Eur Spine J (2012) 21:1290–1300
DOI 10.1007/s00586-011-2144-4

ORIGINAL ARTICLE

Cost-effectiveness of a classification-based system for sub-acute and chronic low back pain

Adri T. Apeldoorn · Judith E. Bosmans · Raymond W. Ostelo · Henrica C. W. de Vet · Maurits W. van Tulder

Received: 24 August 2011 / Revised: 12 December 2011 / Accepted: 30 December 2011 / Published online: 19 January 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract
Purpose Identifying relevant subgroups in patients with low back pain (LBP) is considered important to guide physical therapy practice and to improve outcomes. The aim of the present study was to assess the cost-effectiveness of a modified version of Delitto's classification-based treatment approach compared with usual physical therapy care in patients with sub-acute and chronic LBP with 1 year follow-up.
Methods All patients were classified using the modified version of Delitto's classification-based system and then randomly assigned to receive either classification-based treatment or usual physical therapy care. The main clinical outcomes measured were; global perceived effect, intensity of pain, functional disability and quality of life. Costs were measured from a societal perspective. Multiple imputations were used for missing data. Uncertainty surrounding cost differences and incremental cost-effectiveness ratios was estimated using bootstrapping. Cost-effectiveness planes and cost-effectiveness acceptability curves were estimated.
Results In total, 156 patients were included. The outcome analyses showed a significantly better outcome on global perceived effect favoring the classification-based approach, and no differences between the groups on pain, disability and quality-adjusted life-years. Mean total societal costs for the classification-based group were €2,287, and for the usual physical therapy care group €2,020. The difference was €266 (95% CI €-720 to €1,612) and not statistically significant. Cost-effectiveness analyses showed that the classification-based approach was not cost-effective in comparison with usual physical therapy care for any clinical outcome measure.
Conclusion The classification-based treatment approach as used in this study was not cost-effective in comparison with usual physical therapy care in a population of patients with sub-acute and chronic LBP.

Keywords Low back pain · Economic evaluation · Cost-effectiveness analysis · RCT · Classification

A. T. Apeldoorn (✉) · R. W. Ostelo · H. C. W. de Vet · M. W. van Tulder
Department of Epidemiology and Biostatistics and the EMGO+ Institute for Health and Care Research, VU University Medical Centre, Van der Boechorststraat 7, Amsterdam 1081 BT, The Netherlands
e-mail: a.apeldoorn@vumc.nl
URL: http://www.emgo.nl

R. W. Ostelo
e-mail: r.ostelo@vumc.nl

H. C. W. de Vet
e-mail: hcw.devet@vumc.nl

M. W. van Tulder
e-mail: maurits.van.tulder@falw.vu.nl

J. E. Bosmans · R. W. Ostelo · M. W. van Tulder
Department of Health Sciences and the EMGO+ Institute for Health and Care Research, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, The Netherlands
e-mail: judith.bosmans@falw.vu.nl

Introduction

The high prevalence of low back pain (LBP) and the associated costs, result in a substantial socioeconomic burden to society [1]. Therefore, it is important to evaluate the cost-effectiveness of interventions to help decision makers to determine which intervention should be reimbursed.

For care providers, it is often not possible to reliably make a specific patho-anatomical diagnosis for patients with LBP. Consequently, 85–95% of patients with LBP are diagnosed by their general practitioner as having non-specific LBP [2]. Given this absence of diagnostic precision, and the limited evidence that basing treatment on a specific aetiology improves outcomes, most clinicians use pattern recognition and patient profiling in an attempt to optimize treatment outcomes [3]. A classification system for LBP that has received considerable attention in the literature is the one developed by Delitto et al. [4]. In Delitto's classification system, classification is based on signs and symptoms of the patient and aims to identify subgroups of patients who are most likely to respond to different types of treatment, i.e. direction-specific exercises, spinal manipulation, stabilization exercises or traction. Two randomized controlled trials (RCTs) reported that treatment based on this system resulted in slightly more favourable outcomes compared with other usual physical therapy management strategies [5, 6]. One of these RCTs also conducted an economic evaluation and reported a trend towards reduced total medical costs ($774) compared with usual physical therapy care ($1,004) after 1 year [5]. However, a limitation of this study was that only direct medical costs were included, while in patients with chronic LBP most costs are attributed to indirect costs such as absenteeism from paid labour [1].

The aim of the present study was to determine whether a modified version of Delitto's classification-based treatment approach was cost-effective in patients with sub-acute (6–12 weeks) and chronic (>12 weeks) LBP. This approach was compared with usual physical therapy care according to Dutch physical therapy guidelines [7, 8]. The economic evaluation was performed from a societal perspective. Details of the effectiveness of this intervention are published elsewhere [9].

Methods

This economic evaluation was conducted alongside a RCT performed in The Netherlands between 2008 and 2010 [9]. The Medical Ethics Committee of the VU University Medical Centre in Amsterdam approved this study. The methods used are briefly described here; a detailed description of the design has been published elsewhere [10].

Patients and setting

Patients were recruited by physical therapists working in primary care in the greater Amsterdam area in the Netherlands. Recruitment occurred during patients' first consultation. Eligible patients were referred to one of four research physical therapists, who evaluated patient eligibility, obtained written informed consent, collected baseline questionnaires and conducted clinical examinations. Inclusion criteria were: LBP as the primary complaint (with or without associated leg pain), age between 18 and 65 years, current episode longer than 6 weeks, and able to read and write Dutch. The main exclusion criteria were: known- or suspected-specific LBP (e.g. cauda equina compression, fractures), severe radiculopathy (widespread sensory loss and substantial diminished myotomal strength), serious co-morbidity and psychopathology.

Examination

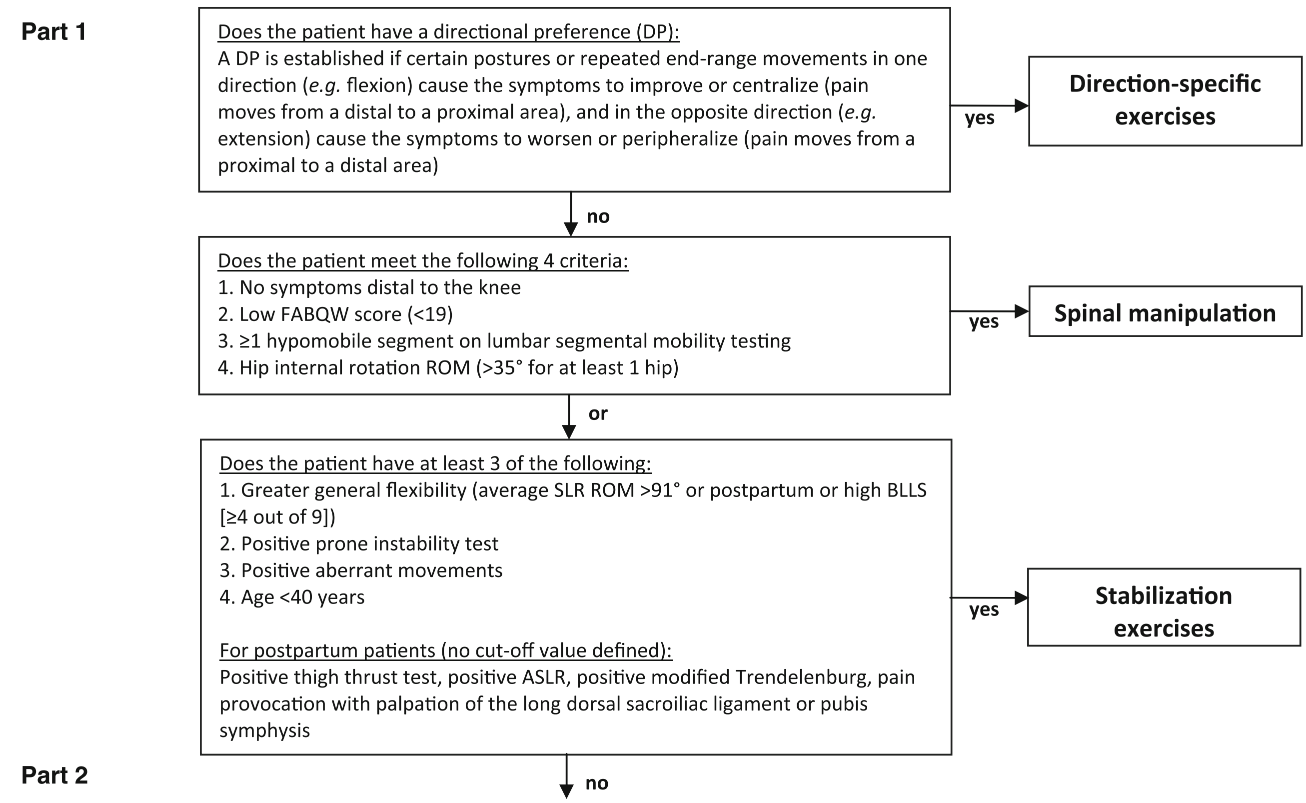
The research physical therapists classified all patients into one of the following treatment subgroups using the criteria of the classification algorithm presented in Fig. 1: direction-specific exercises, spinal manipulation, or stabilization exercises. This algorithm is an updated version of Delitto's algorithm [11] that was modified to fit into the Dutch health care system and the typical Dutch physical therapist practices. Adaptations included; exclusion of patients with acute LBP, because guidelines in the Netherlands discourage physical therapy in this phase [7, 12], and exclusion of the traction subgroup because Dutch private practices generally do not have mechanical traction equipment and traction is not recommended in Dutch guidelines [7]. Following the recommendation by Stanton et al. [13], patients that could be placed in one single treatment subgroup using only the first three boxes of the classification algorithm received the label 'clear classification'. Patients were given the label 'unclear classification' if additional criteria in the second part of the classification algorithm were needed to place them in one of the three subgroups.

Randomisation

After the examination by the research physical therapist, an independent researcher randomized each patient to either the classification-based group or the usual physical therapy care group using a randomization list. This list was generated by computer before the start of the trial and patients were pre-stratified by duration of complaints (more or less than 12 weeks) and by disability (Oswestry disability index [ODI] more or less than 25%).

Classification-based group

Patients assigned to the classification-based group were referred to a physical therapist who was trained in the



Part 2

Which subgroup does the patient fit best into?

Spinal manipulation	Stabilization exercises		Direction-specific exercises	
Factors against	**Factors favouring**	**Factors against**	**Factors favouring**	**Factor against**
• Symptoms below the knee • Increasing episode frequency • No pain with mobility testing • No hypomobility • Less discrepancy in left-to-right hip internal rotation (<10°) • Negative Gaenslen's sign • Peripheralization with motion testing	• Hypermobility with spring testing • Increasing episode frequency • 3 or more prior episodes • Previous severe low back/pelvis incident	• Discrepancy in SLR ROM (>10°) **Does the patient have at least 3 of the following:** • FABQPA score<9 • Negative prone instability test • Absence of aberrant movements • Absence of lumbar hypermobility	• Strong preference for sitting or walking or cycling • Pheripheralization in direction opposite to centralisation **Flexie DP:** • >50 years of age • Imaging evidence of lumbar spinal stenosis **Lateral DP:** • Visible lateral shift	• LBP only

LBP, Low Back Pain; RCT, Randomized Controlled Trial; FABQA and FABQW, Fear-avoidance Beliefs Questionnaire subscales Activity and Work; ROM, Range Of Motion; SLR, Straight Leg Raise; BLLS, Beighton Ligamentous Laxity Scale; ASLR, Active Straight Leg Raise

Fig. 1 The classification-based system for patients with LBP (>6 weeks) as used in this RCT

classification treatment protocol. Patients received treatment according to their classification category (direction-specific exercises, spinal manipulation, or stabilization exercises) for a minimum of 4 weeks. After this period, the physical therapist was allowed to change treatment strategy according to the current Dutch guidelines.

Direction-specific exercises

Patients in this subgroup received exercises in the direction that matched their directional preference: extension, flexion, lateral (side-gliding right or left) or flexion-rotation (right or left). For each direction, a treatment protocol was available which contained patient-generated forces (sustained positions and/or repeated movements) and physical therapist-generated forces (mobilizations in the direction of preference). The protocol could be adapted to the patient's pain responses.

Spinal manipulation

Patients in this subgroup received high-velocity thrusts directed at the affected spinal levels from T12 to L5 and/or the sacroiliac region. Decisions about the choice of the technique and the location of the forces were left to the discretion of the manual therapist. The symptomatic side was manipulated first. During each session, the manual therapist could make a maximum of four attempts (two on each side) to achieve a cavitation (i.e. 'a pop') heard or felt by the manual therapist or the patient.

Stabilization exercises

Patients in this subgroup were first taught to activate their local stability system to control neutral joint position in a supine position ('drawing in' the stomach with normal breathing). Second, the patients were instructed to perform these abdominal bracing exercises during strengthening exercises performed in standing, quadruped, and side-support positions [14]. Third, in addition to these standardized exercises, activity of the local and global muscle systems was trained in positions and movements that aggravated the patient's pain, to restore functional capacity and improve dynamic control.

Usual physical therapy care group

Patients assigned to usual physical therapy care were referred to a physical therapist who was trained in implementation of the current Dutch physical therapy LBP guidelines [7, 8]. These guidelines recommend limiting the number of treatment sessions for patients with a normal course of LBP, giving adequate information, using mainly active interventions (e.g. resistive strengthening, stretching and postural exercises) and setting functional treatment goals.

Clinical outcome measures

The outcome measures were: (1) global perceived effect (GPE), measured by self-assessment on a 7-point Likert scale ranging from 'completely recovered' to 'worse than ever'. The GPE outcome measure was dichotomized a priori into success (completely recovered and much recovered) and non-success (slightly recovered, no change, slightly worse, much worse and worse than ever), (2) pain intensity over the previous week, measured using an 11-point numerical rating scale (NRS, 0 = 'no pain' to 10 = 'worst imaginable pain'), (3) functional status, measured with the 10-item ODI, version 2.1a (score range 0–100) with higher scores indicating lower functional status [11]; Health-related quality of life was measured with the EuroQol (EQ-5D) [15]. The Dutch tariff developed by Lamers et al. [16] was used to calculate utilities. Quality-adjusted life-years (QALYs) were calculated by multiplying the utility of a health state by the time spent in this health state. Transitions between health states were linearly interpolated.

Cost measures

Cost data were collected from a societal perspective at 8, 26, 39 and 52 weeks after the start of treatment. Direct healthcare costs, direct non-healthcare costs and absenteeism from unpaid labour were measured using self-completed cost diaries [17]. Absenteeism from paid labour was measured using the Productivity and Disease Questionnaire (PRODISQ) [18]. Productivity loss costs were estimated using the friction cost method (FCM), which assumes that sick workers are replaced after a certain period of time (friction period used was 154 days) [19]. Mean productivity costs per working hour according to age and sex were used to estimate the costs of absenteeism [19]. Healthcare utilization was valued according to the guidelines published in the updated handbook for economic evaluation in the Netherlands [19]. Medication was valued using prices of the Royal Dutch society for Pharmacy, the G-standard (Z-index, The Hague, The Netherlands). Table 1 lists the cost categories and prices used in the present study. The index year for this study was 2009 [20].

Statistical analysis

The economic evaluation was performed according to the intention-to-treat principle. Multiple imputation (MI) was used to impute missing cost and effect data. In the MI procedure, five imputed data sets were generated, each of which was analysed separately. Using Rubin's rules, effects and costs from the five complete data sets were pooled [21]. All analyses were based on imputed data.

Mean clinical outcome differences were analysed with *t* tests and uncertainty was estimated by 95% confidence intervals. Costs generally have a highly skewed distribution. Therefore, bootstrapping with 5,000 replications was used to estimate 'approximate bootstrap confidence' (ABC) intervals around cost differences [22, 23]. Incremental

Table 1 Prices used in the economic evaluation

Cost category	Price (€, 2009)
Direct healthcare costs	
Primary care costs	
General practitioner, per visit	21.89
General practitioner, per telephone contact	10.94
Physical therapy, per treatment session	24.65
Manual therapy, per treatment session	34.09
Psychologist, per treatment session	83.01
Professional home care, per hour	13.76
Secondary care costs	
X-ray, per image	48.03
MRI scan, per scan	243.94
Outpatient visit specialist, per visit	60.68
Hospitalization, per day	365.18
HNP-operation, per operation	1,271.29
Outpatient rehabilitation, per day	89.94
Epidural injection, per treatment session	109.76
Facet denervation, per treatment session	293.14
Direct non-healthcare costs	
Informal care, per hour	8.89
Paid home help, per hour	8.89
Indirect non-healthcare costs	
Absenteeism paid labour, per hour[a]	21.75–51.82
Absenteeism unpaid labour, per hour	8.89

HNP herniated nucleus pulposus

[a] Depending on age and gender [22]

cost-effectiveness ratios (ICERs) were calculated by dividing the incremental costs by the incremental effects. Similarly, incremental cost-utility ratios (ICURs) were calculated by dividing the incremental costs by the difference in QALYs based on the EQ-5D. Uncertainty around the ICERs and ICURs were estimated using bootstrapping with 5,000 replications. The bootstrapped cost-effective pairs were graphically represented on cost-effectiveness planes (CE planes) [24]. Cost-effectiveness acceptability curves (CEA curves) were estimated that indicate the probability that a treatment is cost-effective at a specific ceiling ratio, which is the amount of money society is willing to pay to gain one extra unit of effect [25].

To assess the robustness of the study results, we performed three sensitivity analyses. Firstly, a per-protocol analysis was conducted. A deviation from the protocol was defined as not receiving treatment after allocation, withdrawal from therapy after three or less visits and not being treated according to patient's subgroup for patients allocated to the classification-based group. Second, the main analyses were repeated using only complete cases (i.e. with complete clinical outcome data and complete cost data). Finally, the impact of using the Human Capital Approach (HCA) instead of the FCM on the results was assessed. With the HCA the total costs of productivity losses are estimated without considering the possibility of replacing a sick worker.

Finally, a post hoc per-protocol analysis was performed among patients with a 'clear classification' label. A previous analysis showed that for this subgroup the clinical outcomes slightly increased in favour of the classification-based group after 1 year [9]. All analyses were performed in SPSS-18 and R [26].

Results

A total of 243 potentially eligible patients were recruited by 62 physical therapists. Of these, 158 patients met all inclusion criteria. Two patients dropped out immediately after randomization and before treatment started, resulting in a sample of 156 patients. The classification procedure placed 85 (54%) patients into the direction-specific exercise subgroup, 42 (27%) into the manipulation subgroup and 29 (19%) into the stabilization exercises subgroup. Baseline characteristics were comparable between the groups, and are described in Table 2. Eighty-two patients were randomized to the usual physical therapy care group and 74 to the classification-based group. Complete clinical outcome data were available for 67 (91%) patients in the classification-based group and 76 (93%) patients in the usual physical therapy care group, and complete QALY data for 60 (81%) and 59 (72%) patients, respectively. Complete cost data were available for 62 (84%) patients in the classification-based group and 65 (79%) patients in the usual physical therapy care group. Patients without complete cost data were younger, had more fearful beliefs about work measured with the Fear-Avoidance Beliefs Questionnaire [27], and were more frequently male and smokers.

Clinical outcomes

Table 3 shows that statistically significant differences were present between groups in terms of recovery status (as measured by GPE). Patients that received classification-based treatment recovered more often than patients in the usual physical therapy care group. No differences were found on the other clinical outcomes.

Resource utilization and costs

Table 4 shows the mean health care utilization after 1 year in patients with complete cost data. Table 5 shows the differences in costs between classification-based treatment

Table 2 Baseline characteristics

Characteristics	Classification-based group ($n = 74$)	Usual physical therapy care group ($n = 82$)
Age (years)	43.2 (11.7)	42.0 (10.9)
Female (%)	54.1	59.8
Dutch nationality (%)	89.2	87.8
History of LBP		
First experience of LBP ever (months)[a]	114 (48–192)	96 (24–216)
Previous episodes of LBP (%)	91.9	80.5
Lower back surgery (%)	1.4	1.2
Duration of current LBP (months)[a]	5 (2–12)	3 (2–12)
Current LBP (6–12 weeks) (%)	17.6	23.2
Current LBP (>12 weeks) (%)	82.4	76.8
Pain intensity in the past week (NRS 0–10)	6.0 (1.7)	6.2 (1.8)
Pain radiated into the leg (%)	43.2	32.9
ODI (0–100) (%)	18.1 (11.5)	21.9 (14.5)
ODI (0–100) (≥25%) (%)	27.0	31.7
Currently taking medication for LBP (%)	8.1	14.6
Smoker (%)	29.2	30.8
Marital status		
Married/living with a partner (%)	75.7	76.8
Single/divorced (%)	24.3	23.2
Education		
Low, *n* (%)	14 (18.9)	13 (15.9)
Middle, *n* (%)	23 (31.1)	38 (46.3)
High, *n* (%)	37 (50.0)	31 (37.8)
Employed (%)	83.8	85.4
Employed and currently working (%)	85.2	80.0
Employed, but currently on sick leave (%)	14.8	20.0
FABQ		
Activity (0–24)	11.6 (5.4)	12.7 (5.2)
Work (0–42)	11.5 (9.5)	15.1 (11.9)
ÖMPSQ (0–210)	80.0 (20.5)	87.2 (27.8)
SF-36		
PCS (0–100)	43.7 (8.3)	40.2 (8.7)
MCS (0–100)	52.3 (8.5)	51.1 (10.6)
EuroQol (EQ-5D)	71.5 (12.5)	69.1 (15.2)
Classification outcome		
Direction-specific exercises, *n* (%)	45 (60.8)	40 (48.8)
Manipulation, *n* (%)	22 (29.7)	20 (24.4)
Stabilization exercises, *n* (%)	7 (9.5)	22 (26.8)

Values are the mean (standard deviation) unless otherwise indicated

LBP low back pain, *NRS* numerical rating scale, *ODI* Oswestry disability index, *FABQ* fear-avoidance beliefs questionnaire, *ÖMPSQ* Örebro musculoskeletal pain screening questionnaire, *SF-36* short-form 36, *PCS* physical component summary, *MCS* mental component summary

[a] Median and inter-quartile range

and usual physical therapy care after 1 year. There was a trend towards reduced primary care costs (e.g. physical therapy and general practitioner) for the classification-based group compared to the usual physical therapy care group, but this difference was not statistically significant. Mean total societal costs (i.e. direct health care, direct non-health care and indirect non-health care) were €2,287 for the classification-based group and €2,020 for the usual physical therapy care group, but this difference was not statistically significant (mean difference €266, 95% CI €-720 to €1,612). Indirect non-health care costs (i.e. absenteeism from paid and unpaid labour) were the main contributor to the difference in total costs.

Cost-effectiveness and cost-utility analysis

Table 6 shows among others the main results of the cost-effectiveness analysis after 1 year. The bootstrapped cost-effect pairs for the ODI, the NRS and the GPE were

Table 3 Multiply imputed and pooled clinical outcomes after 1 year

Outcome	Classification-based group ($n = 74$)	Usual physical therapy care group ($n = 82$)	Difference (95% CI)
Global perceived effect (0/1)	0.68 (0.06)	0.47 (0.06)	0.20 (0.04; 0.37)
ODI improvement (0–100)	−8.2 (1.7)	−7.8 (1.7)	0.5 (−4.4; 5.4)
NRS improvement (0–10)	−2.83 (0.40)	−2.69 (0.35)	0.13 (−0.86; 1.12)
QALYs gained (0–1)	0.82 (0.02)	0.80 (0.02)	0.02 (−0.03; 0.08)

Presented are means (SEs) and differences between classification-based treatment and usual physical therapy care (95% CI)

CI confidence interval, *ODI* Oswestry disability index, *NRS* numerical rating scale, *QALYs* quality-adjusted life years

Table 4 Mean (SD) health care utilization after 1 year in patients with complete cost data

Cost category	Classification-based treatment group ($n = 62$)	Usual physical therapy care group ($n = 65$)
Direct healthcare utilization		
Primary care utilization		
General practitioner (visit and telephone contact)	0.5 (1.0)	1.1 (2.0)
Physical and manual therapy	11.2 (10.6)	16.7 (15.6)
Other paramedic disciplines	0.5 (2.7)	2.6 (9.0)
Complementary and alternative medicine	0.5 (2.7)	1.0 (4.2)
Psychologist	0 (0)	0.5 (3.7)
Professional home care (h)	0.2 (1.9)	0 (0)
Secondary care utilization		
X-ray	0.2 (0.8)	0.5 (1.2)
MRI scan	0.2 (0.4)	0.06 (0.2)
Outpatient visit	0.3 (1.0)	0.3 (0.9)
Hospitalization (days)	0.05 (0.2)	0.03 (0.2)
HNP-operation	0.05 (0.2)	0 (0)
Outpatient rehabilitation (days)	0.4 (3.3)	1.0 (7.9)
Epidural injection and facet denervation	0 (0)	0.03 (0.2)
Direct non-healthcare utilization		
Informal care (h)	1.7 (8.0)	5.0 (34.4)
Paid home help (h)	1.9 (11.0)	2.0 (16.4)
Indirect non-healthcare		
Absenteeism paid labour, days	12.6 (30.4)	14.0 (47.2)
Absenteeism unpaid labour (h)	2.9 (10.6)	1.8 (5.9)

SD standard deviation, *MRI* magnetic resonance imaging, *HNP* herniated nucleus pulposus

Table 5 Multiply imputed and pooled costs after 1 year

Cost category	Classification-based treatment group ($n = 74$)	Usual physical therapy care group ($n = 82$)	Difference (95% CI)
Direct costs	712 (151)	813 (137)	−100 (−469; 323)
Direct health care costs	648 (136)	736 (115)	−88 (−400; 290)
Primary care costs	421 (50)	538 (69)	−118 (−301; 29)
Secondary care costs	227 (105)	198 (86)	29 (−209; 327)
Direct non-health care costs	64 (23)	77 (38)	−12 (−134; 57)
Indirect costs	1,575 (378)	1,208 (289)	367 (−423; 1,545)
Total costs	2,287 (482)	2,020 (331)	266 (−720; 1,612)

Presented are means (SEs) and differences between classification-based treatment and usual physical therapy care (95% CI). 95% 'approximate bootstrap confidence' (ABC) intervals obtained by bootstrapping with 5,000 replications

SE standard error, *CI* confidence interval

Table 6 Results of cost-effectiveness and cost-utility analyses after 1 year

Outcome effect	Sample size (*n*)		Cost difference (€) (95% CI)	Effect difference (95% CI)	ICER/ICUR	Distribution in cost-effectiveness plane (%)			
	CBT	Usual PT care				North east[a]	South east[b]	South west[c]	North west[d]
Main analysis (ITT)									
ODI (0–100)	74	82	266 (−859; 1,391)	0.5 (−4.4; 5.4)	533	40	18	15	27
NRS (0–10)	74	82	266 (−859; 1,391)	0.1 (−0.9; 1.1)	2,057	41	19	14	26
GPE (0–1)	74	82	266 (−859; 1,391)	0.20 (0.04; 0.37)	1,299	67	33	0	0
QALYs (0–1)	74	82	266 (−859; 1,391)	0.03 (−0.03; 0.08)	10,543	54	31	2	13
Per protocol									
ODI (0–100)	66	76	378 (−849; 1,606)	2.3 (−2.7; 7.3)	167	60	23	5	12
NRS (0–10)	66	76	378 (−849; 1,606)	0.3 (−0.7; 1.3)	1,175	54	21	8	18
GPE (0–1)	66	76	378 (−849; 1,606)	0.27 (0.10; 0.44)	1,400	72	28	0	0
QALYs (0–1)	66	76	378 (−849; 1,606)	0.04 (−0.01; 0.09)	9,896	67	28	0	6
Complete cases									
ODI (0–100)	62	62	243 (−995; 1,480)	−0.1 (−4.6; 4.4)	−2,714	33	15	20	32
NRS (0–10)	62	65	322 (−890; 1,534)	0.1 (−0.9; 1.1)	2,732	42	18	13	27
GPE (0–1)	62	65	322 (−890; 1,534)	0.20 (0.03; 0.37)	1,612	68	31	0	1
QALYs (0–1)	59	57	431 (−881; 1,743)	0.04 (−0.02; 0.09)	12,152	64	26	1	10
HCA									
ODI (0–100)	74	82	−106 (−1,735; 1,523)	−0.1 (−4.8; 4.7)	1,851	23	27	26	24
NRS (0–10)	74	82	−106 (−1,735; 1,523)	0.1 (−0.8; 1.1)	−933	27	32	21	19
GPE (0–1)	74	82	−106 (−1,735; 1,523)	0.19 (0.02; 0.36)	−558	46	53	1	1
QALYs (0–1)	74	82	−106 (−1,735; 1,523)	0.02 (−0.03; 0.08)	−3,805	32	49	5	14
Per protocol and a 'clear classification' label									
ODI (0–100)	49	55	34 (−1,271; 1,340)	2.9 (−2.9; 8.7)	12	44	41	8	7
NRS (0–10)	49	55	34 (−1,271; 1,340)	0.3 (−0.9; 1.5)	116	35	34	15	16
GPE (0–1)	49	55	34 (−1,271; 1,340)	0.32 (0.12; 0.52)	11	51	49	0	0
QALYs (0–1)	49	55	34 (−1,271; 1,340)	0.06 (−0.01; 0.12)	596	49	49	0	2

Positive costs indicated more costs for CBT. Positive effect differences indicated a beneficial effect in favour of CBT

CBT classification-based treatment, *PT* physical therapy, *CI* confidence interval, *ICER* incremental cost-effectiveness ratio, *ICUR* incremental cost-utility ratio, *ITT* intention-to-treat, *ODI* Oswestry disability index, *NRS* numerical rating scale, *GPE* global perceived effect, *QALYs* quality adjusted life years, *HCA* human capital approach

[a] CBT more effective and more costly than usual PT care

[b] CBT more effective and less costly than usual PT care

[c] CBT less effective and less costly than usual PT care

[d] CBT less effective and more costly than usual PT care

primarily located in the northeast and southeast quadrant, indicating that the effects in the classification-based treatment were larger, but that the costs did not differ.

The ICER for GPE shows that the extra costs for one recovered patient extra in the classification-based group in comparison with the usual physical therapy care group were €1,299. Figure 2 shows the CE plane and CEA curve for GPE after 1 year. The CE plane confirms that there was significant difference in GPE between the two groups, but not in total costs. The CEA curve shows that the probability of classification-based treatment being cost-effective is 95%, if the society is willing to pay around €10,000 for one extra recovered patient.

The cost-utility analysis showed that classification-based treatment was more costly and slightly more effective than usual physical therapy care (ICUR 10,543, Table 6 'main analysis'). The CEA curve in Fig. 3 shows that the maximum probability of classification-based treatment being cost-effective lies around 0.8. To reach this probability society should be willing to pay around €30,000 per QALY.

Sensitivity analyses

The results of the sensitivity analyses are also presented in Table 6. The results of these analyses followed the same

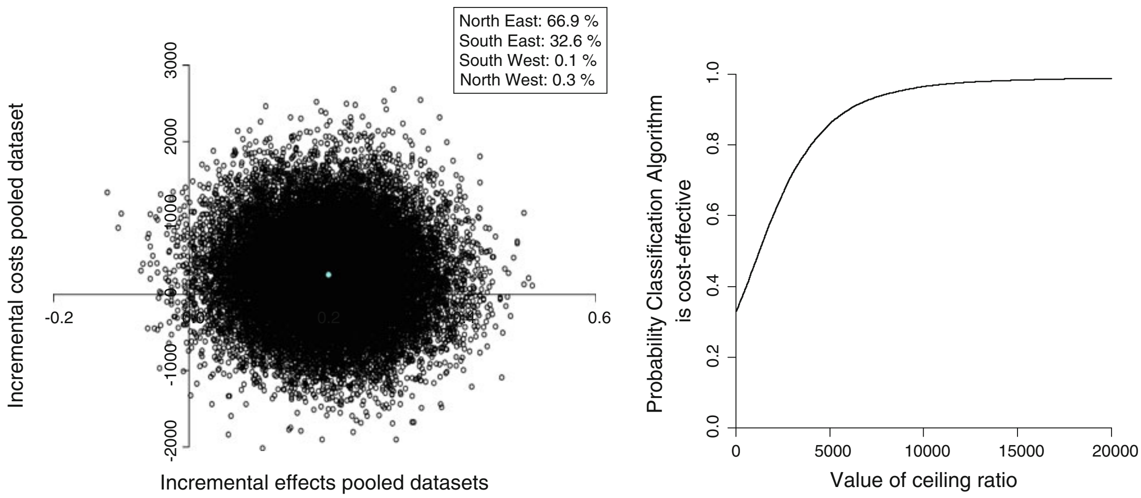


Fig. 2 Cost-effectiveness plane and cost-effectiveness acceptability curve for classification-based treatment in comparison with usual physical therapy care for global perceived effect after 12 months (multiple imputed data)

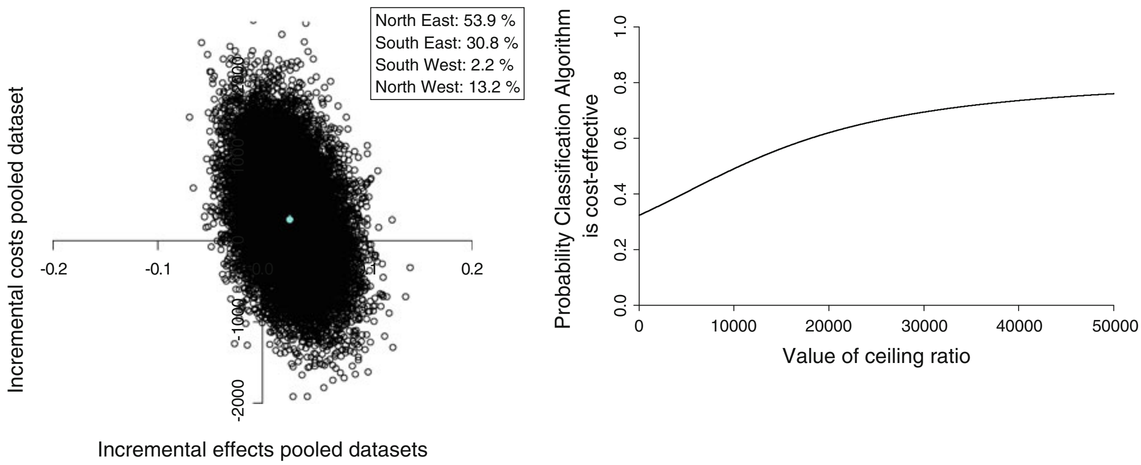


Fig. 3 Cost-utility plane and cost-utility acceptability curve for classification-based treatment in comparison with usual physical therapy care after 12 months (multiple imputed data)

pattern as the main analyses. However, when using the HCA, the total costs of classification-based treatment were slightly less compared with usual physical therapy care (Table 6).

Post hoc analysis

A total of 116 patients met the criteria of a 'clear classification' and 104 of these patients followed the treatment protocols. A per-protocol analysis among patients with a 'clear classification' label showed that there was a trend for larger effects for classification-based treatment compared with the results of the other analyses (Table 6). However, between group differences were again only statistically significant for GPE and the cost-effectiveness results did not differ fundamentally from the main analysis.

Discussion

This is the first economic evaluation that evaluated the cost-effectiveness of this classification-based treatment approach in comparison with usual physical therapy care from a societal perspective, meaning that direct and indirect costs were included. The classification-based treatment approach as used in this study was not cost-effective compared with usual physical therapy care for patients with sub-acute and chronic LBP.

Fritz et al. [5] also evaluated the cost-effectiveness of classification-based treatment, but included only direct health care costs. Fritz's study and our study found lower direct healthcare costs in the classification-based group as compared to usual physical therapy care after 1 year, but the differences were not statistically significant. However, it must be noted that the patient selection criteria, the comparison treatment and the classification algorithm used in the study of Fritz et al. [5] differed from our study and hampers comparison of the two studies.

The results of the three sensitivity analyses were in line with the main analysis, indicating that our findings were reasonably robust. The one exception was that indirect costs were higher in the classification-based approach using the FCA, but lower when using the HCA. This difference was due to two patients in the usual physical therapy care group who were absent from work during the whole year. For these two patients, only the indirect costs of the first 154 days of sick leave were included using the FCA, whereas the indirect costs were included for the whole year using the HCA resulting in higher indirect and total costs. However, the conclusion of the cost-effectiveness and cost-utility analyses did not differ between both approaches.

Patients in the usual physical therapy care group received significantly more physical and manual therapy sessions than patients in the classification-based group at a follow-up period of 1 year (mean difference 4.6, 95% CI 0.7–8.5, $p = 0.02$). This might indicate that the classification-based approach leads to lower physical therapy consumption compared to usual physical therapy care in the long run. However, another explanation could be the differences in the training of the therapists. Almost all physical therapists in the classification-based group were also trained as manual therapists, whereas in the usual physical therapy care group only a minority of physical therapists were also manual therapists. Korthals-de Bos et al. [28] also found higher utilization of physical therapy in those receiving treatment from non-manual therapists compared to those receiving treatment from manual therapists in patients with neck pain. In the present study, the differences between the two groups in utilization of physical and manual therapy did not have major consequences for the total costs, because the physical and manual therapy costs represented only around 15% of the total costs.

An exploratory post hoc per-protocol analysis showed that the cost-effectiveness of classification-based treatment in patients with a 'clear classification' label was slightly better compared with the results of the other analyses. Ideally, a classification algorithm should be able to classify all patients into (only) one subgroup. In our RCT, using the first part of the algorithm 24% of all patients did not meet any of the subgroups and 16% met more than one subgroup. Our findings are consistent with those of Stanton et al. [13]. Although our post hoc results must be interpreted very cautiously, they may indicate that the classification algorithm needs further refinement and possibly extension to be able to classify other relevant subgroups in order to improve clinical outcomes and reduce costs.

Limitations

In this trial 16–21% of the cost data were missing and for some of the baseline characteristics the patients with complete cost data differed from the patients with incomplete cost data. However, the results of the cost-effectiveness and cost-utility analyses for patients with complete cost data did not differ from the results after applying multiple imputation techniques for missing cost and effect data. Therefore, this is unlikely to have any major influence on the results.

Conclusions

Societal costs of classification-based treatment in patients with LBP were comparable with usual physical therapy care. Cost-effectiveness could not be demonstrated with regard to any of the outcome measures. Therefore, based on a societal cost perspective, we do not recommend widespread implementation of this approach.

Acknowledgments For this study we received a grant from the Netherlands Organisation for Health Research and Development (ZonMw, project no: 170882401). The funders had no role in the design and conduct of the study, data analysis or writing of the report. The authors wish to thank all the physical therapists and patients for their participation in this project. The authors acknowledge Julie Fritz, Mark Hancock, Steve Kamper, Nick Henschke and Tasha Stanton who contributed to the discussions and conclusions presented in this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Lambeek LC, van Tulder MW, Swinkels IC et al (2011) The trend in total cost of back pain in the Netherlands in the period 2002 to 2007. Spine 36:1050–1058
2. Waddell G (2005) Subgroups within "nonspecific" low back pain. J Rheumatol 32:395–396
3. Kent P, Keating J (2004) Do primary-care clinicians think that nonspecific low back pain is one condition? Spine 29:1022–1031
4. Delitto A, Erhard RE, Bowling RW (1995) A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Phys Ther 75:470–485
5. Fritz JM, Delitto A, Erhard RE (2003) Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine 28:1363–1371
6. Brennan GP, Fritz JM, Hunter SJ et al (2006) Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. Spine 31:623–631
7. Bekkering GE, Hendriks HJM, Koes BW et al (2005) KNGF-richtlijn fysiotherapie bij lage-rupijn [National practice guidelines

for physical therapy in patients with low back pain]. Ned Tijdschr Fysiother 115:1–32
8. Heijmans WFGL, Hendriks HJM, van der Esch M et al (2003) KNGF-richtlijn manuele therapie bij lage-rugpijn [National practice guidelines for manual therapy in patients with low back pain]. Ned Tijdschr Fysiother 113:1–12
9. Apeldoorn AT, Ostelo RW, van Helvoirt H et al (2011) A randomized controlled trial on the effectiveness of a classification-based system for sub-acute and chronic low back pain (Submitted for publication)
10. Apeldoorn AT, Ostelo RW, van Helvoirt H et al (2010) The cost-effectiveness of a treatment-based classification system for low back pain: design of a randomised controlled trial and economic evaluation. BMC Musculoskelet Disord 11:58
11. Fritz JM, Cleland JA, Childs JD (2007) Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther 37:290–302
12. Dutch Institute for Health Care Improvement (CBO) (2002) Richtlijn aspecifieke lage rugklachten [Guideline aspecific low back complaints]. Van Zuiden Communications, Alphen aan de Rijn
13. Stanton TR, Fritz JM, Hancock MJ et al (2011) Evaluation of a treatment-based classification algorithm for low back pain: a cross-sectional study. Phys Ther 91:496–509
14. Hicks GE, Fritz JM, Delitto A et al (2005) Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. Arch Phys Med Rehabil 86:1753–1762
15. EuroQol Group (1990) EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 16:199–208
16. Lamers LM, Stalmeier PF, McDonnell J et al (2005) Kwaliteit van leven meten in economische evaluaties: het Nederlands EQ-5D-tarief [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. Ned Tijdschr Geneeskd 149: 1574–1578
17. Goossens ME, Rutten-van Molken MP, Vlaeyen JW et al (2000) The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. J Clin Epidemiol 53:688–695
18. Koopmanschap MA (2005) PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. Expert Rev Pharmacoecon Outcomes Res 5:23–28
19. Oostenbrink JB, Boumans CAM, Koopmanschap MA et al (2004) Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg, geactualiseerde versie [Manual for costing: methods and standard costs for economic evaluations in health care, updated version]. Health Care Insurance Board, Diemen
20. Central Bureau of Statistics the Netherlands (CBS)/Dutch Institute for Health Care Improvement (CBO) (2007) Voorburg/Heerlen
21. Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley, New York
22. Burton A, Billingham LJ, Bryan S (2007) Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. Clin Trials 4:154–161
23. Efron B (1994) Missing data, imputation and the bootstrap. J Am Stat Assoc 89:463–475
24. Black WC (1990) The CE plane: a graphic representation of cost-effectiveness. Med Decis Making 10:212–214
25. Fenwick E, O'Brien BJ, Briggs A (2004) Cost-effectiveness acceptability curves—facts, fallacies and frequently asked questions. Health Econ 13:405–415
26. R Development Core Team (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
27. Waddell G, Newton M, Henderson I et al (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
28. Korthals-de Bos IB, Hoving JL, van Tulder MW et al (2003) Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. BMJ 326:911