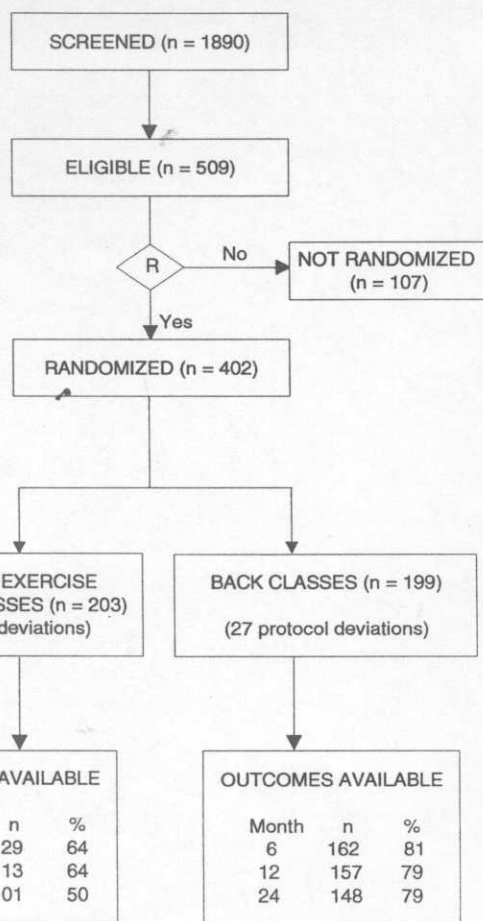


Figure 1. Progress through the stages of the trial including flow of participants, withdrawals, and timing of primary and secondary outcome measures.

Table 1. Losses to followup.

Factor	Months			Total	Chi-squared ₁ (p)
	6	12	24		
Refusals	41	26	32	99	21.460 (< 0.0001)
Experimental	33	18	19	70	
Control	8	8	13	29	0.135 (0.7134)
Moved	3	4	2	9	
Experimental	0	3	1	4	19.185 (< 0.0001)
Control	3	1	1	5	
Total	44	30	34	108	
Experimental	33	21	20	74	
Control	11	9	14	34	

Table 2. Baseline characteristics.

Variable	Experimental (n = 203)	Control (n = 199)
Discrete (f, %)		
Male (yes)	94, 46.3	93, 46.7
Exercise regularly (yes)	112, 55.2	106, 53.3
Low back pain at present* (yes)	2, 1.0	4, 2.0
Low back pain in past 5 yrs** (yes)	110, 56.1	86, 43.4
Continuous (mean, SD)		
Age (yrs)	38.3, 9.4	38.4, 9.0
Grip strength (mm Hg)	409.2, 114.4	413.2, 110.2
Abdominal muscle strength (mm Hg)	220.5, 62.3	219.2, 59.9
Height (cm)	169.3, 9.5	169.7, 10.0
Weight (kg)	72.0, 16.4	71.9, 15.8
Tender point score	1.0, 2.3	0.9, 1.8
Leg length inequality (cm)	0.34, 0.49	0.41, 0.54

*Protocol deviations.

**Chi-squared = 4.638; P2 = 0.03.

months (Tables 4 and 5); however, control subjects showed a statistically significant, but not clinically important, improvement in abdominal muscle strength at 6 months (Table 4). The proportion of subjects returning at least 80% of their diaries at 2 years was significantly lower in the experimental group ($\chi^2_1 = 20.381$, $P2 < 0.001$).

The logistic regression model adjusting for strata and covariate combinations did not change the statistical significance of our findings at the primary endpoint of 2 years, nor at 6 months and one year. Subgroup analysis suggests that young males in both groups who participate in general exercise twice a week or more are less likely to experience LBP episodes. This finding was consistent at 6, 12, and 24 months. Young males in the experimental group who did general exercise less than twice a week had fewer episodes of LBP at 6, 12, and 24 months than their control cohorts, suggesting that abdominal exercise in that group may be protective. Paradoxically, older males in both groups who did general exercise twice a week or more reported a disproportionately greater number of LBP episodes at 6, 12, and 24 months than did their young male counterparts, indicating that regular exercise in that age group may be an important risk factor. Older males who did general exercise less than twice a week fared better in the experimental group at 6 months than did the control group. This benefit, however, was reversed at 12 and 24 months. A similar pattern was noted for older females.

The imputed outcomes and complete outcomes for LBP episodes are shown in Table 3B and Figure 2B. All 3 time points — 6, 12, and 24 months — now contain results for the 203 experimental subjects and the 199 control subjects. These results suggest that at 6 months, the experimental group had a 4.8% lower rate of LBP than the control, although this result was not statistically significant ($P2 = 0.191$) and less than the 15% difference suggested as clinically

Table 3A. Low back pain episodes at 6, 12, and 24 months.

	6 months	12 months	24 months
Group			
Experimental (%)	17/129 (13.2)*	28/113 (24.8)	35/101 (34.7)
Control (%)	26/162 (16.1)	37/157 (23.6)	45/148 (30.4)
Total (%)	43/291 (14.8)	65/270 (24.1)	80/249 (32.1)
Relative risk	0.79	1.07	1.14
Risk difference, %	2.9	-1.2	-4.2
95% CI	-5.3, 11.0	-11.6, 9.2	-16.1, 7.6
Chi-squared	0.470	0.053	0.497
P2 value	0.493	0.818	0.481

*Entries are frequency, f/sample size, n (%).
CI: confidence intervals.

Table 3B. Imputed low back pain episodes at 6, 12, and 24 months.

	6 months	12 months	24 months
Group			
Experimental (%)	28/203 (13.8)*	53/203 (26/1)	72/203 (35.5)
Control (%)	37/199 (18.6)	50/199 (25.1)	64/199 (32.2)
Total (%)	65/402 (16.2)	103/402 (25/6)	136/402 (33.8)
Relative risk	0.70	1.05	1.16
Risk difference, %	4.8	-1.0	-3.3
95% CI	-2.4, 12.0	-9.5, 7.6	-12.6, 5.9
Chi-squared	1.708	0.051	0.491
P2 value	0.191	0.821	0.483

*Entries are frequency, f/sample size, n (%).

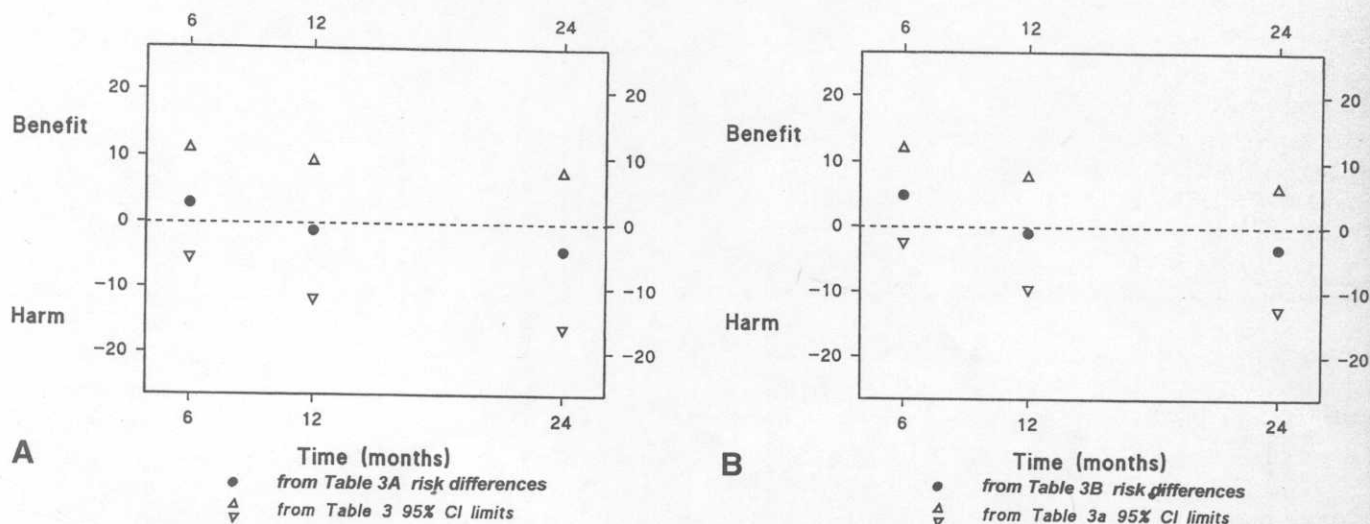


Figure 2. The results of the efficacy analysis (A) and the results of the imputation analyses (B). The left and right axes denote the percentage difference between the experimental group and the control group for rates of LBP episodes. The broken horizontal line at 0 separates the top half of the graph, where the experimental group receives more benefits than the control group, from the bottom half of the graph, where the experimental group receives more harm than the control group. The bottom and top axes label the 3 times, in months, when the LBP outcomes were assessed. ●: the point estimate of difference in LBP percentages between the experimental and control groups. △, ▽ indicate the lower and upper bounds of the 95% CI for the differences in LBP percentages between the experimental and control groups. Point estimates and their CI straddle the zero line all 3 times, suggesting that the exercise plus education intervention is similar to the education alone intervention in preventing LBP episodes.

Table 4. Abdominal muscle strength improvements (mm Hg).

Group	6 months			12 months			24 months		
	n	Mean	SD	n	Mean	SD	n	Mean	SD
Experimental	115,	4.9,	59.4	96,	9.1,	53.5	84,	2.6,	72.5
Control	143,	20.4,	55.3	148,	11.6,	55.6	131,	14.8,	63.0
Total	258,	13.5,	57.6	244,	10.6,	54.7	215,	10.0,	67.0
Difference		-15.5			-2.5			-12.2	
95% CI		-29.5, -1.4			-16.6, 11.6			-30.6, 6.2	
Student's t		2.173			0.353			1.306	
P2 value		0.031			0.720			0.193	

SD: standard deviation.

Table 5. Grip strength improvements (mm Hg).

Group	6 months			12 months			24 months		
	n	Mean	SD	n	Mean	SD	n	Mean	SD
Experimental	117,	11.6,	54.3	101,	1.3,	50.9	89,	-5.1,	51.2
Control	143,	9.2,	50.4	148,	2.4,	55.2	128,	-6.5,	55.8
Total	260,	10.3,	52.1	249,	1.9,	53.4	217,	-5.9,	53.9
Difference		2.4			-1.1			1.4	
95% CI		-10.5, 15.2			-14.7, 12.5			-13.3, 16.1	
Student's t		0.363			0.158			0.192	
P2 value		0.717			0.875			0.848	

SD: standard deviation.

cally important for the sample size justification at 24 months. At 12 months, the benefit is estimated at -1.0% ($P_2 = 0.821$); at 24 months, the benefit is estimated at -3.3% ($P_2 = 0.483$).

DISCUSSION

Because humans walk upright, their spine is different from that of any other species. There is a sharp angle at the lumbosacral junction not present in other primates and not present at birth. When stressed with the low back locked in hyperextension, pain may arise from the interaction between tight ligaments and bone. Because the 5th lumbar vertebra is not represented in the cerebral cortex, the brain can identify neither the precise site of origin of the pain, nor the nature of the injury. The pain is referred, and felt to originate in perfectly innocent structures that happen to share the same segmental nerve supply, and are represented in our body image. There is no intuitive feedback informing us if our response has been appropriate.

If this hypothesis is correct, the object of a preventive or treatment program is to avoid loading the back when it is locked in lumbar hyperextension. There should be adequately strong abdominal muscles, and they should be used to reduce the lumbar curve.

We wished to identify subjects who might be at risk because of weak abdominal muscles, and see if the rate of back pain episodes could be reduced by a program to correct the abdominal muscle weakness. We recognized that the desired reduction might not occur, because of poor compliance, because strong muscles would not necessarily correct posture, or because the underlying hypothesis was wrong.

The results of this study indicate that abdominal exercise and back education, compared to back education alone, does not appear to reduce the risk of LBP episodes over a 24 month experimental period. In fact, the exercise group had a slightly higher rate of episodes, but this difference was not clinically important or statistically significant (34.7 vs 30.4%, $P_2 = 0.481$). The 2.9% risk difference favoring the exercise group at 6 months (4.8 for the imputed data) shows that initially, abdominal exercise may be mildly protective; however, this difference was not statistically significant ($P_2 = 0.493$ complete or $P_2 = 0.055$ imputed), and was lost at 24 months (Tables 3A and B). This lack of longterm benefit from abdominal exercises may be due to the inability of experimental subjects to comply with the prescribed exercise regimen. One of the aims of the study was to provide data on the efficacy of the exercise program. Unfortunately, the results do not settle this question.

the high number of subjects who did not complete the study in the exercise group (Table 1). We expected a certain degree of noncompliance and refusals (which may be interlinked) as the daily performance of exercises, monitoring of symptoms, and measuring abdominal muscle strength can become an onerous task. As well, the repeated reminders to return diaries may have had a negative effect. To minimize noncompliance we kept the exercise program simple and short, and encouraged subjects to perform their abdominal exercises at the same time each day, so that it becomes part of their daily routine. We also gave each subject an abdominal muscle strength measuring device to monitor their abdominal muscle strength at home.

Two other points may have contributed to these results. It is possible that control subject awareness of the experimental maneuver (abdominal exercise) triggered their interest and resulted in cross contamination of the experimental maneuver. As well, since experimental subjects had a statistically significantly higher 5 year prevalence of LBP at baseline, they may have been predisposed to more frequent and severe episodes of LBP.

Paradoxically, we note that the daily diary, intended as a compliance enhancing strategy, became in itself a source of noncompliance among the exercise group subjects, providing another window on their attitudes and exercise performance. Exercise noncompliance may also explain why abdominal muscle strength did not improve from baseline. What is even more surprising is the increase in abdominal muscle strength at 6 months in the control group, which was statistically significant, but not clinically important. It is interesting to note that among exercise group subjects in this study who returned diaries and did not exercise, fatigue was cited as the most prevalent reason for not exercising today (50%), followed by poor health or injury (30%).

One might also conclude that the education in the control group had little effect, since the observed LBP prevalence at 24 months was 30.4% (Table 3A), while our sample size assumption for the control group was 30%. Our result confirms a more recent finding of noncompliance with an educational program. While knowledge increased, behavior did not change¹⁴. Perhaps the combination of a weak intervention and disproportionate noncompliance meant we could not detect the effect of the exercise program. Also, the diary system, the provision of muscle testing devices, and exercise followup by physical therapists at 6, 12, and 24 months may not have been enough. It is possible that more frequent exercise reviews and a more intensive exercise program could have yielded better results, but at a considerably higher cost.

Imputation of results is now common in the statistical literature^{21,22}, but infrequent in the clinical literature. Hence, readers of the clinical literature may view our use of imputation with suspicion. The imputed results suggest that the

modest at 6 months and has been eliminated by 12 and 24 months (Table 3B). The imputed results when compared to the complete results suggest that the benefit of exercise and education over education alone may be larger than we observed in the complete data. Clearly, any loss of outcome data, regardless of how it occurs, will reduce the statistical precision of the trial, and may also introduce bias if the losses are differential by treatment group²³.

Although the benefits of exercise and education over education alone to prevent episodes of low back pain appear to be small, our result is not yet a convincing answer to our original question. A larger trial with more attention to enhanced compliance should be conducted to eliminate the possibility that this result is a random variation. Possibly a metaanalysis of the results of this trial combined with future trials will be needed to show the benefits of exercise to prevent low back pain.

ACKNOWLEDGMENT

We are indebted to study subjects for their participation, to Joanna Crabb, Pat Darling, Jane Farrell, and Tracy O'Rourke for their research assistance; to Lisa Wilson, Ann Rennie, Kelly Welch, Kristen Darling-Lightfoot, Lynda McClatchie, Karen Wassell, Robert Stevens, Mike Murphy, Jennifer MacNab, Guissepe Parisi, Doug Burgoyne, Ian Ritchie, Vickie Balvert, and John Syrovoy, who acted as independent assessors; to Larry Stitt for data management; to Eric Duku for graphics and to Peter Aker for his advice and assistance in the preparation of the manuscript.

REFERENCES

1. Frymoyer JW. Back pain and sciatica. *N Engl J Med* 1988;318:291-300.
2. Frank A. Low back pain. *BMJ* 1993;306:901-9.
3. Spitzer WO. Scientific approach to the assessment and management of activity-related spinal disorders. Quebec Task Force on Spinal Disorders. *Spine* 1987;12 Suppl 1:S1-559.
4. Liebenson CS. Pathogenesis of chronic back pain. *J Manipulative Phys Ther* 1992;15:299-308.
5. Nachemson AL. Prevention of chronic back pain. *Bull Hosp J Dis* 1984;44:1-15.
6. Dwyer AP. Backache and its prevention. *Clin Orthop* 1987;222:35-43.
7. Frymoyer JW. Can low back pain disability be prevented? *Baillieres Clin Rheumatol* 1992;6:595-606.
8. Lahad A, Malter AD, Berg AO, Deyo RA. The effectiveness of four interventions for the prevention of low back pain. *JAMA* 1994;272:1286-91.
9. Linton SJ, Bradley LA, Jensen I, Spangfort E, Sundel L. The secondary prevention of low back pain. *Pain* 1989;36:197-207.
10. Kellet KM, Kellet DA, Nordham LA. Effects of an exercise program on sick leave due to low back pain. *Phys Ther* 1991;4:283-93.
11. Gundewall B, Liljequist M, Hansson T. Primary prevention of back symptoms and absence from work. *Spine* 1993;18:587-94.
12. Donchin M, Woolf O, Kaplan L, Floman Y. Secondary prevention of low back pain: a clinical trial. *Spine* 1990;15:1317-20.
13. Morris JM. Biomechanics of the spine. *Arch Surg* 1973;107:418-23.
14. Daltroy LH, Iversen MD, Larson MG, et al. A controlled trial of an education program to prevent low back injuries. *N Engl J Med* 1997;337:322-8.

15. Helewa A, Goldsmith CH, Smythe HA. Measuring abdominal muscle weakness in patients with low back pain and matched controls. A comparison of 3 devices. *J Rheumatol* 1993;20:1539-43.
16. Voss DE, Ionta MK, Myers BJ. Proprioceptive neuromuscular facilitation: patterns and techniques. 3rd ed. Philadelphia: Harper & Row; 1985.
17. Smythe HA, Buskila D, Gladman DD. Performance of scored palpation, a point count, and dolorimetry in assessing unsuspected nonarticular tenderness. *J Rheumatol* 1993;20:352-7.
18. Dupont WD, Plummer WD. Power and sample size calculations: a review and computer program. *Controlled Clin Trials* 1990;11:116-28.
19. Microsoft FoxPro Relational Database Management System for MS-DOS and Windows: Language Reference. Perrysberg, Ohio: Microsoft Corporation; 1993.
20. SAS Institute Inc. SAS/STAT User's Guide, Version 6, 4th ed, Vol 1 & 2. Cary, NC: SAS Institute; 1989.
21. Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley; 1987.
22. Little RJA, Rubin DB. Statistical analysis with missing data. New York: John Wiley; 1987.
23. Meinart CL, Tonascia S. Clinical trials: design, conduct and analysis. New York: Oxford University Press; 1986.