

Low back pain in general practice: cost-effectiveness of a minimal psychosocial intervention versus usual care

Petra Jellema · Nicole van der Roer · Daniëlle A. W. M. van der Windt · Maurits W. van Tulder · Henriëtte E. van der Horst · Wim A. B. Stalman · Lex M. Bouter

Received: 12 April 2007 / Accepted: 1 July 2007 / Published online: 21 July 2007
© Springer-Verlag 2007

Abstract An intervention that can prevent low back pain (LBP) becoming chronic, may not only prevent great discomfort for patients, but also save substantial costs for the society. Psychosocial factors appear to be of importance in the transition of acute to chronic LBP. The aim of this study was to compare the cost-effectiveness of an intervention aimed at psychosocial factors to usual care in patients with (sub)acute LBP. The study design was an economic evaluation alongside a cluster-randomized controlled trial, conducted from a societal perspective with a follow-up of 1 year. Sixty general practitioners in 41 general practices recruited 314 patients with non-specific LBP of less than 12 weeks' duration. General practitioners in the minimal intervention strategy (MIS) group explored and discussed psychosocial prognostic factors. Usual care (UC) was not protocolized. Clinical outcomes were functional disability (Roland–Morris Disability Questionnaire),

perceived recovery and health-related quality of life (EuroQol). Cost data consisted of direct and indirect costs and were measured by patient cost diaries and general practitioner registration forms. Complete cost data were available for 80% of the patients. Differences in clinical outcomes between both the groups were small and not statistically significant. Differences in cost data were in favor of MIS. However, the complete case analysis and the sensitivity analyses with imputed cost data were inconsistent with regard to the statistical significance of this difference in cost data. This study presents conflicting points of view regarding the cost-effectiveness of MIS. We conclude that (Dutch) general practitioners, as yet, should not replace their usual care by this new intervention.

Keywords Low back pain · Cost-effectiveness · Minimal intervention strategy · Psychosocial factors · General practice

P. Jellema (✉) · D. A. W. M. van der Windt · H. E. van der Horst · W. A. B. Stalman
Department of General Practice, VU University Medical Center,
Van der Boechorststraat 7, 1081 BT Amsterdam,
The Netherlands
e-mail: p.jellema@vumc.nl

P. Jellema · N. van der Roer · D. A. W. M. van der Windt · M. W. van Tulder · H. E. van der Horst · W. A. B. Stalman · L. M. Bouter
EMGO Institute, VU University Medical Center,
Amsterdam, The Netherlands

D. A. W. M. van der Windt
Primary Care Sciences Research Centre, Keele University,
Keele, Staffordshire, UK

M. W. van Tulder
Institute for Health Sciences, Faculty of Earth and Life Sciences,
VU University, Amsterdam, The Netherlands

Introduction

Low back pain (LBP) affects a large number of people each year. Lifetime prevalence rates range from 49 to 70% [45]. LBP causes not only great discomfort, but also great economic loss due to work absenteeism [28, 45]. In the UK, LBP is one of the most expensive conditions for which an economic analysis has been carried out [28]. Economic evaluations in which both the costs and clinical outcomes of two or more interventions are compared, are becoming increasingly important, as health care expenditures rise while budgets remain limited [44]. The importance of economic evaluations is illustrated by the fact that some authors suggest that an intervention might be implemented when it is less effective but saves substantial costs [8].

As especially chronic LBP is associated with substantial costs to society [15], a large amount of costs can be saved by interventions that prevent acute LBP becoming chronic. The rationale is that it will be more cost-effective to address a wider target population early with simple low-cost interventions than to expend considerable time and resources on rehabilitating the smaller group of back pain patients who have become incapacitated by chronic pain. As psychosocial factors have been shown to play an important role in the transition from acute to chronic back pain [25, 33], one may assume that early interventions focusing on these factors prevent chronicity.

We tested this assumption by conducting a cluster-randomized clinical trial in general practice, comparing a minimal intervention strategy (MIS) aimed at psychosocial factors for patients with (sub)acute LBP to usual care (UC) by the GP, which was not standardized. Our theory on the working mechanisms of MIS was that identification and discussion of psychosocial factors would lead to modification of these factors, eventually leading to better functioning. Unfortunately, MIS appeared to be no more effective than UC in improving the following clinical outcomes: the degree of functional disability, the recovery rate and the number of patients on sick-leave due to LBP [20]. These findings are in line with Linton and Andersson [26], who showed that their cognitive-behavioral intervention was not more effective than usual care in reducing the degree of back pain and generic function status. However, their intervention was effective in reducing the number of visits to a physician for spinal pain and number of days of sick-leave, implicating time savings for physicians and thus substantial cost savings for society. These results may indicate an increase in coping or self-care with the pain; patients who received the psychosocial intervention, less often visited a physician and had less days of work while they had the same degree of functional disability as the patients who received usual care. These promising results, in combination with the fact that self-care with the pain was also a goal in our MIS, the recent emphasis on health care budgets, and the call for more high quality economic evaluations on the cost-effectiveness of treatments for LBP [43] stimulated us to conduct an economic evaluation from a societal perspective with a follow-up of 1 year. We hypothesized that MIS would be cost-effective compared to UC.

Materials and methods

Study design

The study is designed as a full economic evaluation alongside a cluster-randomized controlled trial and was

approved by the Medical Ethics Committee of the VU University Medical Center in Amsterdam, the Netherlands.

Randomization and training sessions

Randomization took place at the level of the general practice in blocks of four practices, according to a random numbers table prepared before recruitment of general practitioners. General practitioners were informed about their allocation after they had given final consent to participation. Twenty practices (28 GPs) were randomized to the MIS group and 21 practices (32 GPs) to the UC group. The GPs randomized to the MIS group received two training sessions of 2.5 h each which were given by a GP (HvdH) with extensive expertise in development and training of psychosocial interventions. The training consisted of theory, role-playing and feedback on the practiced skills. In addition, a treatment manual was provided. The contents of the training sessions and its evaluation by GPs have been described in more detail elsewhere [21].

Patients and interventions

Participating GPs were asked to select ten consecutive patients who consulted them for LBP. Inclusion criteria were age 18–65, non-specific LBP of less than 12 weeks' duration (i.e. (sub)acute LBP) or an exacerbation of mild symptoms, and sufficient knowledge of the Dutch language. Exclusion criteria were specific LBP (i.e. LBP caused by specific pathological conditions), LBP currently treated by another healthcare professional, and pregnancy. Patients, but not their GPs, were kept unaware that two different interventions were studied.

Patients received a minimal intervention strategy (MIS) or usual care (UC). The MIS was aimed at identification and discussion of psychosocial prognostic factors. The MIS consultation lasted about 20 min and consisted of three phases: exploration, information and self care. During the exploration phase, the GP explored the presence of psychosocial prognostic factors by asking standardized questions that could be rephrased to fit the style of communication of the doctor and the patient. The following psychosocial prognostic factors were explored: the patient's own ideas on the cause of their LBP, fear avoidance beliefs, worries/distress, pain catastrophising, pain behaviors and reactions from the social environment (family, friends, work). In the information phase the GP provided general information on the cause, course and (im)possibilities of treatment of LBP, thereby giving specific attention to psychosocial factors identified in the exploration phase. Finally, in the self care phase, the GP and patient set specific goals on resuming activities or work. Follow-up consultations were not protocolized, but

we advised GPs to make an appointment for a follow-up visit in case they identified obstacles to recovery and suspected an increased risk of chronic LBP.

GPs in the UC group provided care as usual. We did not protocolize the content and number of UC consultations, and assumed that GPs would generally follow the guideline for LBP of the Dutch College of General Practitioners [13]. For acute LBP (<6 weeks' duration) this guideline advises a wait and see policy. For subacute LBP (6–12 weeks' duration) the guideline advises referral for physical therapy in the case of persistent functional disability. Explicit guidance on psychosocial factors is lacking. The contents of both interventions have been described in more detail elsewhere [20].

Data collection

Clinical outcomes

Baseline data were collected during a home visit by a research assistant, while follow-up data after 12 months were collected using postal questionnaires. Primary clinical outcome measures were functional disability, perceived recovery, and health related quality of life. Functional disability was measured at baseline and after 12 months by the Roland–Morris disability questionnaire (0–24) [35]. Perceived recovery was scored by the patient on a 7-point Likert scale (very much/much/slightly improved, no change, slightly/much/very much worse) after 12 months [42]. As a score of at least “much improved” has been denoted a minimal clinically important change [32], patients were a priori defined as recovered if they reported at least “much improvement”. Health related quality of life was measured at baseline and after 3, 6 and 12 months by the EuroQol (0–1) [12], covering five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The EuroQol scores were transformed into utilities using a representative British sample and time trade-off methods. The utilities were then multiplied by the amount of time a patient spent in this particular health state, with transitions between health states linearly interpolated [7]. This results in quality adjusted life years (QALYs).

In order to describe the study population characteristics in more detail and to compare baseline similarity of both intervention groups, five other outcome measures were assessed at baseline: (1) pain severity during the day (0–10) [27]; (2) perceived general health (1–5), using the first question of the subscale “general health perceptions” of the short form health survey (SF-36) [47]; (3) fear avoidance beliefs, using the 4-item physical activity subscale of the fear avoidance beliefs questionnaire (FABQ, 0–24) [46]; (4) catastrophising thoughts, using the 6-item subscale of the coping strategies questionnaire (CSQ, 0–36) [36] and (5)

distress, measured by the 16-item subscale of the four-dimensional symptom questionnaire (4DSQ, 0–32) [37].

Cost data

The economic evaluation was conducted from a societal perspective, indicating that all costs and consequences of the competing interventions are taken into account regardless of who pays for or benefits from them [9]. A societal perspective incorporates direct health care costs, direct non-health care costs and indirect costs due to LBP. All cost data were collected by prospective cost diaries [16] that patients completed for the periods baseline to 3 months, 3–6 months, 6–9 months, and 9–12 months. In addition, GPs were contacted after 12 months follow-up to provide information on follow-up consultations due to LBP in the last year (date, referrals, medication prescribed by physician). GPs used their medical records to complete these registration forms.

Table 1 summarizes the cost categories, and the prices and sources used for valuing. “Medication” in Table 1 includes both the over-the-counter medication and medication prescribed by physicians, while self-prescribed alternative interventions are included in “complementary care”. Costs of absenteeism from paid labor due to LBP were calculated according to the friction cost approach [23]. This approach is based on a mean income of the Dutch population according to age and gender, and defines a friction period as 154 days [31]. As most cost data were collected during the year 2002, prices were adjusted using consumer price index figures.

Statistical analysis

Firstly, baseline similarity was studied. Secondly, we compared baseline characteristics of patients with complete cost data to those with incomplete cost data by using logistic regression analysis. Thirdly, clinical outcomes and total costs were compared. As cost data are characterized by large variation and irregular distributions and as a complete cost dataset was available for 80% of our participants (250/314), we decided that our primary analysis would be a complete case analysis. Differences between groups (MIS minus UC) were calculated for the clinical outcomes: (1) functional disability, by calculating change scores between baseline and 12 months follow-up; (2) perceived recovery at 12 months follow-up; and (3) health-related quality of life gained over 12 months (i.e. QALYs). Student's *t* tests were used to analyze the change scores between the treatment groups for functional disability and quality of life, and a Chi-square test for perceived recovery. To compare costs between the two groups, confidence intervals (CIs) for the mean differences in costs were cal-

Table 1 Prices used for valuing resources (year 2002)

Resources	Euro
Direct health care	
Primary care costs	
General practitioner (consultation) ^a	19.8
MIS-consultation, including intervention costs (consultation) ^b	40.9
Physical therapist (treatment session) ^a	22.3
Manual therapist (treatment session) ^c	30.8
Exercise therapist (treatment session) ^a	22.5
Back school (treatment session) [18]	66.3
Chiropractor (treatment session) ^d	40.0
Physiofitness (treatment program, 10 sessions) ^b	120.0
Professional homecare (per hour) ^a	21.3
Psychologist (treatment session) ^d	75.0
Secondary care costs	
Outpatient appointment ^a	54.8
Hospitalization general hospital (per day) ^a	330.0
Surgery (per hour) ^b	884.7
Radiograph ^b	48.2
MRI-scan ^b	212.9
Medication	^e
Direct non-health care	
Complementary care	^d
Informal care (per hour) ^a	8.1
Equipment aids	^f
Indirect costs	
Absenteeism from paid work	^g
Absenteeism from unpaid work ^a	8.1

MIS Minimal intervention strategy, MRI Magnetic resonance imaging

^a Price according to Dutch guidelines [30]

^b Calculated unit costs

^c Price according Dutch Central Organisation for Health Care Charges [10]

^d Price according to professional organisation or health care provider

^e Price of medication prescribed by the GP according to Royal Dutch Society for Pharmacy [48]; price of over-the-counter medication according to cost diaries

^f Price according to cost diaries

^g Costs based on mean income of Dutch population according to age and sex [23, 31]

culated using bias-corrected and accelerated bootstrapping (2000 replications) [11]. Bootstrapping incorporates drawing samples with replacement and is a preferred method for the analysis of cost data, as it uses the observed distributions of the data without making assumptions about the shape of the distribution [38]. Fourthly, cost-effectiveness analyses were performed. Incremental cost-effectiveness ratios were calculated in which the mean difference in total costs (MIS minus UC) was divided by the mean difference in improvement on the clinical out-

comes (MIS minus UC). Uncertainty around the ratios was calculated using the bias-corrected percentile bootstrapping method (5,000 replications) [5] and plotted on a cost-effectiveness plane.

Sensitivity analyses

A complete case analysis has some disadvantages. Due to the missing data the power of the analysis is reduced and bias may be introduced due to selective drop-out. Imputation can be used to replace missing data by statistical estimates of the missing values. To explore the robustness of our primary analysis, we performed two sensitivity analyses in which we imputed missing data: in one analysis we imputed all missing cost data, and in the other we imputed only missing days of absenteeism. Imputation was done using the Expectation Maximization algorithm (SPSS 10.1).

Results

Between September 2001 and April 2003, 314 patients were enrolled in our study: 143 in the MIS group and 171 in the UC group. Table 2 shows that baseline characteristics of GPs and patients were largely similar for the two groups. Less than 9% of all patients withdrew from the study during follow-up. Reasons were ‘no time and no complaints anymore’ (MIS $n = 1$, UC $n = 4$), “burden too high due to psychological problems” (UC $n = 3$) or “unknown” (MIS $n = 10$; UC $n = 8$). The flow chart of this study, including information on refusals, exclusions and drop-outs has been published in our previous paper [20].

For 116 patients (81%) in the MIS group and 134 patients (78%) in the UC group complete cost data were available. Logistic regression analysis comparing baseline characteristics of patients with complete cost data to those with incomplete cost data showed that patients with incomplete cost data on average were younger and scored higher on distress at baseline.

Clinical outcomes

After 1 year follow-up both the groups showed similar improvements in clinical outcomes. Sixty-nine percent of the patients were defined as recovered after 1 year. The difference between both the groups in mean improvement on functional disability was -0.74 points on the RDQ (95% CI, -2.31 to 0.83), and -2% (95% CI, -14 to 10%) on recovery rate. Over the follow-up period of 1 year the mean difference in quality of life was 0.004 QALYs (95% CI, -0.04 to 0.03). All differences favored UC, but were neither clinically relevant nor statistically significant (Table 5).

Table 2 Baseline characteristics of general practitioners and patients

GPs	MIS <i>n</i> = 28	UC <i>n</i> = 32
Demographic characteristics		
Age (years) ^a , mean (SD)	43.0 (7.2)	45.7 (7.4)
Gender, female	6/28	12/32
Number of included patients per GP		
0	1/28	5/32
1–5	17/28	14/32
>5	10/28	13/32
Patients		
<i>n</i> = 143		
<i>n</i> = 171		
Demographic characteristics		
Age (years), mean (SD)	43.4 (11.1)	42.0 (12.0)
Gender, % female	47.6	47.4
Nationality, % Dutch	97.2	97.7
Health insurance, % public	70.6	67.8
Educational level and work status		
Education level ^a , %		
Primary	35.0	33.1
Secondary	46.2	52.7
College, university	18.8	14.2
Paid job, % yes	81.8	81.3
Sick leave because of LBP ^a (among the working population), % yes	34.8	41.0
Characteristics of LBP		
Duration current episode (days), median (IQR)	11 (5–21)	14 (7–21)
Frequency of LBP episodes last year, %		
1 or 2 episodes	58.0	60.8
3 or more episodes	19.6	18.7
Exacerbation of persisting LBP	22.4	20.5
Pain radiating below knee ^a , %	12.8	14.6
Clinical outcomes, primary measures		
Functional disability (RDQ, 0–24), mean (SD)	11.7 (5.4)	12.2 (5.0)
Health related quality of life (EuroQol), mean (SD)	0.61 (0.25)	0.61 (0.24)
Clinical outcomes, secondary measures		
Pain severity during the day (0–10) ^b , mean (SD)	4.9 (2.0)	4.8 (2.0)
Severity of the main complaint (0–10) ^a , mean (SD)	7.0 (1.9)	6.8 (2.0)
Perceived general health (SF-36, 1–5) ^a , mean (SD)	2.7 (0.8)	2.8 (0.8)
Clinical outcomes, psychosocial measures		
Fear-avoidance beliefs (FABQ, 0–24), mean (SD)	14.3 (5.6)	15.3 (5.2)
Catastrophising thoughts (CSQ, 0–36), mean (SD)	10.3 (6.6)	11.2 (6.9)
Distress (4DSQ, 0–32) ^b , mean (SD)	8.3 (7.0)	9.5 (7.3)

GP General practitioner, MIS Minimal intervention strategy, UC Usual care, *n* number, SD Standard deviation, IQR Inter quartile range (25th–75th percentile), LBP Low back pain, RDQ Roland–Morris disability Questionnaire, SF-36 Short-form 36, FABQ Fear avoidance and beliefs questionnaire, CSQ Coping strategies questionnaire, 4DSQ 4D symptom questionnaire (higher scores means more functional disability, worse health, more fear-avoidance, more catastrophising or more distress)

^a Two missing values

^b One missing value

Cost data

Table 3 lists per patient the mean utilization of resources (i.e. health care, help, absenteeism). In both the groups, resource utilization was low and largely similar. Only two statistically significant differences were found. In the year following randomization patients in the MIS group had more

consultations with a GP (MIS 2.7 vs. UC 0.9), excluding the consultation leading to recruitment but including the 20 min consultation aimed at psychosocial measures as specified by the study protocol. Patients in the UC group reported more consultations with a manual therapist (MIS 0.1 vs. UC 0.4) but the proportions of patients who received such a treatment (MIS 2.6% vs. UC 9%) were very low.

Table 3 Mean resource use (SD) per patient ($n = 250$) for MIS and UC during 12 months follow-up, and the percentage of patients who made use of that specific resource

Resources	MIS ($n = 116$)		UC ($n = 134$)	
Direct health care				
Primary care				
General practitioner (consultations)*	2.7 (1.0)	99.1%	0.9 (1.2)	44%
Physical therapy (treatment sessions)	2.3 (5.8)	17.2%	3.2 (5.8)	35.1%
Manual therapy (treatment sessions)*	0.1 (0.7)	2.6%	0.4 (1.7)	9%
Exercise therapy (treatment sessions)	0.3 (1.8)	2.6%	0.9 (3.5)	9%
Back school (treatment sessions)	0.5 (0.6)	0.9%	0 (0)	0%
Chiropractor (treatment sessions)	0 (0)	0%	0.1 (0.9)	2.2%
Physiofitness (treatment programs)	0.1 (0.2)	6%	0.1 (0.2)	4.5%
Professional home care (hours)	0 (0)	0%	0.6 (6.5)	0.7%
Psychologist (treatment sessions)	0.1 (0.8)	0.9%	0 (0)	0%
Secondary care				
Outpatient visit (number)	0.1 (0.4)	4.3%	0.2 (0.7)	7.5%
Hospitalization (days)	0.02 (0.2)	0.9%	0.01 (0.1)	0.7%
X-ray (number)	0.1 (0.3)	5.2%	0.1 (0.4)	9.7%
MRI-scan (number)	0.01 (0.1)	0.9%	0.03 (0.2)	2.2%
Medication (% yes)	–	74.1%	–	73.9%
Direct non-health care				
Complementary care				
Homeopathy (treatment sessions)	0.04 (0.3)	1.7%	0 (0)	0%
Osteopathy (treatment sessions)	0.1 (0.5)	3.4%	0.2 (0.3)	0.7%
Massage (treatment sessions)	0.1 (0.8)	1.7%	0.4 (3.0)	3.7%
Acupuncture (treatment sessions)	0 (0)	0%	0.04 (0.4)	1.5%
Magnetizer (treatment sessions)	0.3 (2.3)	1.7%	0.1 (1.0)	0.7%
Orthomanual therapy (treatment sessions)	0.03 (0.3)	1.7%	0.03 (0.4)	0.7%
Podiatrist	0.03 (0.2)	1.7%	0.03 (0.3)	1.5%
Informal care (hours)	0 (0)	0%	0.7 (6.3)	3%
Equipment aids (% yes)	–	10.3%	–	6.0%
Indirect costs				
Absenteeism paid work (days)	4.2 (12.1)	28.4%	8.6 (23.1)	38.8%
Absenteeism unpaid work (hours)	2.2 (10.9)	12.1%	3.3 (10.4)	17.2%

*Significant difference ($P < 0.05$)

SD Standard deviation, n number, MIS Minimal intervention strategy, UC Usual care

Table 4 shows the mean total costs in both the treatment groups and the difference in costs with 95% CI. Total indirect costs, especially absenteeism from paid work, were the largest contributor to the total costs. The difference in total costs amounted to 490 € (95% CI –987 to 92 €) in favor of the MIS group (MIS 799 €; UC 1288 €), but this difference was not statistically significant.

Cost-effectiveness

Table 5 shows the incremental cost-effectiveness ratios (ICERs) for the three outcome measures. MIS resulted in less improvement than UC, but saved money. The ICER for functioning was 690 €, indicating that per point less improvement on the RDQ MIS saved 690 €, while per percent less improvement in recovery rate MIS saved

239 €. The difference in QALY's gained during 1 year between both the groups was very small, resulting in a large ICER of 47,348 €. The large majority of the bootstrapped ICERs presented on the cost-effectiveness planes are located in the southern quadrants (Fig. 1), indicating that the costs of MIS were lower than the costs of UC.

Sensitivity analyses

Imputation of missing cost data led to a mean difference of –628 € (95% CI –1123 to –81 €) in total costs. Imputation of missing data on days of absenteeism led to a mean difference of –545 € (95% CI –1031 to –40 €) in total costs. Both differences are statistically significant and in favor of MIS.

Table 4 Mean costs (SD) in Euros per patient in the MIS and UC group and differences between both the groups during follow-up of 52 weeks

	MIS (<i>n</i> = 116)	UC (<i>n</i> = 134)	Mean difference (95% CI)
Direct health care costs	181 (287)	175 (275)	6 (–65; 73)
Primary care costs	159 (181)	152 (242)	7 (–52; 54)
Secondary care costs	22 (179)	23 (101)	–1 (–44; 35)
Medication costs	9 (18)	9 (14)	0 (–4; 4)
Direct non-health care costs	30 (75)	31 (129)	–1 (–25 ; 27)
Complementary care costs	18 (65)	21 (115)	–2 (–24; 22)
Informal care	6 (34)	8 (58)	–2 (–12; 11)
Equipment aids	5 (22)	2 (14)	3 (–2; 6)
Indirect costs	587 (1,636)	1,081 (2,463)	–495 (–921; 158)
Absenteeism paid work	569 (1,631)	1,055 (2,448)	–486 (–932; 50)
Absenteeism unpaid work	18 (89)	27 (84)	–9 (–31; 18)
Total costs	798 (1,820)	1,288 (2,594)	–490 (–987; 92)

None of the mean differences between MIS and UC are statistically significant ($P > 0.05$)

SD Standard deviation, MIS Minimal intervention strategy, UC Usual care, *n* Number, 95% CI 95% confidence interval

Table 5 Incremental cost-effectiveness ratios for functional disability, perceived recovery and health-related quality of life

Outcome measure	MIS		UC		MIS versus UC
	Costs	Effects ^a	Costs	Effects ^a	ICER
Functional disability (RDQ; 0–24) Change score between baseline and 1 year follow-up	798 € (<i>n</i> = 116)	8.378 (<i>n</i> = 116)	1,307 € (<i>n</i> = 132)	9.116 (<i>n</i> = 132)	MIS results in less improvement, but saves 690 € per RDQ point
Recovery rate Score at 1 year follow-up	798 € (<i>n</i> = 116)	68.1% (<i>n</i> = 116)	1,288 € (<i>n</i> = 134)	70.1% (<i>n</i> = 134)	MIS results in less improvement, but saves 239 € per 1%
Quality of life (EuroQol; QALY) Score over 1 year follow-up	819 € (<i>n</i> = 102)	0.833 (<i>n</i> = 102)	1030 € (<i>n</i> = 111)	0.837 (<i>n</i> = 111)	MIS results in less improvement, but saves 47,348 € per QALY

MIS Minimal intervention strategy, UC Usual care, ICER Incremental cost-effectiveness ratio, *n* Number, QALY Quality adjusted life year

^a Difference in clinical outcomes between baseline and 12 months follow-up

Discussion

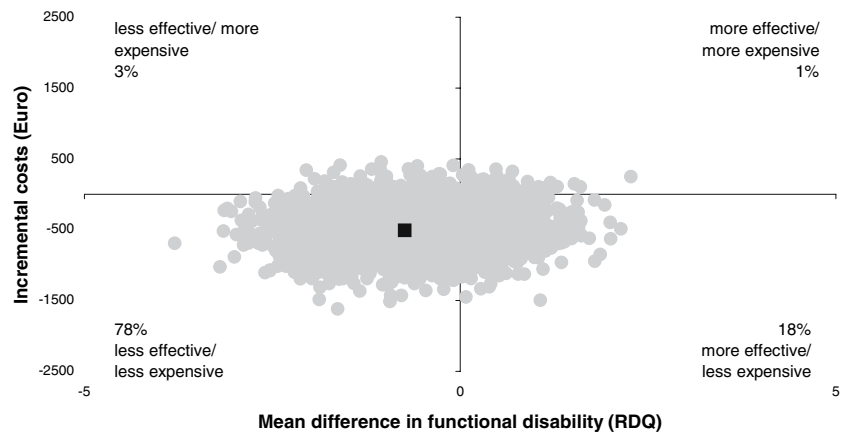
The results of our primary analysis showed no statistically significant differences in total costs or clinical outcomes between our psychosocial intervention and UC in patients with (sub)acute LBP in general practice. However, the results of our sensitivity analyses are inconsistent with those of the primary analysis. In this discussion section we will focus on the interpretation of our cost data and the methodological issues involved when interpreting cost data. In previous papers we have discussed several methodological issues involved when interpreting the clinical outcomes (e.g. the quality of the training sessions and interventions) [20, 21].

Comparison to the literature

A recent review on the cost-effectiveness of treatments for patients with LBP reported that 6 of the 17 studies concluded that the intervention of interest was more cost-effective than the control intervention [43]. Unfortu-

nately, no definite conclusions for the cost-effectiveness of a specific intervention could be given as the number of economic evaluations per type of intervention was limited. More recently the cost-effectiveness analyses of two trials on psychosocial primary care interventions for LBP have been published. The UK BEAM Trial compared spinal manipulation, exercise classes incorporating cognitive behavioral principles, spinal manipulation plus exercise classes, and best care [40, 41]. Although the authors conclude that spinal manipulation is a cost-effective addition to “best care”, as it generates 0.04 more QALYs for an extra 279 € [37], others question this conclusion [6, 39]. Another study in primary care compared advice directed towards promoting self-management and modifying beliefs and behavior, and routine physiotherapy treatment with advice only [14, 34]. The authors conclude that advice should be considered as the first-line treatment, as out-of-pocket expenses of patients receiving routine physiotherapy were significantly higher (41£, 59 €), while differences in either QALYs (0.02) or National Health Service costs (20£, 29 €) were not significantly different

Fig. 1 Cost-effectiveness plane for functional disability (RDQ) in which MIS is compared to UC



[34]. Both the above mentioned studies found lower QALYs and higher health care costs over 12 months than we found in our study. The shorter duration of the LBP episode at baseline of our patients and, thus, their more favorable prognosis may (partly) explain these results. In contrast to our study, both the trials did not include indirect costs in their cost-effectiveness analyses, even though it is generally agreed upon that indirect costs associated with work absence due to LBP account for high economic costs in western societies [28, 45].

Methodological issues and secondary analyses

In our study, we used the British sample as reference population for the utilities as, until recently, these were the only reference data available in the Netherlands. Recently, a EuroQol transformation has been introduced based on a Dutch sample [24]. We decided not to use this Dutch transformation as (1) it has only been internally validated through bootstrapping methods but not yet externally validated; (2) the representativeness of the, rather small ($n = 298$), Dutch sample can be debated [4] and, finally (3) use of the Dutch transformation would complicate comparison across international studies. A comparison between the use of a Dutch or UK sample showed that, although QALYs based on the Dutch sample were slightly higher compared to the British sample, the mean differences between the groups, in this case two primary care interventions for treatment of depression, were largely the same [2].

We based our sample size calculations on demonstrating a clinically relevant difference of 2.5 points on the RDQ instead of demonstrating the cost-effectiveness [20]. Although the usefulness of sample size calculations in economic evaluations has been debated [1], the required sample size in an appropriately powered cost-effectiveness study is expected to be much larger than in a clinical effectiveness trial [3].

Work absence due to LBP was the main contributor to the total costs associated with LBP. The number of days of absenteeism was twice as high in the UC group as in the MIS group. Although not statistically significant, this finding induced us to explore whether MIS was cost-effective among patients having a paid job ($n = 93$ MIS; $n = 110$ UC). This post-hoc subgroup analysis showed that the difference in total costs amounted to -548 € (95% CI -1137 to 200 €) in favor of MIS, which was not statistically significant. As this subgroup analysis was clearly insufficiently powered, and as other early interventions have been shown to reduce sick-leave in workers [17, 19, 22] one may hypothesize that MIS may be cost-effective in an occupational setting. The sensitivity analyses with imputed data were less underpowered and, interestingly, showed statistically significant differences in total costs between groups. The question now arises which analysis we should value most.

Interpretation of the conflicting results

A complete case analysis has the advantage over an imputed analysis that no assumptions have to be made regarding imputation of data that are characterized by large variation and irregular distributions. However, a complete case analysis does not include all participating patients, which reduces the statistical power and which may introduce bias. Advocates of the (conclusions of our) complete case analysis may, among others, be (1) researchers who do not favor imputation techniques when data are available for 80% of the cases; (2) GPs, as application of MIS will increase their workload, with no evidence of improvement of the patients' functional disability, recovery rates, or pain intensity. Advocates of the (conclusions of our) sensitivity analysis may among others be (1) researchers who favor imputation techniques; (2) policy makers, who are interested in all relevant costs and effects regardless of who will pay the costs and who may benefit; or (3) managers in

workplace settings as they may be especially interested in the reduction of costs due to sick-leave.

The cost-effectiveness planes do not provide a solution for the impasse. The plane for functional disability showed that 96% of the bootstrapped cost-effect pairs, although very close to the origin of the plane, were located in the southern quadrants. Does this indicate cost-effectiveness of MIS? The answer to this question will depend on the way one defines cost-effectiveness. Obviously, when a new intervention is clinically more effective and less costly than the old intervention (south–east quadrant), there will be no debate about cost-effectiveness and implementation of the new intervention. When a new intervention is clinically more effective and more costly (north-east quadrant), a new intervention will be considered cost-effective when costs stay below a certain threshold. Dowie (2004) [8], however, questions why an intervention should not also be implemented when it is less effective but does save substantial costs, pleading for defining a threshold that extends into the south-west quadrant. This threshold can be a straight line from the north-east to the south-west quadrant, or kinked at the origin of the cost-effectiveness plane, indicating that the selling price for an unit of effect is greater than the buying price of an additional unit of effect [29].

Although there are various points of view, we conclude that (Dutch) GPs should not replace their UC by MIS in patients with (sub)acute LBP. This conclusion is based on the results of the complete case analysis, because this was our pre-planned analysis. The second reason is that the difference in costs, if any, was mainly caused by a small proportion of the population who reported sick leave due to back pain. We cannot rule out the possibility that this was a chance finding. The third reason is that MIS is a new approach in general practice. Implementation of MIS will meet many difficulties, given the lack of evidence for clinical effects, the need for training of GPs and the increase of GP workload. The final, and maybe most important reason is that our study is the first and, as far as we know, the only study that investigated the cost-effectiveness of a psychosocial intervention in patients with (sub)acute LBP in general practice.

Conclusions

Results of only one study are insufficient to establish firm evidence on the cost-effectiveness of an intervention. More studies on the cost-effectiveness of psychosocial interventions in general practice are needed, thereby taking into account some of the factors that may explain why our intervention was not more successful than usual GP care as described in previous papers [20, 21]. As yet, we conclude

that (Dutch) general practitioners should not replace their usual care by our new intervention.

Acknowledgments We thank all participating patients and GPs for their time and effort. This study was supported by a grant (No. 2200.0095) from The Netherlands Organization for Health Research and Development (ZonMw), The Hague, the Netherlands. The study was approved by the Medical Ethics Committee of the VU University Medical Center in Amsterdam, the Netherlands.

References

1. Al MJ, van Hout BA, Michel BC, Rutten FFH (1998) Sample size calculation in economic evaluations. *Health Econ* 7:327–335
2. Bosmans J (2006) Cost-effectiveness of interpersonal psychotherapy (IPT) for elderly primary care patients with major depression. In: *Cost-effectiveness of treatment of depression in primary care* (thesis). VU University Medical Center, Amsterdam, pp. 77–96
3. Briggs A (2000) Economic evaluation and clinical trials: size matters. *BMJ* 321:1362–1363
4. Buskens E (2005) Beleidsmatige beslissingen vanuit maatschappelijk perspectief: locale of nationale waardering voor gezondheidsgerelateerde kwaliteit van leven niet noodzakelijk. *Ned Tijdschr Geneesk* 149:1553–1555
5. Chaudhary MA, Stearns SC (1996) Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Stat Med* 15:1447–1458
6. Church TR (2005) United Kingdom back pain exercise and manipulation (UK BEAM) trial: manipulation alone costs more than other options so why is it recommended? *BMJ* 330:674
7. Dolan P (1997) Modeling valuations for EuroQol health states. *Med Care* 35:1095–1108
8. Dowie J (2004) Why cost-effectiveness should trump (clinical) effectiveness: the ethical economics of the South West quadrant. *Health Econ* 13:453–459
9. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005) *Methods for the economic evaluation of health care programmes*. Oxford University Press, New York
10. Dutch Central Organisation for Health Care Charges (2002) Maximum tariffs physiotherapy from 1 Jan 2002. Tariffs decision 5800–1900-02-1. CTG, Utrecht, The Netherlands
11. Efron B, Tibshirani RJ (1993) *An introduction to the bootstrap*. Chapman & Hall, New York
12. EuroQol Group (1990) EuroQol—a new facility for the measurement of health-related quality of life. *The EuroQol Group. Health Policy* 16:199–208
13. Faas A, Chavannes AW, Koes BW, Van den Hoogen JMM, Mens JMA, Smeele LJM, Romeijnders ACM, van der Laan JR (1996) NHG-Standaard Lage-Rugpijn. *Huisarts Wet* 39:18–31
14. Frost H, Lamb SE, Doll HA, Carver PT, Stewart-Brown S (2004) Randomised controlled trial of physiotherapy compared with advice for low back pain. *BMJ* 329:708
15. Frymoyer JW, Cats-Baril W (1991) An overview of the incidences and costs of low back pain. *Orthop Clin North Am* 22:263–271
16. Goossens ME, Rutten-van Molken MP, Vlaeyen JW, van der Linden SM (2000) The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol* 53:688–695
17. Hagen EM, Eriksen HR, Ursin H (2000) Does early intervention with a light mobilization program reduce long-term sick leave for low back pain? *Spine* 25:1973–1976

18. Heymans MW (2004) The cost-effectiveness and cost-utility of high intensity versus low intensity back schools in occupational care: a randomized controlled trial and economic evaluation with 12 months follow-up. In: *The cost-effectiveness of back schools in occupational care, the BOC study (thesis)*. VU University Medical Center, Amsterdam, pp 122–146
19. Indahl A, Velund L, Reikeraas O (1995) Good prognosis for low back pain when left untampered: a randomised clinical trial. *Spine* 20:473–477
20. Jellema P, van der Windt DAWM, van der Horst HE, Twisk JWR, Stalman WAB, Bouter LM (2005) Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Results of a cluster-randomised clinical trial in general practice. *BMJ* 331:84–90
21. Jellema P, van der Windt DAWM, van der Horst HE, Blankenstein AH, Bouter LM, Stalman WAB (2005) Why is a treatment aimed at psychosocial factors not effective in patients with (sub)acute low back pain? *Pain* 118:350
22. Karjalainen K, Malmivaara A, Pohjolainen T, Hurri H, Mutanen P, Rissanen P, Pakkajarvi H, Levon H, Karpoff H, Roine R (2003) Mini intervention for subacute low back pain, a randomized controlled trial. *Spine* 28:533–541
23. Koopmanschap MA, Rutten FF (1996) A practical guide for calculating indirect costs of disease. *Pharmacoeconomics* 10:460–466
24. Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Buschbach JJV (2006) The Dutch Tariff : results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 15:1121–1132
25. Linton SJ (2000) A review of psychological risk factors in back and neck pain. *Spine* 25:1148–1156
26. Linton SJ, Andersson T (2000) Can chronic disability be prevented? A randomized trial of a cognitive-behavior intervention and two forms of information for patients with spinal pain. *Spine* 25:2825–2831
27. Lundeberg T, Lund I, Dahlin L, Borg E, Gustafsson C, Sandin L, Rosen A, Kowalski J, Eriksson SVJ (2001) Reliability and responsiveness of three different pain assessments. *Rehabil Med* 33:279–83
28. Maniadakis N, Gray A (2000) The economic burden of back pain in the UK. *Pain* 84:95–103
29. O'Brien BJ, Gertsen K, Willan AR, Faulkner LA (2002) Is there a kink in consumers' threshold value for cost-effectiveness in health care? *Health Econ* 11:175–180
30. Oostenbrink JB, Koopmanschap MA, Rutten FFH (2002) Standardisation of costs: the Dutch Manual for Costing in economic evaluations. *Pharmacoeconomics* 20:443–454
31. Oostenbrink JB, Bouwmans CAM, Koopmanschap MA, Rutten FFH (2004) Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen. Geactualiseerde versie
32. Ostelo RWJG, de Vet HCW (2005) Clinically important outcomes in low back pain. *Best Practice Res Clin Rheumatol* 19:593–607
33. Pincus T, Burton AK, Vogel S, Field AP (2002) A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 27:E109–E120
34. Rivero-Arias O, Gray A, Frost H, Lamb SE, Stewart-Brown S (2006) Cost-utility analysis of physiotherapy treatment compared with physiotherapy advice in low back pain. *Spine* 31:1381–1387
35. Roland M, Morris R (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low back pain. *Spine* 8:141–144
36. Rosenstiel AK, Keefe FJ (1983) The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain* 17:33–44
37. Terluin B, van Rhenen W, Schaufeli WB, de Haan M (2004) The four-dimensional symptom questionnaire (4DSQ): measuring distress and other mental health problems in a working population. *Work Stress* 18:187–207
38. Thompson SG, Barber JA (2000) How should cost data in pragmatic randomised trials be analysed? *BMJ* 320:1197
39. Tveito TH, Eriksen HR (2005) United Kingdom back pain exercise and manipulation (UK BEAM) trial: is manipulation the most cost effective addition to best care? *BMJ* 330:674
40. UK BEAM Trial Team (2004) United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. *BMJ* 329:1377
41. UK BEAM Trial Team (2004) United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: cost effectiveness of physical treatments for back pain in primary care. *BMJ* 329:1381
42. van der Heijden GJ, Leffers P, Wolters PJ, Verheijden JJ, van Mameren H, Houben JP, Bouter LM, Knipschild PG (1999) Effect of bipolar interferential electrotherapy and pulsed ultrasound for soft tissue shoulder disorders: a randomised controlled trial. *Ann Rheum Dis* 58:530–540
43. van der Roer N, Goossens ME, Evers SM, van Tulder MW (2005) What is the most cost-effective treatment for patients with low back pain? A systematic review. *Best Pract Res Clin Rheumatol* 19:671–84
44. van der Roer N, Boos N, van Tulder MW (2006) Economic evaluations: a new avenue of outcome assessment in spinal disorders. *Eur Spine J* 15:S109–S117
45. van Tulder MW, Koes BW, Bombardier C (2002) Low back pain. *Best Pract Res Clin Rheumatol* 16:761–775
46. Waddell G, Somerville D, Henderson I, Newton M, Main CJ (1993) A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 52:157–168
47. Ware JE, Sherbourne CD (1992) The MOS-36-item Short-form Health status Survey (SF-36): 1. conceptual framework and item selection. *Med Care* 30:473–483
48. Z-index (2002) G-Standaard. The Hague, Z-index. The Netherlands www.z-index.nl