7

Effectiveness and cost-effectiveness of radiofrequency denervation for chronic low back pain originating from a combination of the facet joints, sacroiliac joints and/or discs

Esther T Maas, Johan NS Juch, Raymond WJG Ostelo, J George Groeneweg, Jan-Willem Kallewaard, Bart Koes, Arianne P Verhagen, Johanna M van Dongen, Frank JPM Huygen, Maurits W van Tulder

Under review

ABSTRACT

Introduction Radiofrequency (RF) denervation is one of the treatment options for patients with chronic mechanical low back pain (LBP), but evidence for its effectiveness and cost-effectiveness is limited. The aim of this study is to establish whether RF denervation added to a standardised exercise programme is effective and cost-effective compared to a standardised exercise programme alone for patients with chronic low back pain originating from a combination of the facet joint, sacroiliac joints (SI-joint) and/or disc who are referred to a pain clinic.

Methods A multicentre, non-blinded randomised controlled trial and economic evaluation from a societal perspective was performed involving patients with chronic LBP originating from the facet joint, SI-joint and/or disc, irresponsive to conservative care. Participants were randomly assigned to the intervention- or control group based on medical history and a physical examination. Both groups received a three month standardised exercise programme (8-12 hours) and psychological support if needed. Depending on the suspected diagnosis of facet joint, SI-joint and/or disc pain, the intervention group also received a diagnostic block and/or provocative discography and RF denervation if the diagnosis was confirmed. Primary outcomes were pain intensity (0-10 numeric rating scale), global perceived effect (1-7 Likert scale; 1-2 was defined as success), and functional status (0-100 Oswestry Disability Index) three months after the intervention. Participants completed questionnaires at baseline, three and six weeks, three, six, nine and 12 months after start of the treatment. Costs were collected using self-reported cost questionnaires. Longitudinal mixed-model analyses for the effectiveness and seemingly unrelated regression analyses for the cost-effectiveness were performed by intention to treat.

Results Between January 1, 2013 and October 24, 2014, 202 participants were assigned. Statistically significant between-group differences were found three months after the start of the intervention for pain intensity (-0.99; 95%CI:-1.73 to -0.25) in favour of the intervention group, and for functional status (6.44; 95%CI: 0.61 to 12.26) at 12 months after start of the intervention) favouring the control group. After 12 months, societal costs did not statistically significantly differ between groups (€1124.24; 95%CI -911.17 to 4001.44) and the probability of cost-effectiveness was low (i.e. <0.66).

Discussion This study showed statistically significant results for RF denervation added to a standardised exercise programme compared to a standardised exercise programme

alone three months after the start of the intervention for pain when used for patients with chronic low back pain originating from the facet joint, SI-joint and/or disc. At the 12 months follow up there was statistically significant difference in physical functional status in favour of the control group. Furthermore there were no statistically significant differences on any other outcome or any other follow-up moment. The cost-utility analysis indicated that there were no statistically significant differences in cost-effectiveness between a standardised exercise programme with RF denervation versus a standardised exercise programme alone.

INTRODUCTION

Low back pain (LBP) has major social and economic consequences worldwide, and 10-15% of LBP patients have been found to develop chronic (>three months) symptoms.[1-4] Effective LBP treatments are desirable due to the aging population and the associated increasing disability and costs.

The facet joint, sacroiliac joint (SI-joint), and disc are potential sources of chronic LBP (CLBP) from mechanical origin. Each structure can cause pain that is evoked by noxious stimulation of structures in the lumbar spine. Noxious stimulation causes dull, aching pain in the back.[5] The prevalence of facet joint, SI-joint and disc pain, among CLBP patients ranges from 13 to 40%.[6-8] This wide range in prevalence might be due to a lack of a universally accepted gold standard for diagnosing CLBP of any mechanical structure. Furthermore, LBP is not always associated with an isolated structure, and multiple structures may contribute to the pain.

Minimal interventional treatments, such as radiofrequency (RF) denervation, are commonly used in pain clinics in a multidisciplinary pain programme for the treatment of LBP symptoms originating from the disc, facet joint and/or SI-joint. RF denervation involves the use of an electrical current to generate a controlled lesion to safely interrupt nociceptive input and reduce spinal pain.[9] Recent systematic reviews have shown low to moderate quality evidence for small positive effects of facet joint RF denervation on pain and functional status compared to placebo or steroid injections; and very low to moderate quality evidence with conflicting results for SI-joint RF denervation and transdiscal RF denervation.[10-12] Most of these studies had small sample sizes and methodological flaws. In clinical practice, RF denervation is provided to the main suspected cause of the pain (facet joints, SI-joint or discs), and may be repeated in another suspected structure if the pain has not been resolved. Furthermore, current anaesthesiology guidelines describe that providing RF denervation in addition to a standardised exercise programme may optimize the effect. There is an absence of high quality evidence on the effectiveness and cost-effectiveness of RF denervation used for a combination of structures contributing to the LBP and in addition to usual care.[13] Therefore, the aim of this study is to establish whether RF denervation added to a standardised exercise programme is effective and cost-effective compared to a standardised exercise programme alone for patients with CLBP originating from a combination of facet joint, SI-joint and/or disc who are referred to a pain clinic.

METHODS

Study design and participants

This study was part of a larger collective initiative: the MinT (Minimal Interventional Treatment) study.[14] The MinT study consisted of three non-blinded multicentre randomised controlled trials (RCTs) and economic evaluations (1:1 ratio); as well as an observational study. The studies were conducted in 16 pain clinics and 102 physiotherapy practices in the Netherlands. Four RCTs were defined in the study design paper, including patients with facet joint, SI-joint, disc, or a combination of pain problems.[14] No patients with isolated disc problems were included after a recruitment period of five months; therefore this RCT was prematurely terminated. The study design and informed consent procedure were approved by the Medical Ethics Committee of the Erasmus University Medical Centre Rotterdam, The Netherlands (registration number MEC-2012-079). Full details on the study design have been published in the study protocol.[14]

Participants were aged between 18 and 70 years, had CLBP, had no improvement of symptoms after conservative treatment, were referred to a pain clinic, and suspected of LBP originating from a combination of the facet joint, SI-joint and/or disc pain based on patient history and physical examination performed by a standard format. The exclusion criteria were: previous surgery to the painful discs, pregnancy, inability to complete questionnaires, previous anticoagulant drug therapy and/or coagulopathy, body mass index (BMI)> 35, involvement in a work-related conflict, or severe psychiatric or psychological problems.

Randomisation and blinding

A total of 85 participants per group was needed (using a power of .9, alpha .05 and a correlation of .5 for repeated measurements)[15] to detect a clinically relevant mean difference of two points on the NRS for pain intensity (SD 4). Anticipating a potential study withdrawal of 20%, 102 participants per group were required.

A centrally developed computer random number generator was used for randomisation. Randomisation was stratified by pain clinic. Participants who met the inclusion criteria and provided informed consent were allocated by a local research nurse to the intervention or control group (1:1 ratio). Participants and care providers were not blinded as a result of the study design. As all outcome measures were self-reported by the participant, the outcome measurement was not blinded either. Conclusions based on data analysis

made by the project team were blinded for treatment allocation. All participants were assigned a unique number to ensure that data was stored and analysed anonymously.

Interventions

Each participant received a standardised exercise programme and if necessary, psychological support. Intervention group participants also received, based on the suspected diagnosis, a diagnostic facet- or SI-joint block or provocative discography. RF denervation was administered where the diagnostic block or provocative discography was positive.

Exercise programme: All participants received a standardised exercise programme that was in accordance to the LBP guideline of the Royal Dutch Society for Physical Therapy.[16] The programme ranged from eight to 12 hours (one to two sessions per week) over three months. More details are available in the study protocol.[14]

Psychological support: If necessary, the participant was referred to a psychologist. The participants received psychological support at the discretion of the psychologist.

Diagnostic block and radiofrequency denervation: A diagnostic facet joint block, diagnostic SI-joint block or provocative discography were used as definitive diagnostic procedures. These procedures were performed (after patient history and physical examination) according to International Spine Intervention Society (ISIS) criteria.[17,18] An extensive description of these procedures is added in the Appendix S1.

Participants who responded negatively to the diagnostic facet joint block, SI-joint block or provocative discography could receive a second diagnostic block or were referred to the standardised exercise programme. Participants, who responded positively to the diagnostic block or discography, were referred to the standardised exercise programme and received RF denervation (in the intervention group). Participants with CLBP originating from the facet joint received facet joint RF denervation, participants with CLBP originating from the SI-joint received SI-joint RF denervation (Cooled RF technique, Palisade technique or Simplicity III probe) and for participants with pain originating from the discs, Intradiscal Electrothermal Therapy (IDET) or Biacuplasty were provided. An extensive description of these procedures is added in the Appendix S1 as well.

In both treatment groups, no co-interventions were allowed during the three-month

intervention period. Co-interventions between three months and one year follow-up were monitored and evaluated.

Outcomes

All outcomes were assessed at baseline, three, six, nine, and 12 months by web-based questionnaires. Pain intensity, global perceived effect (GPE), and health-related quality of life were assessed at three and six weeks after start of treatment as well.

Primary outcomes consisted of the core outcome set for clinical trials in LBP;[19] including pain intensity (11-point Numerical Rating Scale (NRS)),[20] global perceived effect (GPE) (7-point Likert scale),[21] and functional status (Oswestry Disability Index (ODI)).[22] Pain reduction of more than 30% or two points decrease in NRS was defined as treatment success for a post-hoc responder analysis. Success of treatment in GPE was achieved if a patient scored "complete recovery" or "much recovery".

Secondary outcomes were health-related quality of life (EuroQol (EQ-5D)),[23] patient satisfaction (7-point Likert scale),[24] general health (Rand-36),[25] and chronic pain experiences (Multidimensional Pain Inventory (MPI)).[26] The participants' EQ-5D health states were converted into utilities using the Dutch tariff.[27] Quality Adjusted Life Years (QALYs) were calculated using linear interpolation between measurement points.

Cost data were collected from a societal perspective using three-monthly self-reported cost questionnaires and included the intervention, other healthcare, informal care, work absenteeism, and unpaid productivity costs.[28] Intervention costs comprised the costs for diagnostic procedures (diagnostic facet joint blocks, SI-joint block or provocative discography) and treatment costs (facet joint RF denervation, Cooled RF technique, Palisade technique or Simplicity III probe, IDET and/or Biacuplasty). Bottom-up micro-costing using two hospital accounting records was used to estimate the intervention costs (i.e. data were collected regarding the quantity of diagnostic and treatment procedures used per participant). Healthcare utilization included the use of primary healthcare, secondary healthcare, and the use of prescribed and over-the-counter medications. Dutch standard costs were used to value healthcare utilization.[29] If these were not available, prices according to professional organizations were used. Medication use was valued using unit prices of the Royal Dutch Society of Pharmacy.[30] Informal care was valued using a recommended Dutch shadow price of €13.7/hour and included care by family, friends, and other volunteers.[29] Work absenteeism was measured with the Productivity and Disease Questionnaire (PRODISQ).[31] Work absenteeism was valued in

accordance with the friction cost approach (friction period=23 weeks) and using age- and gender-specific price weights.[29] Unpaid productivity costs were valued using a recommended Dutch shadow price of €13.7/hour and included all hours of volunteer work, domestic and educational activities that patients were not able to perform due to their CLBP.[29] All costs were converted to 2014 Euros using consumer price indices.[32]

Statistical analyses

Effectiveness: Analyses were performed according to the intention-to-treat principle. Baseline characteristics of intervention and control group participants as well as those of participants with complete and incomplete data were compared using descriptive statistics to identify possible selective drop-out. We used maximum likelihood estimation for longitudinal mixed-effects models under "missing at random" assumptions. This included a term for pain clinic if necessary to adjust for the possible effect of missing data on the study results in the analysis of mean changes for continuous outcomes.[33] We dichotomised pain scores in a post-hoc responder analysis, for which we used a generalized linear mixed model (logit link) with the same multilevel structure. Regression coefficients or odds ratios with 95% confidence intervals (CI) were calculated. We adjusted for baseline characteristics and were interested in the time by treatment interaction. A sensitivity analysis estimated an "as-treated" longitudinal analysis, excluding the participants in the control group who received RF denervation, or excluding patients in the intervention groups who did not receive RF denervation. Effectiveness data were analysed in MLwiN (V2.22), with a level of significance of $p<.05$.

Cost-effectiveness: Cost-effectiveness analyses (CEAs) were conducted from a societal perspective. Missing data were imputed using multiple imputation.[34,35] The imputation model included gender, smoking, marital status, age, BMI, complaint history, education, treatment expectations, and baseline and follow-up cost and effect measure values. Using Fully Conditional Specification and Predictive Mean Matching, 10 complete data sets were created (loss-of-efficiency<5%). Pooled estimates were calculated according to Rubin's rules.[36] Unadjusted and adjusted between-group cost differences were calculated for total and disaggregated costs. Seemingly unrelated regression (SUR) analyses were performed in which effect differences were corrected for their baseline values and cost differences for baseline sickness absence while taking into account the correlation between costs and effects.[37] Uncertainty surrounding the cost differences

and Incremental cost-effectiveness ratio (ICER) were estimated using Bias Corrected and Accelerated (BCA) bootstrapping techniques (5000 replications).[38] The ICER was calculated by dividing the adjusted cost difference by that in effect (QALYs). To graphically illustrate the uncertainty around the ICER, bootstrapped incremental cost-effectiveness planes were constructed. A cost-effectiveness acceptability curve (CEAC) was estimated to indicate the intervention's probability of cost-effectiveness at different values of willingness-to-pay.[39] To assess the robustness of the results, three sensitivity analyses were performed. We compared the SF6D and EQ5D (SA1), FCA to the human capital approach (HCA) (SA2), and performed a complete-case analysis (SA3).
The economic evaluation was performed using STATA (V12, Stata Corp, College Station, TX).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, interpretation of data, or writing of the paper. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Study participants

In total 7103 patients were assessed for eligibility of whom 3486 were included in the MinT study (one of the three RCTs or observational study) between January 1, 2013 and October 24, 2014 (the inclusion period for patients in the RCTs assessing complaints of the facet joints, SI-joint and/or discs). Of these 3486 participants, 793 patients were diagnosed with symptoms originating from more than one structure (facet joints, SI-joint and/or discs), 202 participants met the inclusion criteria for this RCT and were randomly assigned to the intervention (N=102) and control group (N=99) (Figure 1).
Baseline characteristics, distributions of the origin of pain, and outcomes at baseline were in balance between both groups (Table 1). In the intervention group, 35% did not receive RF denervation, 45% received one RF treatment, and 20% more than one treatment (with a maximum of four treatments). Complete data on the primary outcomes after three months was obtained from 169 participants (83%).

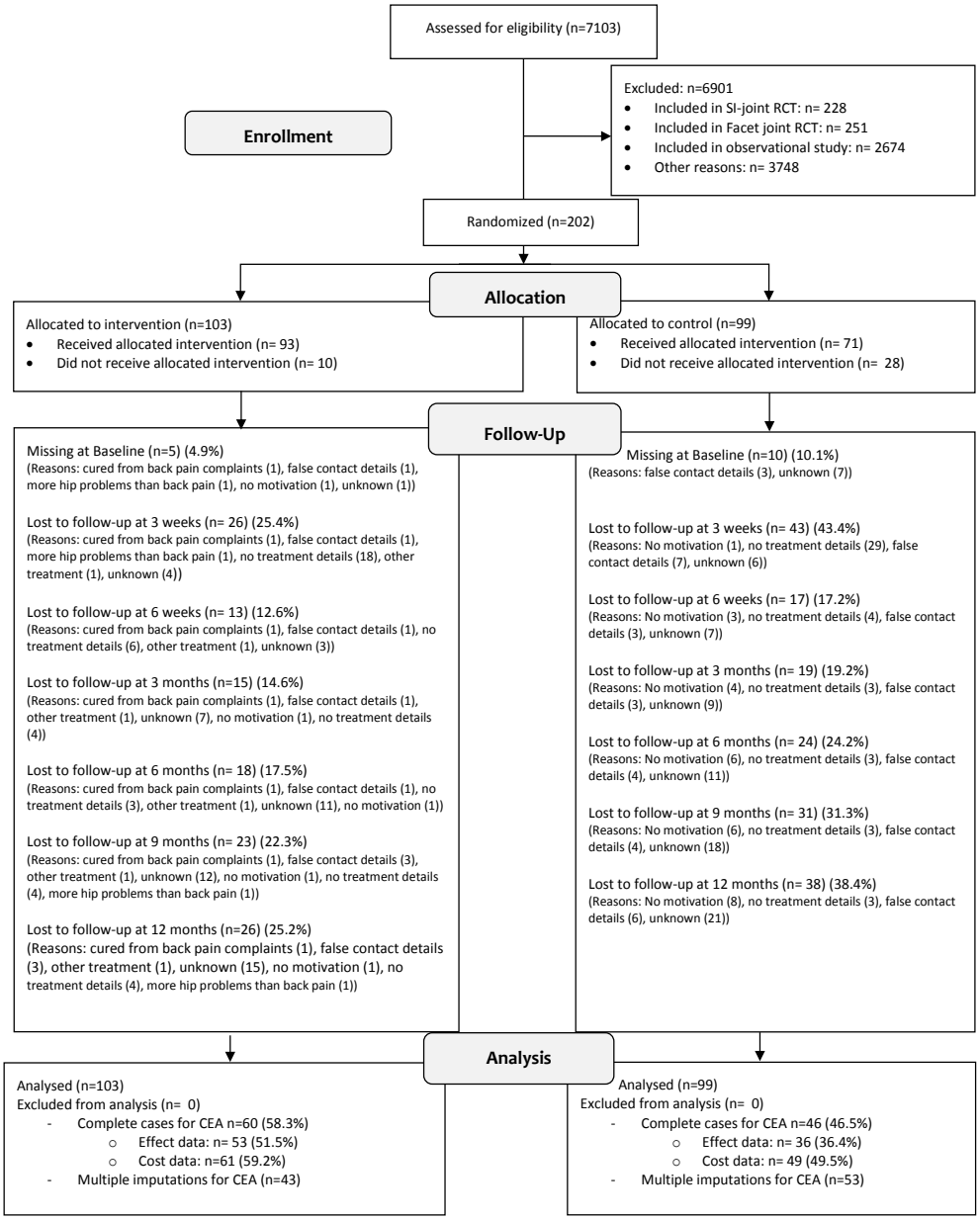

Figure 1. Flow diagram

Table 1. Baseline characteristics

Characteristics	Intervention Randomised: N=103 Complete baseline: N=98	Control Randomised: N=99 Complete baseline: N=89
Age in years (SD)	50.80 (11.33)	53.31 (10.35)
Female (N (%))	64 (65.3%)	66 (74.2%)
BMI (SD)	26.84 (3.82)	26.43 (4.25)
Smoker (N (%))	23 (23.5%)	26 (29.2%)
Education		
• Low (N (%))	52 (53.6%)	43 (48.3%)
• Moderate (N (%))	33 (34.0%)	32 (36.0%)
• High (N (%))	12 (12.4%)	14 (15.7%)
History of back pain complaints		
• Time since first experience with low back pain in months (median (IQR))	120.58 (37.32 – 222.04)	97.33 (32.33 – 192.58)
• Time since current episode with low back pain in months (median (IQR))	36.50 (12.17 – 121.67)	32.33 (8.00 – 97.19)
Married/living with a partner (N (%))	66 (67.3%)	68 (76.4%)
Expectations		
• Credibility (0-27)	20.10 (4.70)	17.07 (5.99)
• Expectancy (0-27)	16.88 (5.78)	14.38 (6.24)
Having a paid job	48 (55.8%)	44 (55.7%)
Origin of back pain (N)		
Facet & SI-joint	69	70
Facet & disc	18	18
SI-joint & disc	6	1
Facet & SI-joint & disc	3	6
Unknown	7	4
Outcomes		
Pain intensity in the past week (NRS 0-10)	7.19 (1.43)	7.43 (1.41)
Oswestry disability index (mean (SD))	39.06 (14.03)	37.20 (13.74)
Quality of life (EQ-5D)	0.49 (0.28)	0.52 (0.28)

Abbreviations: SD, Standard Deviation; N, number; BMI, Body Mass Index; EQ-5D, EuroQol-5D

Complete data on all follow-up assessments during the year, was obtained from 89 participants (44.1%) (intervention n=53; control n=36) on the effect measures and from 110 participants (54.5%) (intervention n=61; control n=49) on the cost measures.
More missing data on the effect measures was caused by missing data at the three and six weeks assessments. Participants with complete data had LBP complaints for a longer period of time, but were similar for all other demographic characteristics (Table S1 in the Supplementary Appendix).
During the three-month intervention period, two participants in the control group received the RF denervation (one SI-joint RF denervation, one facet-joint RF denervation), and two participants did not receive any treatment. In the intervention group, ten participants did not receive RF denervation. These 14 participants were considered protocol violators. Between three- and 12-month follow–up, 31 control group participants (31.3%) received RF denervation (19 participants received facet joint RF denervation, 12 participants received SI-joint RF denervation, 10 participants received two treatments, three received three extra treatments and one participant four during the year follow-up), and eight intervention group participants received another RF denervation (three participants received facet joint RF denervation, five participants received SI-joint RF denervation).

Intention-to-treat analyses

Both groups decreased in pain over the 12-month follow-up. Table 2 shows estimated overall mean changes from baseline and treatment effects between treatment groups of the primary outcomes for all follow-up moments. Multilevel analyses adjusted for baseline variables showed statistically significant differences between the two groups for pain intensity three months after the intervention (-0.99; 95%CI:-1.73 to -0.25), but not for any other follow-up assessments. No differences in functional status were found, except for the 12 months assessment in which participants in the control group showed a statistically significant better functioning than participants in the intervention group (6.44; 95%CI: 0.61 to 12.26). No differences were found in GPE.
For the secondary outcomes, favourable results for the intervention group were found for QALYs at three weeks (0.14; 95%CI 0.07 to 0.22) and three months after the intervention (0.09; 95%CI: 0.01 to 0.16) and for patient satisfaction after three months (-0.52; 95%CI: -0.97 to -0.07) (Table 4).

Table 2. Treatment effects for primary outcomes based on intention-to-treat analyses

		Mean Intervention group	Mean Control group	Treatment effect (95%CI)	P value for difference	
Primary outcomes						
NRS Pain (SD)*	Overall effect			-0.21 (-0.76 to 0.35)	0.47	
	Baseline	7.19 (1.43)	7.43 (1.41)			
	3 weeks	5.45 (2.19)	6.40 (1.80)	-0.65 (-1.47 to 0.17)	0.12	
	6 weeks	5.37 (2.93)	6.09 (1.99)	-0.40 (-1.14 to 0.34)	0.29	
	3 months	4.77 (2.48)	5.94 (2.31)	-0.99 (-1.73 to -0.25)	0.01	
	6 months	4.92 (2.42)	4.95 (2.58)	0.33 (-0.53 to 1.09)	0.39	
	9 months	5.01 (2.47)	5.25 (2.48)	-0.05 (-0.82 to 0.73)	0.90	
	12 months	4.85 (2.65)	4.38 (2.54)	0.69 (-0.10 to 1.49)	0.09	
ODI Functioning (SD)*	Overall effect			1.90 (-2.96 to 6.76)	0.44	
	Baseline	39.06 (14.03)	37.20 (13.74)			
	3 months	28.00 (15.61)	33.63 (16.83)	-4.66 (-10.21 to 0.89)	0.10	
	6 months	30.24 (18.78)	28.61 (16.59)	4.44 (-1.18 to 10.06)	0.12	
	9 months	30.73 (17.76)	28.70 (17.15)	3.55 (-2.17 to 9.26)	0.22	
	12 months	31.20 (17.40)	24.67 (14.65)	6.44 (0.61 to 12.26)	0.03	
		Success % Intervention group	**Success % Control group**	**OR (95%CI)**	**P value for difference**	**NNT**
GPE Success	Overall effect			1.35 (0.77 to 2.37)	0.29	
	3 weeks	17 (22.1%)	4 (7.1%)	2.46 (0.72 to 8.46)	0.15	6
	6 weeks	25 (27.8%)	7 (8.5%)	2.77 (0.99 to 7.70)	0.05	5
	3 months	30 (34.1%)	13(16.3%)	2.46 (0.97 to 6.22)	0.06	5
	6 months	30 (35.3%)	28 (37.3%)	0.67 (0.29 to 1.58)	0.36	-50
	9 months	29 (35.4%)	21 (30.9%)	1.17 (0.48 to 2.86)	0.73	22
	12 months	26 (34.7%)	22 (36.1%)	0.87 (0.35 to 2.16)	0.76	-71

Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.Abbreviation: SD. Standard Deviation; NRS, Numeric Rating Scale; GPE, Global Perceived Effect; ODI, Oswestry Disability Index; NNT, Numbers Needed to Treat
** Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100; for GPE, 1&2 is success*

Table 3. Treatment effects for primary outcomes based on intention-to-treat analyses, in terms of successful treatment

		Success % Intervention group	Success % Control group	OR (95%CI)	P value for difference	NNT
Primary outcomes						
NRS Pain success >30% reduction	Overall effect			1.21 (0.75 to 1.96)	0.43	
	3 weeks	30.7	14.6	3.14 (1.09 to 9.04)	0.03	6
	6 weeks	36.4	29.2	1.24 (0.55 to 2.80)	0.60	13
	3 months	50.0	26.4	2.86 (1.27 to 6.46)	0.01	4
	6 months	43.9	55.9	0.59 (0.26 to 1.30)	0.19	-8
	9 months	46.9	42.6	1.09 (0.48 to 2.48)	0.83	23
	12 months	49.3	57.1	0.73 (0.32 to 1.70)	0.47	-12
NRS Pain success >2 points reduction	Overall effect			1.03 (0.64 to 1.66)	0.90	
	3 weeks	42.7	25.0	2.16 (0.86 to 5.39)	0.10	5
	6 weeks	50.0	45.8	0.92 (0.43 to 1.98)	0.83	23
	3 months	55.8	38.9	1.66 (0.77 to 3.60)	0.20	5
	6 months	59.8	63.2	0.78 (0.35 to 1.74)	0.54	-29
	9 months	59.3	55.7	0.95 (0.42 to 2.15)	0.91	27
	12 months	54.7	66.1	0.58 (0.25 to 1.36)	0.21	-8

Abbreviation: NRS, Numeric Rating Scale; NNT, Numbers Needed to Treat

Table 4. Treatment effects for secondary outcomes based on intention-to-treat analyses, continuous

		Mean Intervention group (SD)	Mean Control group (SD)	Treatment effect (95%CI)	P value for difference
Secondary outcomes					
EQ5D Utilities**	Overall effect			0.04 (-0.01 to 0.09)	0.12
	Baseline	0.48 (0.28)	0.52 (0.28)		
	3 weeks	0.64 (0.27)	0.60 (0.25)	0.06 (-0.02 to 0.14)	0.15
	6 weeks	0.70 (0.20)	0.57 (0.29)	0.14 (0.07 to 0.22)	0.00
	3 months	0.69 (0.22)	0.63 (0.27)	0.09 (0.01 to 0.16)	0.02
	6 months	0.69 (0.25)	0.69 (0.25)	0.01 (-0.06 to 0.09)	0.74
	9 months	0.65 (0.28)	0.70 (0.25)	-0.02 (-0.06 to 0.09)	0.62
	12months	0.64 (0.26)	0.74 (0.21)	-0.07 (-0.15 to 0.01)	0.08
Patient satisfaction*	Overall effect			-0.17 (-0.56 to 0.22)	0.39
	Baseline	-	-		
	3 months	2.98 (1.15)	3.48 (1.39)	-0.52 (-0.97 to -0.07)	0.02
	6 months	3.05 (1.45)	3.13 (1.29)	-0.056 (-0.51 to 0.40)	0.81

		Mean Intervention group (SD)	Mean Control group (SD)	Treatment effect (95%CI)	P value for difference
	9 months	3.19 (1.43)	3.29 (1.41)	-0.13 (-0.59 to 0.34)	0.60
	12 months	3.32 (1.55)	3.08 (1.36)	0.10 (-0.37 to 0.58)	0.67
MPI Pain severity*	Overall effect			0.02 (-0.34 to 0.38)	0.91
	Baseline	4.00 (0.90)	3.96 (1.04)		
	3 months	2.99 (1.36)	3.42 (1.43)	-0.50 (-0.94 to -0.05)	0.02
	6 months	3.06 (1.53)	2.90 (1.42)	0.26 (-0.18 to 0.71)	0.24
	9 months	3.09 (1.56)	3.05 (1.50)	0.10 (-0.35 to 0.55)	0.67
	12 months	3.07 (1.47)	2.61 (1.50)	0.68 (0.22 to 1.15)	0.00
MPI interference*	Overall effect			0.09 (-0.24 to 0.43)	0.58
	Baseline	3.35 (1.22)	3.25 (1.27)		
	3 months	2.84 (1.30)	2.92 (1.49)	-0.18 (-0.58 to 0.21)	0.36
	6 months	2.80 (1.57)	2.57 (1.33)	0.30 (-0.10 to 0.70)	0.14
	9 months	2.78 (1.41)	2.77 (1.51)	0.01 (-0.40 to 0.43)	0.95
	12 months	2.87 (1.44)	2.45 (1.45)	0.31 (-0.11 to 0.73)	0.15
MPI Life control*	Overall effect			0.09 (-0.17 to 0.34)	0.49
	Baseline	3.98 (1.13)	4.09 (0.85)		
	3 months	4.15 (1.17)	3.97 (0.99)	0.28 (-0.05 to 0.61)	0.09
	6 months	4.16 (1.13)	4.10 (1.05)	0.10 (-0.23 to 0.44)	0.54
	9 months	4.06 (1.34)	4.09 (1.15)	-0.02 (-0.36 to 0.32)	0.91
	12 months	4.07 (1.22)	4.25 (1.12)	-0.05 (-0.40 to 0.30)	0.77
MPI Affective distress*	Overall effect			0.03 (-0.15 to 0.21)	0.74
	Baseline	2.66 (0.84)	2.61 (0.72)		
	3 months	2.53 (0.81)	2.68 (0.87)	-0.20 (-0.43 to 0.04)	0.10
	6 months	2.58 (0.81)	2.48 (0.78)	0.11 (-0.13 to 0.35)	0.36
	9 months	2.67 (0.81)	2.61 (0.75)	0.09 (-0.16 to 0.33)	0.49
	12 months	2.57 (0.82)	2.43 (0.73)	0.17 (-0.08 to 0.42)	0.19
MPI Support**	Overall effect			0.13 (-0.15 to 0.40)	0.36
	Baseline	4.56 (1.31)	4.67 (0.96)		
	3 months	4.48 (1.35)	4.54 (1.17)	0.05 (-0.30 to 0.41)	0.76
	6 months	4.36 (1.41)	4.35 (1.27)	0.17 (-0.18 to 0.53)	0.34
	9 months	4.47 (1.48)	4.40 (1.18)	0.16 (-0.20 to 0.53)	0.38
	12 months	4.56 (1.39)	4.51 (1.18)	0.11 (-0.27 to 0.49)	0.57

		Mean Intervention group (SD)	Mean Control group (SD)	Treatment effect (95%CI)	P value for difference
RAND-36 Physical health**	Overall effect			-2.44 (-7.20 to 2.33)	0.32
	Baseline	45.61 (18.35)	48.35 (18.36)		
	3 months	54.66 (19.54)	50.06 (20.99)	4.20 (-1.49 to 9.89)	0.15
	6 months	52.87 (22.13)	57.73 (19.70)	-5.39 (-11.18 to 0.39)	0.07
	9 months	52.87 (22.04)	54.69 (21.53)	-2.20 (-8.11 to 3.70)	0.46
	12 months	52.73 (22.24)	62.25 (18.33)	-8.72 (-14.77 to -2.67)	0.00
RAND-36 mental health**	Overall effect			-0.23 (-4.16 to 3.56)	0.88
	Baseline	72.49 (15.71)	77.55 (11.95)		
	3 months	74.02 (17.00)	74.84 (16.33)	1.57 (-3.02 to 6.17)	0.50
	6 months	73.04 (16.39)	76.32 (16.02)	-0.82 (-5.48 to 3.84)	0.73
	9 months	74.00 (17.76)	76.18 (13.99)	-0.46 (-5.20 to 4.28)	0.85
	12 months	72.96 (19.17)	76.53 (14.98)	-1.94 (-6.79 to 2.91)	0.43

*Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline. Abbreviation: SD, Standard Deviation; EQ5D; Utility scores based on the EuroQol5D; MPI, Multidimensional Pain Inventory; RAND-36, Research and Development 36 item health survey. * Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100; for GPE, 1-7; for patient satisfaction, 1-7, for MPI 0-6. ** Higher score indicates more quality of life. Range for EQ5D utility: 0-1; for RAND36 0-100.*

Responder analyses

In terms of success, statistically significant more participants in the intervention group decreased in pain by 30% or more three weeks (OR 3.14; 95%CI 1.09 to 9.04) and three months (OR 2.86; 95%CI: 1.27 to 6.46) after the start of the intervention compared to the control group (Table 3). No between-group differences were found when success was defined as ≥2 points reduction on pain.

Sensitivity analyses

Replicating the analyses without the 14 protocol violators, slightly increased the contrast between the intervention- and control group, and significantly more people in the intervention group recovered based on GPE after three months (OR 2.6; 95%CI: 1.03 to 6.87) (Table S2 in the Supplementary Appendix). Furthermore, no significant changes

between the groups based on pain were found when replicating the analyses without participants receiving the intervention after the three month intervention period. Other outcomes did not change from the main analysis (Table S2 in the Supplementary Appendix).

In total, 89 participants (44.1%) had complete follow-up data. There were more missing data in the control group. The complete-cases analysis showed no significant between group differences for pain and GPE. Moreover, the control group showed better results on the ODI for functional status at six months (6.86; 95%CI 1.25 to 12.47), nine months (6.36; 95%CI: 0.75 to 11.97) and 12 months (6.56; 95%CI: 0.95 to 12.17) (Table S2 in the Supplementary Appendix). One complication was recorded during the year follow-up in the intervention group. This was a hematoma, causing extra pain. The participant completely recovered without permanent injury.

Cost-utility analyses

Average total annual societal costs were €10,664 per intervention participant and €11,954 per control group participant. Diagnostic block and RF costs were statistically significantly higher in the intervention group than in the control group. All other between-group differences in total and disaggregate costs were not statistically significant (Table 5).

An ICER of -96,813 was found for QALYs. This indicates that one QALY gained was associated with a societal cost saving of €96,813 (Table 6). Note that this large ICER was due to very small differences in QALYs. Also, the uncertainty surrounding the ICER was large. The CEAC (Figure 2) indicated that the maximum probability of the intervention being cost-effective in comparison with control was 66%, irrespective of the willingness-to-pay.

The overall conclusion of this study does not change when using the results of SA2 (HCA) or SA3(SF6D), however for SA1 (complete-case analysis) the results showed that the intervention group had higher total costs than the control group (Table 6).

Table 5. Mean costs per participant in the intervention and control group, and mean cost differences between both groups during the 12-month follow-up period from a societal perspective

Cost category	Intervention group n=125; mean (SEM)	Control group n=126; mean (SEM)	Mean cost difference crude (95%CI)	Mean cost difference adjusted (95%CI)
Intervention				
• **Diagnostic block**	276.73 (19.88)	174. 17 (24.39)	102.56 (35.71 to 160.27)*	120.97 (35.95 to 186.72)*
• **RF denervation**	768.14 (70.53)	0 (NA)	768.14 (635.14 to 907.50)*	772.55 (619.55 to 950.20)*
Primary healthcare	2123.39 (385.23)	2903.10 (1101.65)	-779.81 (-4112.83 to 650.77)	-882.85 (-4383.10 to 671.52)
Secondary healthcare	1067.83 (154.12)	1521.47 (306.52)	-453.64 (-1196.40 to 168.51)	-417.04 (-1254.74 to 277.19)
Medication	171.23 (218.36)	218.36 (56.53)	-47.13 (-178.13 to 82.74)	-43.23 (-193.51 to 103.22)
Informal care	1026.18 (254.70)	2422.97 (1205.23)	-1396.79 (-4811.29 to 80.37)	-1255.71 (-4866.43 to 194.34)
Absenteeism	3214.50 (959.25)	2610.68 (949.58)	603.82 (-1875.19 to 3100.70)	669.60 (-1923.31 to 3440.17)
Unpaid productivity	2016.21 (436.88)	2103.35 (594.90)	-87.14 (-1366.30 to 1002.44)	-88.55 (-1455.20 to 1165.00)
Total	**10664.19 (1477.29)**	**11954.08 (2589.56)**	**-1289.89 (-7852.24 to 3453.22)**	**-1124.24 (-9111.17 to 4001.44)**

*Abbreviations: n, number; SEM, Standard Error of the Mean; CI, Confidence Interval; NA, Not Applicable; SD, Standard Deviation. Note: Costs are expressed in 2014 Euros. *p<0.05*

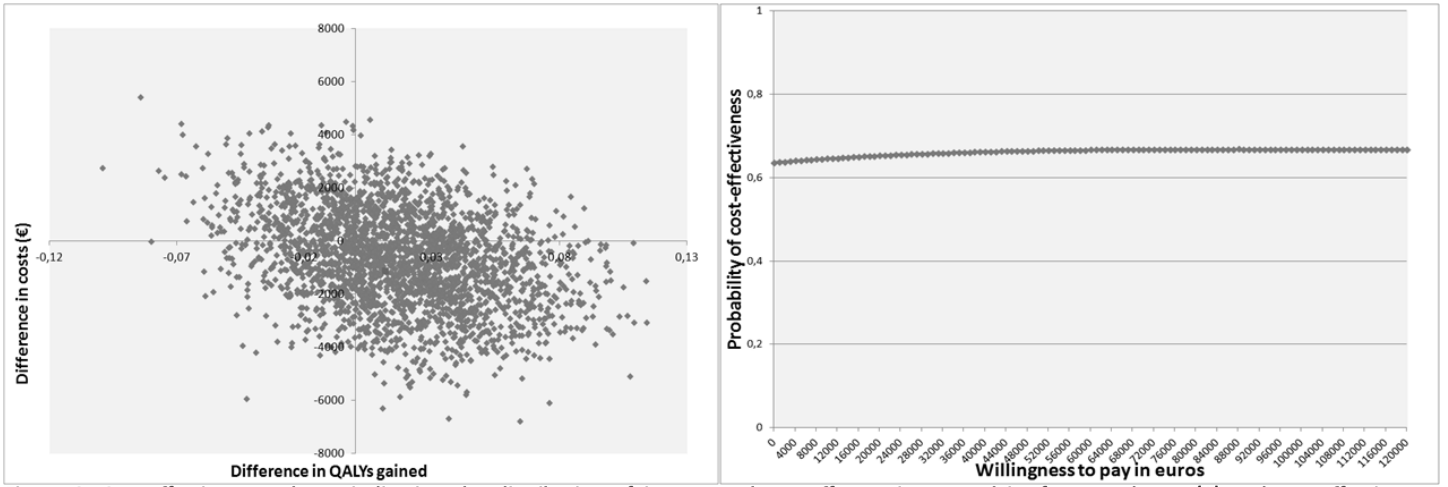

Figure 2. Cost-effectiveness planes indicating the distribution of incremental cost-effect pairs around its four quadrants (1) and cost-effectiveness acceptability curves indicating the probability of adding radiofrequency denervation added to a standardised exercise programme being cost-effective in comparison with a standardised exercise programme alone for different values (€) of willingness-to-pay (2) for QALYS (EQ5D)

Table 6. Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes

Analysis	Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
	Int.	Control		€	Point	€/point	NE	SE	SW	NW
Main analysis – Imputed dataset	103	99	QALY EQ5D (0-1)	-1124.24 (-9034.72 to 3959.14)	0.01 (-0.06 to 0.08)	-96813,2	17.9	44.5	18.7	18.9
SA1 – complete cases	60	46	QALY EQ5D (0-1)	2454.53 (-62.26 to 4971.33)	-0.04 (-0.11 to 0.03)	-58981.6	12.6	16.4	31.4	82.7
SA2 - HCA	103	99	QALY EQ5D (0-1)	-881.46 (-8803.26 to 4195.05)	0.01 (-0.06 to 0.08)	-75906.1	20.4	42.0	17.3	20.3
SA3 – SF6D	103	99	QALY SF6D (0-1)	-1124.24 (-9034.72 to 3959.14)	-0.007 (-0.03 to 0.01)	150292.8	10.3	11.2	52.1	26.5

Abbreviations: CI: Confidence Interval; ICER: Incremental Cost-Effectiveness Ratio; CE-lane: Cost-Effectiveness plane; SA: Sensitivity Analysis; HCA: Human Capital Approach. Note: Costs are expressed in 2014 Euros

DISCUSSION

The aim of this study is to establish whether RF denervation added to a standardised exercise programme is effective and cost-effective compared to a standardised exercise programme alone for patients with CLBP originating from a combination of the facet joint, SI-joint and/or disc who are referred to a pain clinic. We found statistically significant between-group differences for pain (-0.99; 95%CI:-1.73 to -0.25) at three months after start of the intervention favouring the intervention group; and functional status (6.44; 95%CI: 0.61 to 12.26) at 12 months after start of the intervention favouring the control group. Both groups decreased in pain during the 12 months follow-up, but in both groups pain remained (intervention 4.85, and control 4.38). No differences in GPE were found at all follow-up assessments. The maximum probability of a standardised exercise programme with RF denervation being cost-effectiveness for QALYs in comparison with a standardised exercise programme alone was low (<0.66).

The strengths of this RCT are its adequately powered sample size, well-balanced study groups, and 12 month follow-up. Furthermore, this study was the first to evaluate the cost-effectiveness of RF denervation. For the cost-utility analysis, state-of-the-art statistical methods were used, including SUR analyses to account for the possible correlation between costs and effects, and bootstrapping to estimate the uncertainty surrounding cost differences as well as cost-effectiveness. Furthermore, the pragmatic design set in routine clinical care improved the external validity of the results, whereas the internal validity was insured by randomisation. At the same time, the pragmatic design is one of the limitations of this study. Not all participants in the intervention group received RF denervation, because they did not respond to the diagnostic block or provocative discography. This decreased the contrast between the groups. The as-treated analyses (Table S2 and S3) show mostly similar results, however no statistically significant between-group differences in pain after three months were shown. This might be due to the reduced power of this analysis. Secondly, blinding was not possible in this design and patient-reported outcomes were used, which might have resulted in an overestimation of the treatment effect.[40] Thirdly, all effect measures and cost-effect measures were assessed using retrospective questionnaires, which could have introduced recall bias. Fourthly, the amount of missing data is a limitation of this RCT, as complete data could only be collected from 52.5% of the participants. No differences between completers and non-completers were evident in terms of outcomes, and for costs only medication

costs were higher in the non-completer group. However, because of the large amount of drop-outs at 12 months, the complete case analysis can be biased. Therefore, we believe the results of the imputed dataset in the main analysis are more valid. Multiple imputation techniques were used in the main analysis to handle the missing data in the CEA, however still cost and effect estimates are less reliable than those based on a complete dataset. Moreover, it is important to keep in mind that economic evaluation results are not directly transferable between countries due to differences in healthcare and/or social security systems.[41]

Recent systematic reviews that focused on RF denervation for CLBP mostly evaluated the effect of RF denervation on isolated pain sources (pain from the facet joints, SI-joint or discs) based on a positive diagnostic block. Recent reviews have shown evidence of low to moderate quality for small positive effects of facet joint RF denervation on pain and functional status compared to placebo or steroid injections. There is very low to moderate quality evidence and conflicting evidence for SI-joint RF denervation and transdiscal RF denervation.[10,11] In this pragmatic trial, we randomised participants after diagnosis based on medical history and a physical examination and only patients with a multifactorial problem instead of LBP problems originating from an isolated source were included. The latter might have resulted in the inclusion of participants with more complex problems. Furthermore, participants in our trial had back pain complaints for on average nine years and severe baseline pain. A high level of pain is one of the predictors for poor outcomes.[42,43] The complexity of the back pain problems and the severe baseline pain might be reasons for the remaining back pain over the year. In a study by Tubach et al, for example, a patient acceptable symptom state (PASS) for pain of four (NRS 0-10) was found.[44] During the 12-month follow-up, participants in our trial remained above this PASS, although almost 35% of the patients in the intervention group and 36% in the control group indicated to be recovered.

When the results of this RCT are compared to the other RCTs of the MinT study (studying patients with isolated facet joint or SI-joint pain), it shows participants is this RCT also had higher total societal costs, and work absenteeism and unpaid productivity costs in particular compared to those with an isolated source of their LBP. This might be explained by the lower overall functional status scores for participants in this RCT. Furthermore, fewer participants indicated they had recovered during the 12 month follow-up. Although no differences in pain were shown compared to the other RCTs, more problems in functional status and lower success rates might be related to higher

levels of work absenteeism and higher costs.

In conclusion, this pragmatic trial showed that RF denervation added to a standardised exercise programme compared to a standardised exercise programme alone shows statistically significant effects for pain intensity three months after the intervention and that these effects do not remain on the long term. At the 12-month follow-up assessment there was statistically significant difference in functional status in favour of the control group. Furthermore, there were no statistically significant differences on any other outcome or any other follow-up assessment. Finally, the cost-utility analysis indicated that there were no statistically significant differences in cost-effectiveness between a standardised exercise programme with RF denervation versus a standardised exercise programme alone.

Acknowledgements

The authors thank the Netherlands Organisation for Health Research and Development (171202013), the Dutch Society for Anaesthesiology, and the Dutch health insurance companies for funding this study. Merel van Raamt and Robert van Cingel (Sports Medical Centre Papendal) for their help in developing the exercise programme. The authors would also like to thank all participants, physiotherapists, and everyone working on the MinT study at one of the participating pain clinics.

REFERENCES

1. Martin B, Deyo R, Mirza S, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008; 299:656-64
2. Manchikanti L, Singh V, Datta S, Cohen S, Hirsch J. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician.* 2009; 12:35-70
3. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73:968-974.
4. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2013; 9859:2163-96
5. Bogduk N, Dreyfuss P, Govind J. A narrative review of lumbar medial branch neurotomy for the treatment of back pain. *Pain Med*. 2009; 10:1035-45
6. Bogduk N. The anatomical basis for spinal pain syndromes. *J Manipulative Physiol Ther.* 1994; 18:603-605
7. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. *Spine*. 1995; 20:1878-83
8. Manchikanti L, Boswell MV, Singh V, Pampati V, Damron KS, Beyer CD. Prevalence of facet joint pain in chronic spinal pain of cervical, thoracic, and lumbar regions. *BMC Musculoskeletal Disord.* 2004;5:15
9. Kline M. Radiofrequency techniques in clinical practice. Interventional Pain Management. In: Waldman S, Winnie A, eds. *Interventional Pain Mangement.* PhilidelphiaL Sanders; 1996: 1985-217.
10. Poetscher AW, Gentil AF, Lenza M, Ferretti M. Radiofrequency denervation for facet joint low back pain: a systematic review. *Spine*. 2014; 39:E842-E849
11. Center P, Manchikanti L. A Systematic Review and Best Evidence Synthesis of Effectiveness of Therapeutic Facet Joint Interventions in Managing Chronic Spinal Pain. *Pain Physician*. 2015; 18:E535-E82
12. Maas ET, Ostelo RWJG, Niemisto L, et al. Radiofrequency denervation for chronic low back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 10. Art. No.: CD008572. DOI: 10.1002/14651858.CD008572.pub2.
13. Itz C, Willems P, Zeilstra D, Huygen F. Dutch Multidisciplinary Guideline for Invasive Treatment of Pain Syndromes of the Lumbosacral Spine. *Pain Practice.* 2016; 16(1):90-110.
14. Maas E, Juch J, Groeneweg J, et al. Cost-effectiveness of minimal interventional procedures for chronic mechanical low back pain: design of four randomised controlled trials with an economic evaluation. *BMC Musculoskelet Disord.* 2012. 13:260
15. Ostelo R, Deyo R, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine*. 2008;33:90-94
16. Bekkering G, Hendriks H, Koes B, Oostendorp R, Ostelo R, Thomassen-Hilgersom I, van Tulder M. KNGF Guideline low back Pain. *Ned Tijdschr Fysiother.* 2005; 111:1-24.
17. Bogduk N. ISIS Practice Guidelines for Spinal Diagnostic and Treatment Procedures: 2nd Edition. International Spine Intervention Society; 2014.

18. Kallewaard JW, Terheggen MA, Groen GJ, et al. Discogenic low back pain. *Pain Practice.* 2010; 10:560-79
19. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J.* 2015; 24:1127-42
20. Downie W, Leatham P, Rhind V, Wright V, Branco J, Anderson J (1978) Studies with pain rating scales. *Ann Rheum Dis*. 1978; 37:378-81
21. Devilly G, Borkovec T. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry*. 2000; 31:73-86
22. Fairbank J, Pynsent P. The Oswestry disability index. *Spine.* 2000; 25:2940-53
23. Kind P. The EuroQoL instrument: an index of health-related quality of life. *Quality of life and pharmacoeconomics in clinical trials*. Edited by Spilker B. Philadelphia: Lippincott-Raven Publihsers; 1996:191-201
24. Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine.* 2000; 25:3100-03
25. Hays R, Morales L. The RAND-36 measure of health-related quality of life. *Ann Med*. 2001; 33:350-57
26. Lousberg R, Van Breukelen G, Groenman N, Schmidt A, Arntz A, Winter F. Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther* 1999; 37:167-82
27. Lamers LM SP, McDonnell J, Krabbe PFM, van Busschbach JJ. Kwaliteit van leven met in economische evaluaties: het Nederlands EQ-5D tarief. *Ned Tijdschr Geneeskd.* 2005; 149:1574-78
28. Goossens M, Rutten-van Mölken M, Vlaeyen J, van der Linden S. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol.* 2000; 53:688-95
29. Hakkaart-van Roijen L, Tan S, Bouwmans C. Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen; 2011.
30. Riegelman RK. Studying a study and testing a test: how to read the medical evidence. Lippincott Williams & Wilkins, Philadelphia; 2005
31. Koopmanschap MA (2005) PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res*. 2005; 5:23-28
32. Manchikanti L, Pampati V, Fellows B, Bakhit CE (2000) The diagnostic validity and therapeutic value of lumbar facet joint nerve blocks with or without adjuvant agents. *Curr Review Pain.* 2000; 4:337-44
33. Twisk J. Applied multilevel analysis: a practical guide for medical researchers. Cambridge Univ Press, Cambridge; 2006
34. Rubin D. Multiple imputation for nonresponse in surveys, vol 81. John Wiley & Sons, Hoboken New Jersey; 2004
35. Groenwold R, Donders AR, Roes K, Harrell F, Moons K (2011) Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol*. 2012; 175: 210-7
36. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009. 338:b2393
37. Willan A, Briggs A, Hoch J. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health economics*. 2004. 13:461-75

38. Black W. The CE plane a graphic representation of cost-effectiveness. *Med Decis Making*. 1990. 10:212-14
39. Fenwick E, O'Brien B, Briggs A. Cost-effectiveness acceptability curves–facts, fallacies and frequently asked questions. *Health economics*. 2004. 13:405-15
40. Kahan BC, Cro S, Doré CJ, Bratton DJ, Rehal S, Maskell NA, Rahman N, Jairath V (2014) Reducing bias in open-label trials where blinded outcome assessment is not feasible: strategies from two randomised trials. *Trials*. 15:456
41. Tompa E, Culyer AJ, Dolinschi R. Economic evaluation of interventions for occupational health and safety: developing good practice. Oxford University Press, Oxford, 2008
42. Grotle M, Foster NE, Dunn KM, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain*. 2010;151:790-97
43. Costa L, Maher CG, McAuley JH, et al.. Prognosis for patients with chronic low back pain: inception cohort study. *BMJ*. 2009; 339:b3829
44. Tubach F, Ravaud P, Martin-Mola E, et al. Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: results from a prospective multinational study. *Arthit Care & Res.* 2012; 64:1699-07

APPENDIX 1. DESCRIPTION OF DIAGNOSTIC AND TREATMENT PROCEDURES

Diagnostic procedures

1. Facet joint diagnostic block[1]: A 22G needle was inserted to the posterior primary root of the spinal nerve (medial branch) under C-arm fluoroscopy. L3, L4, and L5 were selected for diagnostic blocks. The lateral image was checked to confirm the correct position of the needle, after which 0.5ml 2% lidocaine was injected. No corticosteroids were administered. The diagnostic block was considered positive if the patient reported at least 50% subjective pain reduction 30-90 minutes after the block, after which the patients were randomised.
2. Sacroiliac joint diagnostic block: under C-arm fluoroscopy a 25G needle was inserted 3-10mm laterally (3:00 on the right, 09:00 on the left side) of the sacral foramina S1, S2, and S3. The lateral image was checked to confirm correct depth of the needle, after which 0.5ml lidocaine (2%) was injected. The dorsal ramus of L5 was also blocked as described in the ISIS guidelines,[1] using 0.5ml lidocaine (2%). No corticosteroids were administered. The diagnostic block was considered positive if the participant reported at least 50% subjective pain reduction 30-90 minutes after the blockade. Participants were only randomised if the diagnostic blockade was positive.
3. Provocative discography[2]: The current criterion standard for diagnosing discogenic pain is pressure-controlled provocative discography using strict criteria and at least one negative control level. A full description is stated in the article of Kallewaard et al., 2010.[2] The symptomatic level and the two adjacent levels are examined. The least degenerated or more likely asymptomatic levels are studied first. The patient should be blinded to the discus level and should not be aware of the start of the discus stimulation. The C-arm is first positioned with the direction of the radiation beam parallel to the subchondral plate of the lower vertebral plate of the discus. In the discs above L5-S1, the C-arm is then rotated ipsilaterally until the lateral aspect of the processus articularis overlies the axial middle of the discus to be punctured, and the discus height is at its maximum. In this projection, the needle can be inserted parallel to the direction of the radiation beam and brought into position (tunnel view). The target for the puncturing of the AF is the lateral-middle side of the discus, just lateral to the lateral edge of the processus articularis superior. At the

L5-S1 level, the crista iliaca does not allow access to the discus using a down-the-beam approach. The fluoroscopy tube is rotated until the lateral edge of processus articularis superior of S1 is positioned approximately 25% over the posterior to anterior distance of the corpus vertebrae.

- **Needle positioning** After anesthetizing the skin and the underlying tissue, a one-needle or a two-needle technique can be used to approach the discus. In a two-needle technique, a 20-G needle is advanced over the lateral edge of the processus articularis superior. A 25-G hollow needle is inserted through the needle and into the AF until it reaches the middle of the nucleus. The two-needle technique may help reduce the incidence of discitis and allow entering the discus with needles of a small diameter (e.g., 27G) which might help prevent the incidence of iatrogenic disc degeneration. The needle is carefully advanced to the needle-point end position. Beyond the processus articularis superior, the needle passes through the foramen intervertebrale in the vicinity of the ramus ventralis. In case of paresthesia, the needle must be repositioned. A strong resistance is felt as the needle passes through the annulus. The needle is pushed through the annulus to the centre of the discus. The needle's progress is followed first in AP and then in lateral projection. Ideally, after placement, the needle is situated in the middle of the disc's nucleus, as seen in the AP as well as in the lateral projection
- **Discus stimulation** After verification of the correct needle position, the stylet is removed from the needle and the needle is connected to a contrast agent delivery system which can measure the intradiscal pressure (manometry). The rate of infusion of the contrast agent should not exceed 0.05 mL/s.31–33. This rate reflects a static flow that corresponds to the distension pressure in the discus intervertebralis. If a higher flow is used, false positive discographies can occur due to the resultant pressure peaks. Pain is often provoked by these pressure peaks due to vertebral end-plate compression and distention of the adjacent facet joint. The following parameters must be carefully monitored during the injection of the contrast solution: the opening pressure (OP), the pressure at which contrast is first visible in the discus; the provocation pressure, the pressure greater than the opening pressure at which complaints of pain arise; and the peak pressure or the final pressure at the end of the procedure. The procedure, per level, is continued until the following events: (1) Concordant pain is reproduced at a level of 7 or greater (on a 0 to 10 numeric rating scale; NRS), and subsequent injected volume confirms the response, (2) The volume

infused reaches the 3.0 mL, (3) The pressure rises to 50 psi above opening pressure in discs with a Grade 3 annular tear, (4) If contrast leaks through the outer annulus or through the endplates, one may not be able to pressurize the disc to a pressure sufficient to test the disc sensitivity. In these cases, the rapid manual injection may be acceptable, but must be noted and a negative response is a more defendable response. Assessment criteria. The guidelines of the IASP (International Association for the Study of Pain), as well as those of the ISIS (International Spine Intervention Society), state that two levels must always be tested as controls when performing provocative discography (except if the target disc is that of L5-S1).[1]

- The diagnosis of discogenic pain can only be made if there is reproduction of concordant pain resulting from a pressure that does not produce pain in a normal disc or in an asymptomatic patient.

Treatment procedures:

1. Facet joint radiofrequency denervation: A C-arm image intensifier was positioned in a slightly (10–15°) oblique position to identify skin entry points with the patient in prone position. A 22 G SMK needle with a 10-mm active curved tip was introduced at each entry point. The position of the cannula was checked on the lateral and AP fluoroscopic projection. The depth was adjusted until the tip of the cannula was at the level of a line connecting the posterior aspects of the intervertebral foramen. Sensory stimulation was positive if muscle contraction occurred below 0.6V. Second stimulation at 2 Hz was used in which contraction of the musculus multifidus and no leg contractions should occur. Once the position of the electrode was satisfactory, 1-2 ml per level ml 2% lidocaine was injected and a 90°C 90 seconds RF lesion was made of the medial ramus dorsalis of L3-4, L4-5, and L5-S1.

2. Sacro-iliac joint radiofrequency denervation:
   a. Cooled RF technique[3]: under C-arm fluoroscopy a P/A view of the foramina of S1, S2 and S3 was obtained. A 25G needle was placed as reference point along the inside lateral wall of each foramen, with the tip at the opening of each foramen. An Epsilon ruler (Kimberly Clark Health Care) was used together with the reference needles as landmarks for the lesions. Using the introducer, stylet and probe provided by the manufacturer RF lesions were made at a maximum temperature of 60°C for 2.5 minutes per lesion.

b. Palisade Technique[4]: is done by drawing a cranial-caudal line between the lateral aspect of the sacral foramina and the SI joint line. Under lateral fluoroscopic view 6 20G canulae with 10mm active tips were placed parallel to each other, 10mm apart and perpendicular to the sacrum. Stimulations to 2.0V were done to be sure there was no motor response. Then eight bipolar lesions (90°C for 90 seconds per lesion) were made using adjacent parings of the cannulas.

c. SIMPLICITY III Probe[5]: is a multi-electrode radiofrequency probe, which was inserted at the lateral, inferior border of the sacrum, 10mm below the S4 foramen under fluoroscopic view. The electrode probe was advanced in a cephalad direction along the sacrum, lateral of the foramina, medial to the sacroiliac joint and ventral to the ileum. Using A/P and lateral fluoroscopy the correct position of the electrodes were checked and the RF lesion (85°C for 90 seconds per step) was made. In all three techniques RF lesion of the ramus dorsalis of L5/S1 was carried out monopolar[6]. The lesion sites were adequately anesthetized using 10-20ml of 2% lidocaine.

3. Disc radiofrequency denervation:

a. IDET: Using fluoroscopic control, with the patient prone on the operating table, a needle was passed into the injured disc via the side, with the fewest symptoms. With the needle in the right place, the catheter containing the heating coil was carefully manipulated inside the disc to treat the injured area. The heating generator was connected and the temperature inside the disc was raised to 90°C in 12 minutes, and maintained 90°C for another four minutes.

b. Biacuplasty: Two internally cooled 17 G needles were inserted at the level of the annulus fibrosis. Two RF currents were inserted to generate bipolar configuration. The ideal temperature profile is 55/60°C in the inner posterior disc decreasing to 45°C in the peripheral edge of the posterior disc.

APPENDIX REFERENCES

1. Bogduk N. ISIS Practice Guidelines for Spinal Diagnostic and Treatment Procedures: 2nd Edition. International Spine Intervention Society; 2014.
2. Kallewaard JW, Terheggen MA, Groen GJ, Sluijter ME, Derby R, Kapural L, Mekhail N, Van Kleef M. Discogenic low back pain. *Pain Practice.* 2010; 10:560-79
3. Cohen SP, Hurley RW, Buckenmaier III CC, Kurihara C, Morlando B, Dragovich A. Randomized placebo-controlled study evaluating lateral branch radiofrequency denervation for sacroiliac joint pain. *Anesthesiology*. 2008; 109:279-88
4. Cosman Jr ER, Gonzalez CD. Bipolar radiofrequency lesion geometry: implications for palisade treatment of sacroiliac joint pain. *Pain Practice*. 2011; 11:3-22
5. Schmidt PC, Pino CA, Vorenkamp KE. Sacroiliac joint radiofrequency ablation with a multilesion probe: A case series of 60 patients. *Anesth & Analg*. 2014; 119:460-62
6. Gauci CA, Jankowiak B. Manual of RF Techniques: A Practical Manual of Radiofrequency Procedures in Chronic Pain Management. CoMedical; 2011

Table S1. Baseline characteristics of completers versus non-completers

Characteristics	Intervention Randomised: N=103 Complete baseline: N=98	Intervention Complete N=60 Complete baseline: N=60	Intervention Incomplete N=43 Complete baseline: N=38	Control Randomised: N=99 Complete baseline: N=89	Control Complete N=46 Complete baseline: N=46	Control Incomplete N=53 Complete baseline: N=43
Age in years (SD)	50.80 (11.33)	51.02 (11.51)	50.39 (11.56)	53.31 (10.35)	54.65 (11.00)	52.92 (9.66)
Female (N (%))	64 (65.3%)	48 (80%)	27 (72.9%)	66 (74.2%)	35 (76.1%)	38 (71.7%)
BMI (SD)	26.84 (3.82)	26.70 (3.92)	27.06 (3.69)	26.43 (4.25)	26.29 (3.60)	26.57 (4.86)
Smoker (N (%))	23 (23.5%)	12 (20.0%)	11 (28.9%)	26 (29.2%)	11 (24.4%)	15 (34.1%)
Education						
• Low (N (%))	52 (53.6%)	32 (55.2%)	19 (54.3%)	43 (48.3%)	21 (47.7%)	21 (50.0%)
• Moderate (N (%))	33 (34.0%)	18 (31.0%)	12 (34.3%)	32 (36.0%)	17 (38.6%)	13 (31.0%)
• High (N (%))	12 (12.4%)	8 (13.8%)	4 (11.4%)	14 (15.7%)	6 (13.6%)	8 (19.0%)
History of back pain complaints						
• Time since first experience with low back pain in months (median (IQR))	120.58 (37.32 – 222.04)	124.92 (55.46 – 240.29)	76.03 (30.83 – 124.71)	97.33 (32.33 – 192.58)	121.67 (34.42 – 231.17)	85.17 (27.83 – 176.37)
• Time since current episode with low back pain in months (median (IQR))	36.50 (12.17 – 121.67)	36.50 (14.18 – 119.63)	36.50 (11.88 – 121.67)	32.33 (8.00 – 97.19)	32.33 (11.08 – 121.67)	32.33 (8.00 – 96.77)
Married/living with a partner (N (%))	66 (67.3%)	43 (71.7%)	23 (60.5%)	68 (76.4%)	36 (80.0%)	32 (72.7%)
Expectations						
• Credibility (0-27)	20.10 (4.70)	20.45 (4.10)	19.54 (5.55)	17.07 (5.99)	16.33 (5.95)	17.90 (5.99)
• Expectancy (0-27)	16.88 (5.78)	17.57 (5.85)	15.75 (5.55)	14.38 (6.24)	13.38 (6.42)	15.50 (5.92)
Having a paid job	48 (55.8%)	32 (53.3%)	16 (61.5%)	44 (55.7%)	24 (52.2%)	20 (60.6%)
Outcomes						
Pain intensity in the past week (NRS 0-10)	7.19 (1.43)	7.30 (1.28)	7.03 (1.65)	7.43 (1.41)	4.47 (1.25)	7.39 (1.58)
Oswestry disability index (mean (SD))	39.06 (14.03)	38.00 (13.12)	40.74 (15.38)	37.20 (13.74)	37.16 (15.57)	37.25 (11.54)
Quality of life (EQ-5D)	0.49 (0.28)	0.50 (0.28)	0.47 (0.28)	0.52 (0.28)	0.54 (0.28)	0.49 (0.29)

Abbreviations: SD, Standard Deviation; N, number; BMI, Body Mass Index; EQ-5D, EuroQol-5D

Table S2. As treated analysis for primary outcomes, without 14 protocol violators based on intention-to-treat analyses

		Mean Intervention group (N=93)	Mean Control group (N=95)	Treatment effect (95%CI)	P value for difference	
Primary outcomes						
NRS Pain (SD)*	Overall effect			-0.19 (-0.78 to 0.40)	0.53	
	Baseline	7.28 (1.44)	7.45 (1.43)			
	3 weeks	5.46 (2.26)	6.38 (1.84)	-0.58 (-1.44 to 0.28)	0.19	
	6 weeks	5.37 (2.35)	6.03 (2.00)	-0.31 (-1.10 to 0.47)	0.43	
	3 months	4.74 (2.58)	5.96 (2.35)	-1.04 (-1.82 to -0.25)	0.01	
	6 months	4.84 (2.48)	4.89 (2.63)	0.36 (-0.44 to 1.16)	0.37	
	9 months	4.99 (2.50)	5.26 (2.51)	-0.09 (-0.90 to 0.72)	0.83	
	12 months	4.85 (2.66)	4.41 (2.60)	0.67 (-0.17 to 1.52)	0.12	
ODI Functioning (SD)*	Overall effect			2.52 (-1.79 to 6.83)	0.25	
	Baseline	39.22 (14.12)	37.21 (13.52)			
	3 months	28.03 (15.86)	33.66 (17.15)	-4.54 (-9.69 to 0.61)	0.08	
	6 months	30.75 (19.26)	28.51 (16.93)	5.31 (0.08 to 10.54)	0.05	
	9 months	31.16 (17.99)	28.63 (17.08)	4.24 (-1.06 to 9.54)	0.12	
	12 months	31.38 (17.43)	24.70 (14.84)	7.18 (1.71 to 12.64)	0.01	
		Success % Intervention group	**Success % Control group**	**OR (95%CI)**	**P value for difference**	**NNT**
GPE Success	Overall effect			1.34 (0.75 to 2.40)	0.32	
	3 weeks	23.6	7.5	2.50 (0.71 to 8.81)	0.15	6
	6 weeks	28.6	8.9	2.78 (0.98 to 7.88)	0.05	5
	3 months	37.5	17.1	2.65 (1.03 to 6.87)	0.04	5
	6 months	36.4	38.0	0.69 (0.29 to 1.68)	0.42	-62
	9 months	35.5	31.8	1.12 (0.45 to 2.78)	0.81	27
	12 months	35.3	37.9	0.79 (0.31 to 2.04)	0.63	-38

Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.
Abbreviation: SD, Standard Deviation; NRS, Numeric Rating Scale; GPE, Global Perceived Effect; NNT: Number Needed to Treat
** Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100*

Table S3. Treatment effects for primary outcomes based on an as-treated after 3 months

		Mean Intervention group	Mean Control group	Treatment effect (95%CI)	P value for difference	
Primary outcomes						
NRS Pain (SE)*	Overall effect			-0.22 (-0.85 to 0.40)	0.48	
	Baseline	7.19 (1.43)	7.47 (1.50)			
	3 weeks	5.45 (2.19)	6.66 (1.52)	-0.75 (-1.70 to 0.19)	0.12	
	6 weeks	5.37 (2.29)	5.96 (2.13)	-0.08 (-0.91 to 0.75)	0.85	
	3 months	4.77 (2.48)	5.64 (2.45)	-0.73 (-1.56 to 0.11)	0.09	
	6 months	4.92 (2.42)	5.36 (2.61)	-0.07 (-0.93 to 0.78)	0.87	
	9 months	5.01 (2.47)	5.37 (2.43)	-0.26 (-1.14 to 0.61)	0.55	
	12 months	4.85 (2.65)	4.53 (2.74)	0.56 (-0.34 to 1.45)	0.23	
ODI Functioning (SE)*	Overall effect			1.59 (-2.97 to 6.15)	0.49	
	Baseline	29.06 (14.03)	36.34 (13.50)			
	3 months	28.00 (15.61)	30.96 (17.19)	-3.10 (-8.55 to 2.34)	1.74	
	6 months	30.24 (18.78)	30.09 (18.57)	2.31 (-3.25 to 7.88)	0.42	
	9 months	30.73 (17.76)	27.74 (17.69)	2.89 (-2.83 to 8.60)	0.32	
	12 months	31.20 (17.40)	23.94 (15.83)	6.03 (0.78 to 11.89)	0.04	
		Success % Intervention group	**Success % Control group**	**OR (95%CI)**	**P value for difference**	**NNT**
GPE Success	Overall effect			1.54 (0.77 to 3.07)	0.22	
	3 weeks	22.1	6.1	3.49 (0.68 to 17.85)	0.13	6
	6 weeks	27.8	9.6	2.47 (0.74 to 8.32)	0.14	5
	3 months	34.1	20.0	2.16 (1.96 to 1.96)	<0.0001	7
	6 months	35.3	33.3	0.88 (0.31 to 2.53)	0.81	50
	9 months	35.4	31.7	1.21 (0.40 to 3.62)	0.74	27
	12 months	34.7	30.6	1.19 (0.38 to 3.72)	0.76	24

Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.Abbreviation: NRS, Numeric Rating Scale; GPE, Global Perceived Effect; NNT: Number Needed to Treat
** Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for GPE, 1-7*

Table S4. Treatment effects for complete cases for primary outcomes based on intention-to-treat analyses

		Mean Intervention group	Mean Control group	Treatment effect (95%CI)	P value for difference	
Primary outcomes						
NRS Pain (SD)*	Overall effect			0.17 (-0.52 to 0.85)	0.63	
	Baseline	7.30 (1.28)	7.47 (1.25)			
	3 weeks	5.40 (2.13)	6.39 (1.87)	-0.85 (-1.87 to 0.17)	0.10	
	6 weeks	5.73 (2.09)	5.93 (2.11)	0.03 (-0.89 to 0.95)	0.95	
	3 months	5.10 (2.36)	5.67 (2.41)	-0.60 (-1.52 to 0.31)	0.20	
	6 months	4.98 (2.38)	4.37 (2.51)	0.81 (-0.10 to 1.73)	0.08	
	9 months	5.17 (2.40)	5.02 (2.48)	0.31 (-0.60 to 1.23)	0.50	
	12 months	4.88 (2.71)	4.28 (2.52)	0.80 (-0.12 to 1.72)	0.09	
ODI Functioning (SD)*	Overall effect			4.23 (-0.34 to 8.80)	0.07	
	Baseline	38.00 (13.12)	37.16 (15.57)			
	3 months	38.63 (15.28)	32.87 (15.84)	-2.85 (-8.46 to 2.76)	0.32	
	6 months	30.60 (17.58)	24.96 (14.25)	6.86 (1.25 to 12.47)	0.02	
	9 months	31.03 (17.34)	26.66 (15.15)	6.36 (0.75 to 11.97)	0.03	
	12 months	30.27 (17.33)	25.30 (13.55)	6.56 (0.95 to 12.17)	0.02	
		Success % Intervention group	**Success % Control group**	**OR (95%CI)**	**P value for difference**	**NNT**
GPE Success	Overall effect			0.88 (0.43 to 1.80)	0.72	
	Baseline					
	3 weeks	20.4	6.5	2.77 (0.52 to 14.69)	0.23	7
	6 weeks	23.3	11.4	1.83 (0.50 to 6.68)	0.36	8
	3 months	28.3	17.4	1.85 (0.56 to 6.15)	0.32	9
	6 months	33.3	47.8	0.42 (0.14 to 1.21)	0.11	-6
	9 months	31.7	34.8	0.69 (0.23 to 2.04)	0.50	-32
	12 months	31.7	34.8	0.69 (0.23 to 2.04)	0.50	-32

Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.Abbreviation: SD. Standard Deviation; NRS, Numeric Rating Scale; GPE, Global Perceived Effect; ODI, Oswestry Disability Index Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100; for GPE, 1-7*