Introducing new implants and imaging techniques for lumbar spinal stenosis

Wouter A. Moojen

Colofon

Introducing new implants and imaging techniques for lumbar spinal stenosis

ISBN/EAN: 978-90-6464-815-1

Funding: Paradigm Spine/InSpine funded this trial
Cover illustration: "*Zeelandbrug*" by Agnes Ruijs-Lamp.
Lay-out: Ferdinand van Nispen, Citroenvlinder-dtp.nl, Bilthoven, The Netherlands
Printed by GVO drukkers&vormgevers B.V. | Ponsen & Looijen, Ede, The Netherlands

Introducing new implants and imaging techniques for lumbar spinal stenosis

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 19 november 2014
klokke 13.45 uur

door

Wouter Anton Moojen

geboren te Utrecht, in 1983

Promotiecommissie

Promotor:	Prof. dr. W.C. Peul
Co-promotores:	Dr. M.P. Arts
	Dr. C.L.A.M. Vleggeert-Lankamp
Overige Leden:	Prof. dr. R. Brand
	Prof. dr. J.C. Fairbank (University of Oxford, United Kingdom)
	Prof. dr. F.C. Öner (Universiteit Utrecht)
Paranimfen:	Drs. H.R. Bredenoord
	Dr. R.F. Viergever

Introducing new implants and imaging techniques for lumbar spinal stenosis

Content

Chapter 1	General Introduction	9
Chapter 2	Management of lumbar spinal stenosis: a survey among Dutch spine surgeons	15
Chapter 3	Effectiveness of interspinous implant surgery: a systematic review and meta-analysis	31
Chapter 4	The Felix Trial: design and protocol.	51
Chapter 5	The Felix Trial: clinical results after one year and subgroup analysis	69
Chapter 6	The Felix Trial: clinical results after two years	93
Chapter 7	The Felix Trial: cost-utility analysis	113
Chapter 8	Preoperative MR imaging in patients with intermittent neurogenic claudication: relevance for diagnosis and prognosis	129
Chapter 9	General discussion – scientific evaluation of spinal implants; an ethical necessity	149
Chapter 10	Summary	161
Chapter 11	Samenvatting	167
	List of publications	172
	About the author	174
	Acknowledgements	175

Chapter 1

General Introduction

Intermittent neurogenic claudication (INC) is a complex of symptoms caused by degenerative lumbar spinal stenosis (LSS). [1] The disease was first described by Van Gelderen in 1948 and Verbiest in 1950.[1-4] The characteristic symptoms are leg pain (frequently both legs), which can be exacerbated with prolonged walking and standing and/or lumbar extension, and associated back pain.[1-4] As in other acquired diseases, INC is usually seen in the elderly.[1, 5] Severe stenosis is common in elderly spine: 30.4% in the population have a severe stenosis (average age 67.3).[5] However, only 17.5% of them have typical INC symptoms.

Diagnosis and imaging techniques:

The explanation of the symptoms is often disputed: Verbiest et al stated that the narrowing of the canal (due to degenerative tissue) leads to compression of the nerves and therefore causes symptoms of nerve impairment.[4] Measuring the diameter of the lumbar spinal canal was the main tool to diagnose LSS and subsequently INC. Verbiest started in the 1950s with in vivo measuring techniques.[4] In the 1960s Evans described a complete cerebral spinal fluid stop, on myelography, at the lower lumbar levels in patients with intermittent neurogenic claudication.[1] Imaging techniques were used, after that discovery, preoperatively to select patients for surgical treatment. First myelography was used to prove a complete spinal fluid stop and later Computer Tomography images were used. Nowadays Magnetic Resonance (MR) imaging techniques are used to select patients with INC for treatment.[6] Many different grading scales (such as the Schizas scale) exist to differentiate between mild and severe lumbar spinal stenosis on MR images. Nevertheless, good prospective studies have not been performed to assess the effectiveness and quality of MR imaging techniques and stenosis grading scales.[6]

Surgical treatment:

Surgical treatment is considered to be the gold standard for patients with INC caused by LSS.[1-4] However, the first prospective comparative study proving that surgical therapy was superior in comparison to conservative treatment was published in 2007.[7, 8] The first technique that was described to widen the lumbar spinal canal was wide bony decompression (laminectomy).[9] Until today this technique is widely used. However since, INC is often accompanied by back pain, and thus to postoperative back-pain, it is hypothesized that a wide decompression is a ground for potential instability. Therefore, less invasive techniques, such as laminotomy (partial removal of the lamina), were developed and implemented.[10-12]

Nevertheless, long-term clinical results after surgery are quite poor: only 64% of the patients are satisfied after surgical treatment.[13] Numerous patients still complain about back pain after surgery. In order to solve this 'problem', a French group introduced a new, non-rigid fixation for patients with LSS and associated back pain in 1984: the Wallis system.[14-16] It was a new idea to implant non-rigid implants to indirectly decompress the lumbar canal and to 'unload' the facet joints. The idea was based on their experience of implanting non-rigid implants in other joints. The Wallis system was first implanted in 1986.[14, 15] Nowadays, (other) interspinous process devices (IPDs) are used in the treatment for LSS and also others for back pain.[17-19] The X-stop and Coflex implants were first used in the USA in an FDA trial.[17] In contrast, in the European countries surgeons started to implant these devices right away (not in any prospective study design).[14, 15] As a result, these implants are widely used for almost 30 years to treat patients with INC caused by LSS.

Objective and outline of this thesis

The main objective of this thesis is to compare bony decompression with implantation of interspinous process devices (IPDs) in patients with intermittent neurogenic claudication (INC) caused by lumbar spinal stenosis (LSS). At the start of this research project, no double blind randomized study on this subject was published. However, implantation of IPDs was already part of the daily practice in some Dutch neurosurgical and orthopedic clinics. In **chapter 2** a national survey among Dutch spine surgeons is presented about the usual care of patients with intermittent neurogenic claudication caused by lumbar spinal stenosis. Surgeons' expectations of different treatment options are presented. The existing evidence on interspinous implant surgery will be systematically reviewed in **chapter 3.** Results of treatment with IPDs are compared with other (conservative) treatment options. In **chapter 4** the design of the Foraminal Enlargement Lumbar Interspinosus distraXion (FELIX) trial is described. This double-blind, multicenter, randomized (cost)effectiveness study was designed to answer the question whether treatment with IPDs would be more (cost) effective compared with conventional bony decompression. Short-term results (eight weeks), long-term results (one year) and results in different subgroups are described in **chapter 5.** The two-year results are presented in **chapter 6.** The analysis based on total direct and indirect costs of both procedures (treatment with IPD and bony decompression) are presented in **chapter 7**.

Measuring the amount of lumbar stenosis via estimating the compression on lumbar magnetic resonance imaging is now considered standard in the preoperative work-up in patients with INC suitable for surgical treatment. However, correlation between baseline clinical complaints and the compression on MR images has never been evaluated. Furthermore, the prognostic value of compression on MR images at baseline for the extent of possible recovery at long-term clinical follow-up is also not known. Results of the degree of compression on MR images compared with clinical functioning at baseline and long-term follow-up are presented in **chapter 8.**

Without any good prognostic evidence, new devices have been implanted to treat patients with INC. Drugs are introduced using robust introduction models. However, it is possible to use spinal implants in patients without any good clinical evidence. Ethical considerations of the introduction of these devices are presented in **chapter 9**.

Summary and conclusions are presented in **chapter 10** in English and in **chapter 11** in Dutch.

Reference List

1. EVANS JG. Neurogenic Intermittent Claudication. Br Med J 1964;2(5415):985-987.
2. Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. Ned Tijdschr Geneeskd 1950;94(33):2415-2433.
3. Verbiest H. [Further reports on primary stenosis of the lumbar spinal canal in adults.]. Ned Tijdschr Geneeskd 1951;95(27):1965-1970.
4. Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. J Bone Joint Surg Br 1954;36-B(2):230-237.
5. Ishimoto Y, Yoshimura N, Muraki S et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20(10):1103-1108.
6. de Schepper EI, Overdevest GM, Suri P et al. Diagnosis of lumbar spinal stenosis: an updated systematic review of the accuracy of diagnostic tests. Spine (Phila Pa 1976) 2013;38(8):E469-E481.
7. Malmivaara A, Slatis P, Heliovaara M et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 2007;32(1):1-8.
8. Weinstein JN, Tosteson TD, Lurie JD et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008;358(8):794-810.
9. Mixter WJ, Barr J. Rupture of the intervertebral disc with involvement of the spinal canal. N Engl J Med 1934;211:210-215.
10. Fox MW, Onofrio BM, Onofrio BM, Hanssen AD. Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. J Neurosurg 1996;85(5):793-802.
11. Fox MW, Onofrio BM. Indications for fusion following decompression for lumbar spinal stenosis. Neurosurg Focus 1997;3(2):e2.
12. Thome C, Zevgaridis D, Leheta O et al. Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. J Neurosurg Spine 2005;3(2):129-141.
13. Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. Spine (Phila Pa 1976) 1992;17(1):1-8.
14. Senegas J, Etchevers JP, Vital JM, Baulny D, Grenier F. [Recalibration of the lumbar canal, an alternative to laminectomy in the treatment of lumbar canal stenosis]. Rev Chir Orthop Reparatrice Appar Mot 1988;74(1):15-22.
15. Senegas J. [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. Acta Orthop Belg 1991;57 Suppl 1:221-226.
16. Senegas J. Mechanical supplementation by non-rigid fixation in degenerative intervertebral lumbar segments: the Wallis system. Eur Spine J 2002;11 Suppl 2:S164-S169.
17. Davis RJ, Errico TJ, Bae H, Auerbach JD. Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective, randomized, multicenter, Food and Drug Administration Investigational Device Exemption trial. Spine (Phila Pa 1976) 2013;38(18):1529-1539.
18. Moojen WA, Arts MP, Jacobs WC et al. Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. BMJ 2013;347:f6415.
19. Richter A, Halm HF, Hauck M, Quante M. 2-year Follow-up After Decompressive Surgery With and Without Implantation of an Interspinous Device for Lumbar Spinal Stenosis: A Prospective Controlled Study. J Spinal Disord Tech 2012.

Chapter 2

Management of lumbar spinal stenosis: a survey among Dutch spine surgeons

Extracted from Acta Neurochir (Wien). 2014 Aug 7;

Overdevest GM (A1) & Moojen WA (A1/A2)
Arts MP (A2)
Vleggeert-Lankamp CLAM (A1)
Jacobs WCH (A1)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands

Abstract

Background Various surgical and nonsurgical treatments for lumbar spinal stenosis (LSS) are widely adopted in clinical practice but high quality randomized controlled trials to support these are often lacking, especially in terms of their relative benefit and risk compared with other treatment options. Therefore, an evaluation of agreement among clinicians regarding the indications and the choice for particular treatments seems appropriate.

Methods One hundred six Dutch neurosurgeons and orthopedic spine surgeons completed a questionnaire, which evaluated treatment options for LSS and expectations regarding the effectiveness of surgical and non-surgical treatments.

Results Responders accounted for 6971 decompression operations and 831 spinal fusion procedures for LSS annually. Typical neurogenic claudication, severe pain/neurologic symptoms, and a pronounced constriction of the spinal canal were considered the most important indications for surgical treatment by the majority of responders. Non-surgical treatment was generally regarded ineffective and believed to be less effective than surgical treatment. Interlaminar decompression was the preferred technique by 68% of neurosurgeons and 52% orthopedic surgeons for the treatment of LSS. Concomitant fusion was applied in 12% of all surgery for LSS. Most surgeons considered spondylolisthesis as an indication and spinal instability as a definite indication for additional fusion.

Conclusions The current survey demonstrates a wide variety of Dutch spine surgeons' preferred treatment of symptomatic LSS. To minimize variety, national and international protocols based on high-quality randomized controlled trials and systematic reviews are necessary to give surgeons more tools to support everyday decision-making.

Introduction

Lumbar spinal stenosis (LSS) is the most common reason for lumbar surgery among people over the age of 65 [10]. Surgery rates of LSS have increased markedly over the past decade [35]. The incidence of lumbar surgery for in the United States of America was 1.4 per 1000 and patient costs amounted nearly 1.65 billion US dollars in 2009 [10]. In comparison, the annual incidence of lumbar surgery for LSS in the Netherlands was 0.9 per 1000 according to Dutch hospital payment data in 2007. A recent meta-analysis of randomized controlled trials concluded that surgical decompression with or without fusion, or implantation of interspinous process devices (IPD) is more effective than non-surgical treatment [21, 36]. The gold standard surgical treatment of symptomatic LSS is a facet-preserving laminectomy [13]. More recently, less invasive techniques such as a unilateral and bilateral laminotomies, spinous process osteotomies and indirect decompression with IPDs, have been developed to minimize tissue damage and prevent surgery-induced instability [3, 6, 12, 14, 24, 28, 30, 37, 38]. In cases of severe spinal deformity and marked instability, concomitant fusion procedures may be necessary [20]. Currently, various surgical and non-surgical treatments are widely adopted in clinical practice but high quality randomized controlled trials to support these are often lacking, especially in terms of their relative benefit and risk compared with other treatment options. Various authors state that the lack of evidence is reflected by large geographic variation of spine surgery rates and choices for particular surgical procedure [4, 7, 9, 10, 35]. Therefore, an evaluation of agreement among clinicians about the indications and the choice for particular treatments seems appropriate. Accordingly, a survey on the management of patients with LSS was conducted among neurosurgeons and orthopedic spine surgeons in the Netherlands.

Methods

In 2011, all 195 neurosurgeons of The Netherlands Neurosurgeon Association and orthopedic surgeons of the Dutch Spine Society, were asked to fill in a questionnaire. Subjects who did not respond were sent a reminder after 2 months. The questionnaire referred to various aspects of the treatment of symptomatic LSS. The questionnaire consisted of 22 questions regarding: 1) Surgeons' characteristics: age, sex, years of clinical experience and number of annually performed surgical procedures for LSS; 2) Preferred surgical technique and use of concomitant fusion or indirect decompression with IPD's; 3) Perceived effectiveness, severity of postoperative low back pain and complication incidence of various surgical techniques; 4) Considerations for surgery; 5) Effectiveness of non-surgical treatment; 6) Considerations for non-surgical treatment. On six items we asked the surgeon's opinion according to a 5-point Likert scale, ranging from 'least' to 'most' or 'worse' to 'excellent'. For the analysis, these five categories were reduced to three, by merging the opposite categories and retaining the intermediate/neutral category. Data were analyzed using descriptive statistics. Comparisons were made using a t-test in case of continuous outcome data and a chi-square test in case of categorical data. All frequencies were based on the total number of valid responders. IBM SPSS software, version 20.0, was used for all statistical analysis.

Results

Surgeons' characteristics

106 out of 195 questionnaires were returned (response rate 54%). Among the responders there were 102 male and 4 female surgeons. The majority of the responders were neurosurgeon (62%).

The median clinical experience of the responding neurosurgeons and orthopedic surgeons was 13.2 years. On average the responding neurosurgeon performed 82 decompressions annually, compared with 41 decompressions by orthopedic spine surgeons ($p=0.001$, t-test). The responders performed a total of 6971 decompression operations annually (Table 1).

Table 1. Baseline characteristics.

Responder's characteristics	Number of responders (%)
Number of responders	106 of 195 (54%)
Male	102 (96%)
Neurosurgeon	66 (62%)
Orthopedic surgeon	40 (38%)
Years of experience	13.2 – Interquartile range (IQR) 14
Number decompressions/year	
0-10	14 (13.2%)
11-25	14 (13.2%)
26-50	31 (29.2%)
51-75	11 (10.4%)
76-100	18 (17.0%)
>100	18 (17.0%)
Perform (occasional) concomitant fusion procedures	
0	48 (45.3%)
1-10	31 (29.2%)
11-20	18 (17.0%)
21-30	6 (5.7%)
>30	3 (2.8%)

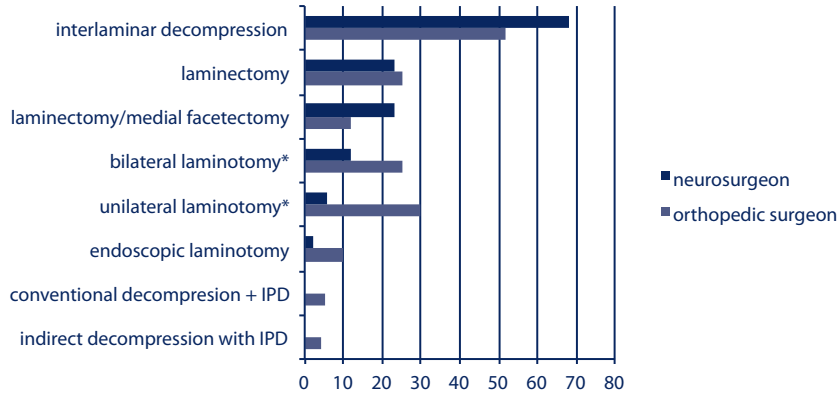
Indications for surgery and effectiveness of non-surgical treatment

Typical neurogenic claudication, severe pain/symptoms and a pronounced constriction of the lumbar spinal canal were considered the most important indications for surgical treatment by more than half of responders (figure 1). 43% of the responders required failure of non-surgical treatment before scheduling surgery among all cases of LSS. 33% required prior non-surgical treatment only in case of mild symptoms or moderate severity of stenosis, and 25% did not require prior non-surgical treatment at all. The majority of responders required 3 months of adequate non-surgical treatment (55%), whilst 17% required non-surgical treatment for at least 6 months. As part of the non-surgical treatment regime, pain medication was most frequently reported as effective (31% of responders). Physical therapy, counseling by the general practitioner or neurologist, epidural injections, and the use of an orthosis were generally regarded ineffective (Figure 2). Non-surgical treatment was regarded most effective among subjects <70 years, whilst symptoms refractory to prior decompression, long-lasting symptoms, concomitant spondylolisthesis, and multiple level stenosis were regarded to be associated with worse treatment outcome (Figure 3).

Surgical Procedure Characteristics

Both neurosurgeons (68%) and orthopedic surgeons (52%) preferred an interlaminar decompression in most cases of LSS (Figure 4).

Figure 4. Surgeons' preferred technique to treat patients with LSS



A total laminectomy was the preferred technique in 24% of responders and a laminectomy including medial facetectomy was the preferred technique in 19% of responders. Only one orthopedic surgeon regularly performed indirect decompression with IPDs and one orthopedic surgeon routinely used IPDs in addition to conventional decompression. Endoscopic laminotomy was performed by only 4% of responders. Unilateral laminotomy was the preferred technique in 30% of orthopedic spine surgeons compared with 6% of neurosurgeons (p=0.001). Bilateral laminotomy was the preferred technique in 25% of orthopedic spine surgeons compared with 12% of neurosurgeons (p=0.001). Other differences regarding preferred techniques among neurosurgeons and orthopedic surgeons were not significant.

Expectations for surgical outcome

The surgeons' expectations regarding the effectiveness of surgery, postoperative low back pain severity, and complication incidence of different lumbar decompression techniques are listed in Table 2. Laminectomy, including medial facetectomy, and interlaminar decompression were expected to be most effective (83% of surgeons), whereas IPD and endoscopic laminotomy were expected to be least effective (54 and 51% of surgeons, respectively). Most postoperative back pain was

expected after laminectomy with medial facetectomy and laminectomy without medial facetectomy (36% and 25% of surgeons, respectively), and least after IPD and endoscopic laminotomy (80% and 77% of surgeons, respectively). Surgical complications were expected to be highest with laminectomy including medial facetectomy and endoscopic laminotomy (19% and 14% of surgeons, respectively) and lowest with interlaminar decompression laminectomy, unilateral or bilateral laminotomy and IPD (82, 81, 80 and 92% of surgeons, respectively). Overall, surgeons expected a satisfactory outcome of 76% after lumbar decompression for LSS at short-term follow-up (8 weeks) and 65% at long-term follow-up (5 years).

Table 2. Surgeons' expectations about the results of different lumbar decompression techniques.

	Expected effectiveness			Expected low back pain			Expected complications		
	most	neutral	least	most	neutral	least	most	neutral	least
Laminectomy including medial facetectomy	83.3	7.1	9.5	36.0	34.8	29.2	19.3	19.3	61.4
Laminectomy	67.9	17.9	14.3	25.3	42.2	32.5	7.2	26.5	66.3
Interlaminar decompression	83.0	13.6	3.4	6.8	29.5	63.6	1.1	17.0	81.8
Unilateral laminotomy	58.8	19.1	22.1	9.0	26.9	64.2	2.9	15.7	81.4
Bilateral laminotomy	52.2	21.7	26.1	14.3	27.1	58.6	5.6	14.1	80.3
IPD	20.3	25.4	54.2	6.6	13.1	80.3	4.7	3.1	92.2
Endoscopic laminotomy	25.6	23.3	51.2	5.9	17.6	76.5	14.0	17.5	68.4

The numbers shown are percentages of valid responses. IPD; interspinous process devices.

Table 3. Indirect decompression with IPDs.

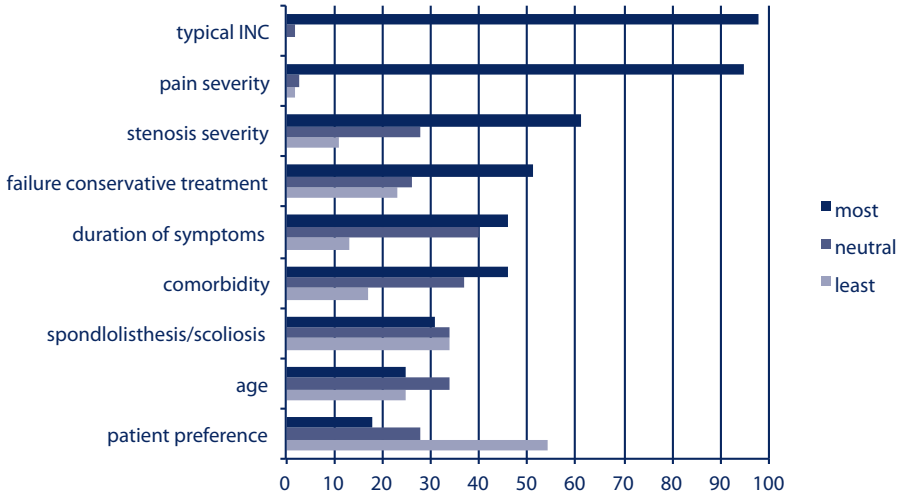
	Number of responders (%)
Number of surgeons performing operations with IPD	33 (31%)
Number of operations with IPD annually	303
Indication for IPD placement	
· Neurogenic claudication	29 (88%)
· Low back pain	2 (6%)
· Lateral foramen stenosis	13 (39%)
· Part of a rigid fixation	1 (3%)
· Part of a dynamic fixation	3 (9%)
· After discectomy with herniated disc	2 (6%)
Implant used	
· Coflex	19 (58%)
· X-stop	4 (12%)
· Aperius	5 (15%)
· Diam	2 (6%)
· ISS	2 (6%)
· other	1 (3%)

IPD; interspinous process devices.

Concomitant spinal fusions

58 of the 106 responders performed concomitant fusion procedures if deemed necessary, of which 49.2% were neurosurgeon and 65.0% orthopedic surgeon (p=0.115). These 58 surgeons annually performed 831 spinal fusion procedures for LSS. Additional fusion was required in 12% of 6971 procedures for LSS in our survey. Additional fusion was deemed necessary in the presence of spondylolisthesis grade 2 or more and documented spinal instability by 79.3% and 74.1% of respondents, respectively. Only 24.1% considered fusion necessary in the presence of low back pain, and 17.2% in case of isthmic spondylolisthesis without instability. 58.6% of responders considered fusion in case of spondylolisthesis grade 1.

Figure 1. Indications for surgical treatment



Frequency and indications for indirect decompression with IPDs

33 spinal surgeons (33%) performed indirect lumbar decompression with IPDs. They accounted for a total of 303 IPD implantations annually, among a total of 6971 procedures for LSS in our survey. This number does not include the use of 70 IPDs within study protocols [23]. Neurogenic claudication was the most frequent indication for treatment with IPD (Table 3).

Figure 2. Effectiveness of conservative treatment

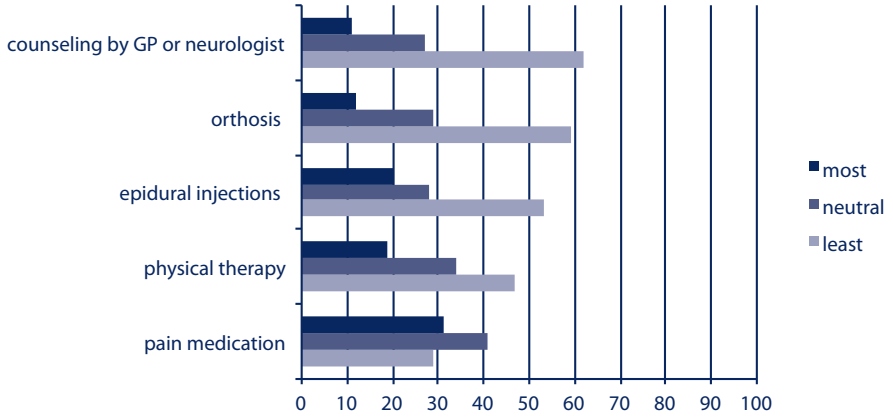
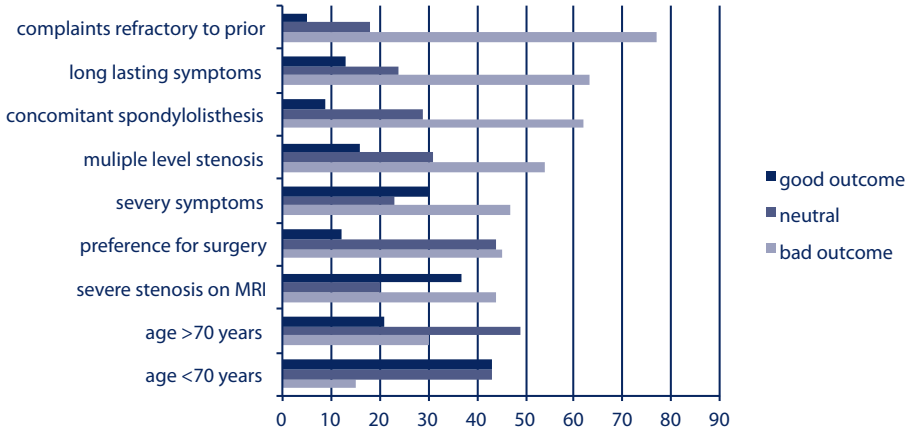


Figure 3. Predictive factors for good outcome



Discussion

The results of our survey among Dutch neurosurgeons and orthopedic surgeons reflect the current clinical management of patients with LSS in the Netherlands. Typical neurogenic claudication and severe pain/symptoms were considered the most important indications for surgical treatment by nearly all responders. In agreement with literature, these symptom characteristics are considered important predictors of good outcome of surgical treatment [19, 25]. The extent of stenosis was regarded as an important indication for surgery by slightly fewer responders. Accordingly, previous surgical studies have demonstrated a significant correlation between good treatment outcome and a pronounced constriction of the spinal canal [2, 15, 16]. Conversely, the duration of symptoms [16, 27] and the presence of comorbid conditions (i.e. cardiovascular, diabetes, and hip joint arthrosis) [2, 11, 18] negatively influence the results of decompression surgery.

Limited evidence is available regarding the efficacy of non-surgical treatments in patients with LSS [33, 34]. Nonsurgical treatments have most frequently been compared with other treatments, rather than to the natural course of LSS [31]. Currently, the choice of treatment is mainly guided by clinical judgment, observational studies, and in analogy to other spine studies. Physical therapy, pain medication, epidural injections, orthosis, and counselling are widely adopted in the current clinical practice, despite questionable evidence of their efficacy. The responders in our survey reported the effectiveness of these non-surgical treatment strategies to be generally low and predictors of non-surgical treatment outcome varied considerably. Complaints refractory to prior decompression, long-lasting symptoms, concomitant spondylolisthesis, and multiple level stenosis were associated with bad non-surgical treatment outcome according to the majority of the responders.

The most commonly performed surgical technique for the treatment of LSS was interlaminar decompression. As narrowing of the spinal canal predominantly takes place at the interlaminar region involving the facet joints and ligamentum flavum, resection of the whole lamina may not always be necessary to effectively decompress the spinal canal [8, 29]. Recent publications documented the use of even less invasive techniques that avoid the removal of midline structures (spinous processes and inter- and supraspinous ligaments), which are designed

to preserve spinal integrity, through minimizing the disruption of back muscles and the posterior ligaments [6, 12, 14, 26, 30, 37]. These studies reported favorable treatment outcomes and post-operative low back pain of (endoscopic) unilateral and bilateral laminotomy compared with conventional laminectomy, and no significant differences regarding the incidence of postoperative complications [6, 12, 14, 26, 30, 37]. In contrast to the findings of these studies the expected effectiveness of these techniques was lower than of conventional midline decompression techniques in our survey. Possibly, the small patient populations and generally poor methodologic quality of these studies caused this discrepancy. Further, laminotomy as opposed to laminectomy may not be suitable to all patients with LSS. In cases of extensive stenosis on MRI or intraoperative doubt of adequate decompression it may be necessary to perform a laminectomy [26]. Therefore, results from studies which compare the results of laminotomy with laminectomy should be interpreted cautiously [26]. Nevertheless, it seems appropriate to tailor the technique for decompression to the severity of stenosis. The reported long-term effectiveness of surgery in our survey (65% satisfactory outcome after 5 years) was in agreement with the pooled estimate of a meta-analysis, which documented a long-term satisfactory outcome after surgery in 64% of the patients [32].

In our survey, 58 (55%) surgeons performed 831 fusion procedures in patients with LSS annually. This amounts to 12% of all 6971 procedures for LSS in our survey. Deyo et al. reported that the rate of spinal stenosis surgery in the Medicare population remained stable, but the rate of concomitant fusion procedures increased 15-fold to 15% of all surgery for spinal stenosis in the period 2002-2007 [10]. These concomitant fusion procedures were independently associated with increased perioperative mortality, major complications, rehospitalization and costs, and therefore should only be performed when necessary. However, the indications for additional spinal fusion are often unclear [20], and occasionally surgeons' preferences may outweigh patients' and disease characteristics in the choice of surgical procedures [5, 17]. Concomitant fusion procedures were deemed necessary in the presence of spondylolisthesis grade 2 or more by 79.3% of responders, and documented spinal instability by 74.1% of responders. Low back pain as an indication for fusion is controversial [1], and was reported by 24.1% of the spine surgeons. Thirty-one percent of the responders occasionally performed indirect decompression with IPDs, despite limited evidence to support this surgical treatment [22]. Previous randomized controlled trials compared the effectiveness

of decompression with IPDs with nonsurgical treatment, rather than conventional decompression [3, 38]. Only recently, a randomized controlled trial comparing decompression with IPDs with conventional midline decompression reported no advantage of decompression with IPDs over conventional decompression. Moreover, a significantly increased reoperation rate among patients treated with IPDs was reported [24].

The present study is subject to several limitations. Spine surgeons' responses may be subject to a recall bias and reliability regarding controversial subjects in particular (e.g. indications for concomitant fusion, use of IPDs) cannot be assured. Definitions of surgical techniques were not provided in detail, which may result in misclassification or heterogeneity amongst the techniques reported. Furthermore, 106 out of 195 spine surgeons completed the questionnaire, thus our data may not be representative for all neurosurgeons and orthopedic spine surgeons in the Netherlands.

The current survey demonstrates a wide variety of Dutch spine surgeons' preferred treatment of symptomatic LSS. The observed variation reflects the limited available evidence regarding the effectiveness of non-surgical treatment, the effectiveness of various surgical techniques, and indications for concomitant fusion. Generally, failure of non-surgical treatment was considered as a pre-requisite to a surgical procedure to be performed, but paradoxically non-surgical treatment was often regarded ineffective. Surgeon's expectations of surgical treatment were in agreement with the modest long-term outcomes reported in literature. Most variation was observed regarding surgeons' preferences for surgical and non-surgical treatments of symptomatic LSS and indications for concomitant fusion procedures. To minimize variety, national and international protocols based on high-quality randomized controlled trials and systematic reviews are necessary to give surgeons more tools to support everyday decision-making.

Reference List

1. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 15 Suppl 2:S192-300
2. Airaksinen O, Herno A, Turunen V, Saari T, Suomlainen O (1997) Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. Spine (Phila Pa 1976) 22:2278-2282
3. Anderson PA, Tribus CB, Kitchel SH (2006) Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. J Neurosurg Spine 4:463-471
4. Birkmeyer NJ, Weinstein JN (1999) Medical versus surgical treatment for low back pain: evidence and clinical practice. Eff Clin Pract 2:218-227
5. Carragee EJ (2010) The increasing morbidity of elective spinal stenosis surgery: is it necessary? JAMA 303:1309-1310
6. Celik SE, Celik S, Goksu K, Kara A, Ince I (2010) Microdecompressive laminatomy with a 5-year follow-up period for severe lumbar spinal stenosis. J Spinal Disord Tech 23:229-235
7. Chen E, Tong KB, Laouri M (2010) Surgical treatment patterns among Medicare beneficiaries newly diagnosed with lumbar spinal stenosis. Spine J 10:588-594
8. Delank KS, Eysel P, Zollner J, Drees P, Nafe B, Rompe JD (2002) [Undercutting decompression versus laminectomy. Clinical and radiological results of a prospective controlled trial]. Orthopade 31:1048-1056; discussion 1057
9. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI (2005) United States trends in lumbar fusion surgery for degenerative conditions. Spine (Phila Pa 1976) 30:1441-1445; discussion 1446-1447
10. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG (2010) Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 303:1259-1265
11. Freedman MK, Hilibrand AS, Blood EA, Zhao W, Albert TJ, Vaccaro AR, Oleson CV, Morgan TS, Weinstein JN (2011) The impact of diabetes on the outcomes of surgical and nonsurgical treatment of patients in the spine patient outcomes research trial. Spine (Phila Pa 1976) 36:290-307
12. Fu YS, Zeng BF, Xu JG (2008) Long-term outcomes of two different decompressive techniques for lumbar spinal stenosis. Spine (Phila Pa 1976) 33:514-518
13. Gibson JN, Waddell G (2005) Surgery for degenerative lumbar spondylosis. Cochrane Database Syst Rev:CD001352
14. Gurelik M, Bozkina C, Kars Z, Karadag O, Ozum U, Bayrakli F (2012) Unilateral Laminotomy For Decompression of Lumbar Stenosis is Effective and Safe: A Prospective Randomized Comparative Study. Journal of Neurological Sciences (Turkish) 29:744-753
15. Herno A, Airaksinen O, Saari T, Miettinen H (1994) The predictive value of preoperative myelography in lumbar spinal stenosis. Spine (Phila Pa 1976) 19:1335-1338
16. Jonsson B, Annertz M, Sjoberg C, Stromqvist B (1997) A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part II: Five-year follow-up by an independent observer. Spine (Phila Pa 1976) 22:2938-2944
17. Katz JN, Lipson SJ, Lew RA, Grobler LJ, Weinstein JN, Brick GW, Fossel AH, Liang MH (1997) Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. Spine (Phila Pa 1976) 22:1123-1131
18. Katz JN, Stucki G, Lipson SJ, Fossel AH, Grobler LJ, Weinstein JN (1999) Predictors of surgical outcome in degenerative lumbar spinal stenosis. Spine (Phila Pa 1976) 24:2229-2233
19. Kleinstueck FS, Fekete T, Jeszenszky D, Mannion AF, Grob D, Lattig F, Mutter U, Porchet F (2011) The outcome of decompression surgery for lumbar herniated disc is influenced by the level of concomitant preoperative low back pain. Eur Spine J 20:1166-1173

20. Knaub MA, Won DS, McGuire R, Herkowitz HN (2005) Lumbar spinal stenosis: indications for arthrodesis and spinal instrumentation. Instr Course Lect 54:313-319
21. Kovacs FM, Urrutia G, Alarcon JD (2011) Surgery versus conservative treatment for symptomatic lumbar spinal stenosis: a systematic review of randomized controlled trials. Spine (Phila Pa 1976) 36:E1335-1351
22. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC (2011) Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. Eur Spine J 20:1596-1606
23. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC (2010) The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord 11:100
24. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW, Vleggeert-Lankamp CL, Peul WC (2013) Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. BMJ 347:f6415
25. Pearson A, Lurie J, Tosteson T, Zhao W, Abdu W, Weinstein JN (2012) Who Should Have Surgery for Spinal Stenosis? Treatment Effect Predictors in SPORT. Spine (Phila Pa 1976) 37:1791-1802
26. Postacchini F, Cinotti G, Perugia D, Gumina S (1993) The surgical treatment of central lumbar stenosis. Multiple laminotomy compared with total laminectomy. J Bone Joint Surg Br 75:386-392
27. Radcliff KE, Rihn J, Hilibrand A, Dilorio T, Tosteson T, Lurie JD, Zhao W, Vaccaro AR, Albert TJ, Weinstein JN (2011) Does the duration of symptoms in patients with spinal stenosis and degenerative spondylolisthesis affect outcomes?: analysis of the Spine Outcomes Research Trial. Spine (Phila Pa 1976) 36:2197-2210
28. Rajasekaran S, Thomas A, Kanna RM, Prasad Shetty A (2013) Lumbar spinous process splitting decompression provides equivalent outcomes to conventional midline decompression in degenerative lumbar canal stenosis: a prospective, randomized controlled study of 51 patients. Spine (Phila Pa 1976) 38:1737-1743
29. Rompe JD, Eysel P, Zollner J, Nafe B, Heine J (1999) Degenerative lumbar spinal stenosis. Long-term results after undercutting decompression compared with decompressive laminectomy alone or with instrumented fusion. Neurosurg Rev 22:102-106
30. Thome C, Zevgaridis D, Leheta O, Bazner H, Pockler-Schoniger C, Wohrle J, Schmiedek P (2005) Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. J Neurosurg Spine 3:129-141
31. Tran de QH, Duong S, Finlayson RJ (2010) Lumbar spinal stenosis: a brief review of the nonsurgical management. Can J Anaesth 57:694-703
32. Turner JA, Ersek M, Herron L, Deyo R (1992) Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. Spine (Phila Pa 1976) 17:1-8
33. van Tulder MW, Koes BW, Bouter LM (1997) Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. Spine (Phila Pa 1976) 22:2128-2156
34. Watters WC, 3rd, Baisden J, Gilbert TJ, Kreiner S, Resnick DK, Bono CM, Ghiselli G, Heggeness MH, Mazanec DJ, O'Neill C, Reitman CA, Shaffer WO, Summers JT, Toton JF (2008) Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. Spine J 8:305-310
35. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES (2006) United States' trends and regional variations in lumbar spine surgery: 1992-2003. Spine (Phila Pa 1976) 31:2707-2714
36. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goldberg H, Berven S, An H (2008) Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 358:794-810
37. Yagi M, Okada E, Ninomiya K, Kihara M (2009) Postoperative outcome after modified unilateral-approach microendoscopic midline decompression for degenerative spinal stenosis. J Neurosurg Spine 10:293-299

38. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, Johnson DR, 2nd, Skidmore GA, Vessa PP, Dwyer JW, Puccio ST, Cauthen JC, Ozuna RM (2005) A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. Spine (Phila Pa 1976) 30:1351-1358

Chapter 3

Effectiveness of interspinous implant surgery: a systematic review and meta-analysis

Extracted from Eur Spine J. 2011 Oct;20(10):1596-606

Moojen WA (A1/A2)
Arts MP (A2)
Bartels RHMA (A3)
Jacobs WCH (A1)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Neurosurgery, Radboud University, Nijmegen, The Netherlands

Abstract

Background: Despite an increasing implantation rate of interspinous process distraction devices (IPD) devices in the treatment of intermittent neurogenic claudication (INC), definitive evidence on the clinical effectiveness of implants is lacking. The main objective of this review was to perform a meta-analysis of all systematic reviews, randomized clinical trials and prospective cohort series to quantify the effectiveness of IPDs and to evaluate the potential side-effects.

Methods: Data from all studies prospectively describing clinical results based on validated outcome scales and reporting complications of treatment of patients with INC with IPD placement. We searched MEDLINE, EMBASE, Web of Science, Cochrane (CENTRAL), CINAHL, Academic Search Premier, Science Direct up to July 2010. Studies describing patients with INC caused by lumbar stenosis, reporting complication rate and reporting based on validated outcome scores, were eligible. Studies with only instrumented IPD results were excluded.

Results: Eleven studies eligible studies were identified. Two independently RCTs and eight prospective cohorts were available. In total 563 patients were treated with IPDs. All studies showed improvement in validated outcome scores after six weeks and one year. Pooled data based on the Zurich Claudication Questionnaire of the RCTs were more in favor of IPD treatment compared with conservative treatment (pooled estimate 23.2, SD 18.5-27.8). Statistical heterogeneity after pooled data was low (I-squared 0.0, p=0.930). Overall complication rate was 7%.

Conclusions: As the evidence is relatively low and the costs are high, more thorough (cost-) effectiveness studies should be performed before worldwide implementation is introduced.

Introduction

Intermittent neurogenic claudication (INC) is a complex of symptoms, the most important being leg pain and numbness (frequently in both legs) with possibly associated back pain[1-4]. The symptoms can be diminished by flexion of the lumbar spine[5-8]. Lumbar spinal arthrosis inducing arthrosis of the facet is associated with INC[8;9]. Traditionally, bony decompression of the canal and the lateral recessus seems to be the golden standard in the treatment of INC[3;4]. There is some evidence that bony decompression is a proven superior therapy compared with non-surgical therapy, such as steroid injections or fysiotherapy[10;11]. Less invasive strategies have been developed to minimize the perioperative damage, such as unilateral laminotomy or endoscopic procedures[12]. Although surgery is frequently offered, detailed outcome results are not available and spine surgeons try to develop innovative less invasive surgical approaches to gain better outcome than the results observed in daily practice.

Parallel to these developments, interspinous implants for interspinous process distraction devices (IPD) have been developed to achieve indirect decompression[13;14]. The design of the implants aims at limitation of lumbar extension and increasing the interlaminar space of the affected level[15-19]. Nowadays, the technique is widely used. Kyphon Inc. had a worldwide X-STOP™ net sale, in the first quarter of 2007, of 18.1 million USD. Paradigm Spine Inc. reported in May 2010 a worldwide sale of 13,128 Coflex™ devices in 2009[20]. The existing evidence seems to be poor; almost no comparative studies between conventional surgical decompression and surgery with IPD are done[17;21-24]. Some claim, performing IPD placement in day surgery and with local anesthesia will lower the costs. However, a thorough cost-analysis has never been performed.

The main objective of this systematic review was to evaluate if surgery with IPD is more effective compared with bony decompression in the treatment of patients with INC or at least more effective compared with conservative (e.g. steroid injections) treatment.

Methods

This systematic review was performed according to the Cochrane systematic review methodology, up-dated by Furlan and Van Tulder and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) criteria[25-27].

Search strategies

On July 1st 2010, a search of relevant systematic reviews on IPD in the Cochrane Library and, in addition, observational cohort studies (with and without control group), systematic reviews and randomized clinical trials was conducted in MEDLINE, EMBASE, Web of Science, Cochrane (CENTRAL), CINAHL, Academic Search Premier, Science Direct. Keywords used for the search were: interspinous implant surgery, interspinous implants, interspinous distraction devices, interspinous decompression device, interspinous process decompression, intermittent neurogenic claudication, neurogenic claudication, lumbar stenosis, or spinal stenosis. The full search strategy is available upon request from the corresponding author. References of retrieved articles and relevant overview articles were checked to identify additional studies.

Inclusion criteria

Prospective cohort studies, systematic reviews and/or RCTs written in English were considered eligible for inclusion if they fulfilled all of the following:

1. The study population consists of patients with INC caused by lumbar stenosis.
2. Patients with INC without or with degenerative spondylolisthesis to a maximal grade I. One of the treatments consists of non-instrumented IPD for treating symptoms of INC (excluding pedicle screw fixations combined with IPD).
3. A validated outcome score is used to evaluate the outcome after surgery, the Zurich Claudication Questionnaire or the Modified Roland Disability Questionnaire for Sciatica, Oswestry Disability Index, VAS leg and back pain[28-34;34-39].

Studies, in which subgroups met our inclusion criteria, were included in our results if the results for these subpopulations were reported separately.

Study selection

Two reviewers independently applied the inclusion criteria to select potential relevant studies from the titles and abstracts or if necessary the complete publication of the references retrieved by the literature search. Where necessary, a third reviewer was consulted to resolve a disagreement.

Categorization of the relevant literature

Relevant literature was categorized under three different headers: systematic reviews, RCTs, and prospective cohort studies of high quality. The header '*systematic reviews*' describes all systematic reviews. The header '*RCTs*' contains all published RCTs on the same intervention comparing IPD with decompression or conservative treatment. Additional prognostic cohort studies were included. The header '*observational cohort studies*' contains all prospective cohorts with adequate description of the follow-up period and validated outcome measurements. When, due to lack of evidence, pooling data was not possible a descriptive review would be performed based on RCTs and prospective observational cohort studies.

Methodological quality assessment

Systemetic reviews were validated using the steps defined by Furlan and Van Tulder[27;40]. To identify potential risks of bias of the included RCTs two reviewers independently assessed the methodological quality of each RCT according to the Cochrane quality measurements adapted by Furlan and Van Tulder[25;27;41;42]. Each item was scored as "yes", "no", or "don't' know". High quality was defined as a score of 50% or more on the methodological quality assessment. The Dutch Cochrane Centre Quality Assessment (DCCQA) scale was used for the validation process for observational studies. According to the Dutch Cochrane Centre Quality Assessment scale, a score below six was defined as low methodological quality on the DCCQA scale. A third reviewer could be consulted to solve disagreement between the reviewers.

Data extraction

Independently, data were extracted by two reviewers. Information was collected on the study population, intervention(s) performed, outcome measures and outcome. The follow-up time was categorized into short-term outcome (six weeks after intervention) and long-term outcome (at least one year). Furthermore, complication rate and device failure (a re-intervention or other surgical technique was necessary) were recorded. Despite the often mentioned spinal process fractures, all other causes for surgical re-interventions were also recorded[43-46].

Outcome measurements

There are various classifications to describe neurological and functional outcome of patients with intermittent neurogenic claudication. Articles were filtered on presence

of one of the four mostly used outcome scales. Firstly, articles were included on the Zurich Claudication Questionnaire (ZCQ), also known as the Brigham Spinal Stenosis Questionnaire and Swiss Spinal Stenosis Questionnaire[35;38;39]. The ZCQ scale consists of three subscales: symptom severity, physical function and patient satisfaction. Domain scores ranges from 1 to 5, 1 to 4, and 1 to 4 respectively. Like in the study of Tuli in 2006, we chose threshold scores for each scale based on prior work[35;38;39;47]. In the symptom severity scale and in the physical function scale the minimal clinically important difference (MCID) is 0.5[38;39]. A mean patient satisfaction score of less than 2.5 has been shown previously to represent a satisfied patient[38;39]. Secondly, articles were used on the Modified Roland Disability Questionnaire for Sciatica (MRDQ). The 23-points MRDQ is the most widely used patient-assessed measure of health for low back pain and leg pain[29-34;34;36]. This questionnaire consists of 23 questions with higher scores indicating increased disability[48]. The Visual Analogue Scale (VAS) is one of the most used follow-up measurement tools for back pain and leg pain[49]. This parameter will measure the experienced back and leg pain intensity in the week before visiting the research nurse. Pain will be assessed on a horizontal 100 millimeters scale varying from 0 millimeter, "no pain", to 100 millimeters, "the worst pain imaginable"[49]. This parameter has a minimum clinically important difference (MCID) of 2 points on a scale of 0 to 10[50]. Finally, the Oswestry disability index (ODI), where 0 indicates no disability and 100 indicates worst possible disability, was included for our analysis[51]. This parameter has a minimum clinically important difference (MCID) of 10.0-12.4 points[50;52;53].

Data synthesis

A meta-analysis was performed if two or more RCTs were available with clinical homogeneous patient groups and statistical homogeneous results. When not possible, due to small amount of studies or heterogeneity, a best-evidence synthesis was used. Best evidence synthesis was performed stratified for studies meeting 50% or more opposed to those meeting less than 50% of the quality criteria of the Van Tulder list[27]. The study was only included in the best-evidence synthesis if a comparison was made between the groups (IPD placement versus conservative treatment or IPD placement versus surgical decompressive treatment). When meta-analysis or best evidence synthesis based on RCT is not possible, a data extraction based on observational studies (with or without control group) will be performed. Although a high risk of bias is possible, if possible we performed a data extraction from observational studies based on the "best-of-the-rest" principle.

Results

Study selection

The search revealed 253 references. 222 articles were excluded on the basis of the abstract, title and keywords. 20 articles were excluded after reading the complete articles because of the following reasons: the reports did not consist original patient data (4),[17;23;54;55] articles were not written in English (2),[56;57] there were no outcome results given (9),[24;58-65] studies with a retrospective study design(5).[43;66-69] As a result, only three RCTs and eight prospective cohorts were included for methodological quality assessment in this review (Figure 1 – Flowchart)[17;21-23;70-76].

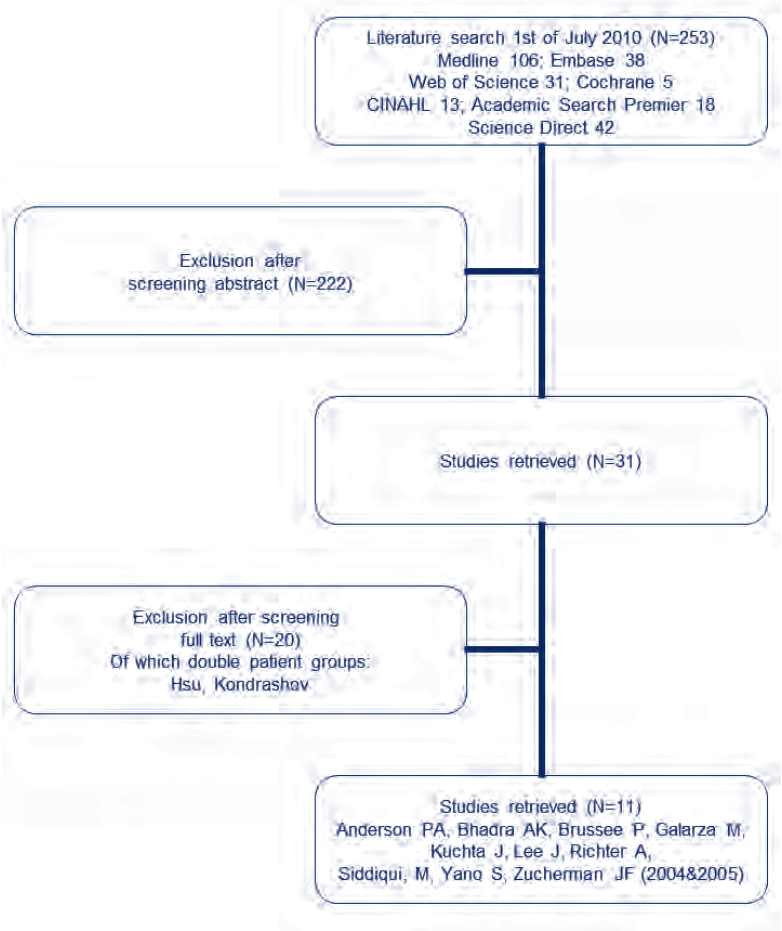


Figure 1. Flowchart

Description of study characteristics

No systematic reviews could be found. Three reports of randomized clinical trials (RCTs) and eight prospective cohorts were found. Three reports described two RCTs comparing non-operative treatment to treatment with IPD; one observational cohort described IPD treatment versus non-operative treatment after bony decompression in both groups; seven cohorts described treatment with IPD only. Two RCTs described the results of the same patient sample. The first study published follow-up data after one year and the last published study after two years, both are shown in Table 1[17;23].

Methodological quality assessment

The methodological quality of the studies is summarized in Table 1 and 2. Two reports of one RCT (of the three RCTs) had a methodological quality score of 5 (low quality) and one RCT had a quality of 6 (a high quality study) according to the Furlan and Van Tulder criteria[25;27]. Only one observational study had a methodological quality of 6 out of 8 (reflecting high quality),[76] thus the remaining seven observational studies are of low methodological quality and with high risk of bias[22;70-75].

Table 1. RCT validation according to Van Tulder validation scale

	Adequate randomization	Allocation concealment	Baseline comparability	Blinding of patient	Blinding of care provider	Blinding outcome assessor	Cointerventions were avoided	Acceptable compliance between groups	Drop-out rate is described	Similar timing of outcome assessment	Intention-to-treat analysis	Overall quality (max 11)
Anderson PA	no	no	yes	no	no	don't know	yes	yes	yes	yes	yes	6
Zucherman JF 2004	no	no	yes	no	no	don't know	yes	yes	yes	yes	don't know	5
Zucherman JF 2005	no	no	yes	no	no	don't know	yes	yes	yes	yes	don't know	5

Data extraction

In tables 3 and 4, relevant data on the selected studies is shown with the baseline and postoperative follow-up scores at six weeks and 1 year. Two RCTs with different patients samples (the two RCTs of Zucherman were conducted on the same patient sample), Anderson et al. and Zucherman et al., could be used for best evidence synthesis[17;21]. Both RCTs compared conservative treatment with IPD placement (Figure 2 – Meta-analysis). Both studies measured follow-up data on the ZCQ. In the study by Zucherman et al, however, overall success rates and standard deviation (SD) values were not shown. A calculation was made, based on the ZCQ values of symptom severity and physical function ZCQ. SD values were calculated estimated from the SD values of Anderson et al. Both studies report 2% complication rate and on top of that 2% device failure (Anderson et al) and 6% device failure (Zucherman). Both studies measured follow-up data on the ZCQ. In the study by Zucherman et al, however, overall success rates and standard deviation (SD) values were not shown.

Table 2. Prospective cohort validation according to Dutch Cochrane validation scale

		Patients defined	Absence of selection bias	Treatment defined	Appropriate outcome measurements	Blinded outcome	Sufficient follow-up time	No selective loss to follow-up	Groups comparable confounding factors	Overall quality (max 8)
1	Bhadra AK	+	+	+	+	-	+	-	-	5
2	Brussee P	+	-	+	+	-	-	+	-	4
3	Galarza M	-	-	+	+	-	+	-	-	3
4	Kuchta J	+	+	+	+	-	+	-	-	5
5	Lee J	+	-	+	+	-	-	+	-	4
6	Richter	+	-	+	+	-	+	-	-	4
7	Siddiqui M	+	-	+	+	-	+	-	-	4
8	Yano S	+	+	+	+	-	+	+	-	6

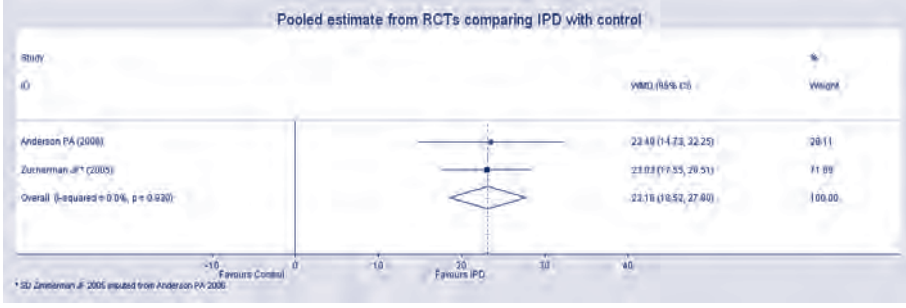
max – maximal points available

Table 3. RCTs – Outcome measurements and complications

RCT	sample size	control sample size	FU	ZCQ baseline IPD (SD)	ZCQ baseline control (SD)	ZCQ short-term IPD	ZCQ short-term control	ZCQ 1year IPD (SD)	ZCQ 1 year control (SD)
Anderson	42	33	99/92	50.4 (±2.0)	51.3 (±2.4)	*46%	*27%	*23.05 (±3.14)	*47.40 (±3.18)
Zucherman	100	91	88/68	SS 3.14	SS 3.12	47%	10%	45,4%	7,4%
				PF 2.48	PF 2.49 overall success	50% *52%	10% *10%	44,3% *48,4%	-0,4% *4,9%

FU – Complete follow-up; ZCQ – Zurich Claudication Questionnaire; SS – ZCQ symptom severity domain; PF – ZCQ physical function domain; * Values of ZCQ overall success domain

Figure 2. Meta-analysis



IPD – interspinous process decompression; SD – standard deviation; WMD – weighted mean difference.

A calculation was made, based on the ZCQ values of symptom severity and physical function ZCQ. SD values were calculated estimated from the SD values of Anderson et al. Both studies report 2% complication rate and on top of that 2% device failure (Anderson et al) and 6% device failure (Zucherman). Both studies favored treatment with IPD placement, pooled ZCQ improvement by 23.2 (SD 18.5-27.8). Statistical heterogeneity after pooled data was low (I-squared 0.0, p=0.930). According to the statistical heterogeneity, baseline criteria in both studies showed a good clinical homogeneity. Richter et al. compared two surgical decompression cohorts: one group with surgical decompression and no IPD placement, one group with surgical decompression with IPD placement[22]. Both groups showed clinical improvement in the ODI, MRDQ and VAS. At six weeks and at one year follow-up there were no statistical significantly differences between both groups. The remaining seven prospective cohort studies showed improvement from baseline after treatment

with IPD[70-76]. However, these groups did not compare other treatment modalities (such as conservative treatment) with IPD follow-up results. Due to the use of multiple follow-up scales, pooling of data was not possible (Table 4 – prospective cohort studies).

In our search of literature, 563 patients underwent implantation with IPD. Complication rates and device failure rates were available from 513 patients (Table 3 – RCT and 4 – prospective cohorts). A total of 31 devices failed (6%) and had to be replaced or were re-operated with bony decompression and stabilization. Six (1%) other complications were also reported (infections and postoperative leakages).

Discussion

The literature has been systematically reviewed to evaluate the outcome for patients with intermittent neurogenic claudication treated with IPD versus bony decompression or conservative non-surgical treatment. To our knowledge, this is the first systematic review and meta-analysis on this subject. After a literature search, two independent RCTs and eight prospective cohorts, one with a control group, were eligible for validation and data-extraction. The methodological quality of the RCTs were 5 (Zucherman) and 6 (Anderson)[17;21]. The methodological quality of the remaining prospective cohort studies was relatively low (only one reached 6 out of 8)[22;70-76]. In total 563 patients were treated with IPD. All studies showed improvement in validated outcome scores after six weeks and one year. Pooled data of the RCTs were more in favor of IPD treatment compared with conservative treatment.

The review of the literature showed that very little is known about treatment with IPD. Only one comparative study with good methodological quality fulfilling our selection criteria was found[21]. Different indications are used for these devices, such as described by Richter who used an IPD in combination with surgical decompression[22]. Some studies show beneficial effect of surgical technique compared to conservative treatment for patients with degenerative lumbar spinal stenosis and neurogenic intermittent claudication[10;11]. More centers, however, perform complex techniques rather than only a decompression technique. Between 2002 and 2007, complex fusion procedures showed a 15-fold increase in

the USA. Furthermore, the overall procedure rate slightly decreased with 1.4%[77-79]. Coflex worldwide implants increased from 1,717 in 2005 to 13,128 in 2009. Even without evidence of implantation of an IPD as a treatment strategy for INC, some centers use it in a combination with other techniques[22]. Despite the fact that no arguments exist in the literature about the effectiveness of treatment with IPD versus bony decompression, many centers throughout the world use IPD for the treatment of INC.

Overall complication and failure rate of (7%, including 6% reoperations rate after device failure) tended to be relatively low compared to the complication rate of standard bony decompression. For example, Weinstein and Malmivaara reported a complication rate of 17-24% in the standard bony decompression operation cohorts[10;11]. The most frequently reported complications in these series are dural tears and wrong level surgery. Due to the use of standard X-rays in the operation theater with IPD treatment, wrong level surgery in interspinous decompression surgery is rare. Most techniques of interspinous decompression are indirect and with some distance from the dura, therefore causing a dural tear is difficult by regular surgical methods. Despite the large numbers of case reports on complications after IPD treatment, complication rates tends to be low[43;44;80;81]. This, however, might be induced by selection bias of published studies. Despite the relatively low complication rate, device failure rate needing reoperation is high (6%). This number can be higher because of the publication bias, but also due to the lack of long-term follow up.
This conclusion is difficult to confirm due to the fact that no comparative studies are done on this subject. Combined with the 6% device failure rate complication rate, the IPD complication rate is 7%. In the literature, implantation surgery is associated with complication rate of 8% (2-6% failure rate)[79]. The complication rate would be possible higher when complications would be monitored thirty days after discharge. Not all studies included in our review reported complication rate thirty days after hospital stay. Prospective reporting of complication should be made standard in future trials.

Table 4a. Prospective cohorts comparing conservative treatment with surgery with IPD – Outcome measurements and complications

	sample size	Device	FU	Age	Follow-up	VAS baseline	VAS follow-up	ODI baseline	ODI follow-up 1	ODI follow-up 2	ZCQ baseline SS/PF	ZCQ follow-up 1 SS/PF/PS	ZCQ follow-up 2 SS/PF/PS	complications	device faillure
Bhadra AK	45	X-Stop	100%	61.5		69	28	42	16					4%	2%
Brussee P	65	X-Stop	95%	64.4										0%	9%
Galarza M	40	Aperius	100%	72.7	1 year	71	22					90%			
Kuchta J	175	X-Stop	100%	69.4	6 weeks/ 1 year	61	39	33	23	15				0%	5%
Lee J	10	X-Stop	100%	71.0	9-18 months							50%		0%	
Siddiqui M	37	X-Stop	65%	71.5	3 months/ 1 year			48	35	37	3.37/ 2.45	2.42/2.05/ 1.90	2.83/2.19/ 2.12	0%	11%
Yano S	19	Ceramic	95%	70.1	mean 37.4 months	69	30				2.94/ 2.51	1.92/1.73		0%	11%

Age – patients' age in years; FU – Complete follow-up; Follow-up – Follow-up periods; ZCQ – Zurich Claudication Questionnaire; ODI – Oswestry Disability Index; VAS – Visual Analogue Scale; SS – ZCQ symptom severity domain; PF – ZCQ physical function domain; PS – ZCQ patient satisfaction domain;

The most important limitation of this review concerned the methodological weaknesses and selection biases of the included studies: the vast majority was observational, without independent outcome assessment, and without complications well defined. Additionally, we combined two different RCTs for our meta-analysis[17;21]. Both studies did not mention a thorough power or sample size design, resulting in a 191 patients in one RCT and 75 in the other. Furthermore, only one study was of relatively high methodological quality. Therefore possible information bias could be introduced. Furthermore we excluded 242 studies, introducing selection bias. Due to the retrospective design of some of these studies, possible interesting patient data had to be excluded. Studies that were published in abstract or poster format only were excluded. The present study was aimed at identifying published peer-reviewed literature, so that influence of publication bias cannot be ruled out. Due to the small number of studies, possible publication bias (using e.g. funnel plot) could not adequately be assessed. Due to the anticipated low number of RCTs, prospective studies were also included. Most of these studies were of low methodological quality (Table 2 – Validation). Due to the inclusion of studies of low methodological quality, information bias is easily introduced. Furthermore, methodological quality assessment does not take into account the author's disclosure. For example, two studies in our review stated that one of the authors is a consultant and, in one article, stockholder of the company manufacturing the IPD device they were using for their study[21;70]. The remaining studies did not mention any conflict of interest or disclosure. Seven studies did not even describe the possible conflicts of interest. Assessing possible conflict of interest is not incorporated in both validation scales[25;27]. Standard adjusting both scales based on possible conflict of interest is advisable.

This review of the literature shows that surgical decompression with interspinous process devices is superior to conservative non-surgical treatment in patients with lumbar degenerative spinal stenosis with INC. However, the level of evidence for this conclusion is debatable due to the low quality of some of the included studies. Furthermore, no data is presently available comparing interspinous process decompression with standard bony decompression. We suggest that more studies will be done on this subject comparing the surgical treatment with IPD versus bony decompression. Despite the fact that we could give a Grade A recommendation, according to the Oxford-Centre for Evidence Based Medicine, we suggest that further studies have to be performed before a thorough recommendation can be

given regarding the treatment of INC with IPDs[80]. These studies should also include analysis on complication rate and device failure rate. As the evidence is relatively low and the costs are high, more thorough cost-effectiveness studies should be performed before worldwide implementation is introduced. Because the golden standard for surgical decompression seems to be absent, patients with lumbar spinal stenosis should be guarded against instrumented surgery or the use of IPD on the basis of the current evidence.

Reference List

1. Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. Spine (Phila Pa 1976) 1991;16:615-619.
2. Amundsen T, Weber H, Lilleas F, Nordal HJ, Abdelnoor M, Magnaes B. Lumbar spinal stenosis. Clinical and radiologic features. Spine (Phila Pa 1976) 1995;20:1178-1186.
3. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. Spine (Phila Pa 1976) 1999;24:1820-1832.
4. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: updated Cochrane Review. Spine (Phila Pa 1976) 2005;30:2312-2320.
5. EVANS JG. Neurogenic Intermittent Claudication. Br Med J 1964;2:985-987.
6. Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. Ned Tijdschr Geneeskd 1950;94:2415-2433.
7. Verbiest H. [Further reports on primary stenosis of the lumbar spinal canal in adults.]. Ned Tijdschr Geneeskd 1951;95:1965-1970.
8. Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. J Bone Joint Surg Br 1954;36-B:230-237.
9. Arnoldi CC, Brodsky AE, Cauchoix J et al. Lumbar spinal stenosis and nerve root entrapment syndromes. Definition and classification. Clin Orthop Relat Res 1976;4-5.
10. Malmivaara A, Slatis P, Heliovaara M et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 2007;32:1-8.
11. Weinstein JN, Tosteson TD, Lurie JD et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008;358:794-810.
12. Thome C, Zevgaridis D, Leheta O et al. Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. J Neurosurg Spine 2005;3:129-141.
13. Mariottini A, Pieri S, Giachi S et al. Preliminary results of a soft novel lumbar intervertebral prothesis (DIAM) in the degenerative spinal pathology. Acta Neurochir Suppl 2005;92:129-131.
14. Tsai KJ, Murakami H, Lowery GL, Hutton WC. A biomechanical evaluation of an interspinous device (Coflex) used to stabilize the lumbar spine. J Surg Orthop Adv 2006;15:167-172.
15. Simons P, Werner D, Krause P, Mark P, Godde G, Reinhard A. Treatment of neurogenic intermittent claudication (NIC), secondary to lumbar spinal stenosis (LSS), with the COFLEX® Interspinous Process Decompression (IPD) system. Early ZCQ outcome results from the European registry 2006.
16. Wilke HJ, Drumm J, Haussler K, Mack C, Steudel WI, Kettler A. Biomechanical effect of different lumbar interspinous implants on flexibility and intradiscal pressure. Eur Spine J 2008;17:1049-1056.
17. Zucherman JF, Hsu KY, Hartjen CA et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. Eur Spine J 2004;13:22-31.
18. Siddiqui M, Karadimas E, Nicol M, Smith FW, Wardlaw D. Influence of X Stop on neural foramina and spinal canal area in spinal stenosis. Spine (Phila Pa 1976) 2006;31:2958-2962.
19. Siddiqui M, Karadimas E, Nicol M, Smith FW, Wardlaw D. Effects of X-STOP device on sagittal lumbar spine kinematics in spinal stenosis. J Spinal Disord Tech 2006;19:328-333.
20. Sale numbers directly obtained from Kyphon Inc and Paradigm Spine Inc
21. Anderson PA, Tribus CB, Kitchel SH. Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. J Neurosurg Spine 2006;4:463-471.
22. Richter A, Schutz C, Hauck M, Halm H. Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. Eur Spine J 2009.
23. Zucherman JF, Hsu KY, Hartjen CA et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. Spine (Phila Pa 1976) 2005;30:1351-1358.

24. Stromqvist B, Johnsson R, Sahlstrand T, Tullberg T. Comparison of patients with neurogenic intermittent claudication treated with the X-Stop interspinous process decompression system vs. decompressive surgery--early follow-up results. The Spine Journal 2007;5:S168.
25. Furlan JC, Singh J, Hsieh J, Fehlings MG. Methodology of Systematic Reviews and Recommendations. J Neurotrauma 2010.
26. Stroup DF, Berlin JA, Morton SC et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008-2012.
27. van Tulder M., Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine (Phila Pa 1976) 2003;28:1290-1299.
28. Deyo RA. Comparative validity of the sickness impact profile and shorter scales for functional assessment in low-back pain. Spine (Phila Pa 1976) 1986;11:951-954.
29. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. Spine (Phila Pa 1976) 1986;11:28-30.
30. Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. Med Care 1995;33:AS286-AS291.
31. Hutchinson PJ, Laing RJ, Waran V, Hutchinson E, Hollingworth W. Assessing outcome in lumbar disc surgery using patient completed measures. Br J Neurosurg 2000;14:195-199.
32. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006;332:1430-1434.
33. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989;27:S217-S232.
34. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine (Phila Pa 1976) 1995;20:1899-1908.
35. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. Spine (Phila Pa 1976) 2002;27:84-91.
36. Roland M, Morris R. A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care. Spine (Phila Pa 1976) 1983;8:145-150.
37. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976) 1983;8:141-144.
38. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369-1378.
39. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine (Phila Pa 1976) 1996;21:796-803.
40. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44:1271-1278.
41. Verhagen AP, de Vet HC, de Bie RA, Boers M, van den Brandt PA. The art of quality assessment of RCTs included in systematic reviews. J Clin Epidemiol 2001;54:651-654.
42. Furlan AD, Pennick V, Bombardier C, van TM. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976) 2009;34:1929-1941.
43. Barbagallo GM, Olindo G, Corbino L, Albanese V. Analysis of complications in patients treated with the X-Stop Interspinous Process Decompression System: proposal for a novel anatomic scoring system for patient selection and review of the literature. Neurosurgery 2009;65:111-119.
44. Bowers C, Amini A, Dailey AT, Schmidt MH. Dynamic interspinous process stabilization: review of complications associated with the X-Stop device. Neurosurgical Focus 2010;28.
45. La Marca F. Analysis of complications in patients treated with the X-Stop interspinous process decompression system: Proposal for a novel anatomic scoring system for patient selection and review of the literature - Commentary. Neurosurgery 2009;65:120.
46. Resnick DK. Analysis of complications in patients treated with the X-Stop interspinous process decompression system: Proposal for a novel anatomic scoring system for patient selection and review of the literature - Commentary. Neurosurgery 2009;65:119.

47. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. Spine (Phila Pa 1976) 2006;31:1276-1280.
48. Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. Phys Ther 1996;76:268-275.
49. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983;16:87-101.
50. Phelan EA, Deyo RA, Cherkin DC et al. Helping patients decide about back surgery: a randomized trial of an interactive video program. Spine (Phila Pa 1976) 2001;26:206-211.
51. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66:271-273.
52. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007;7:541-546.
53. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. Spine J 2008;8:968-974.
54. Hsu KY, Zucherman JF, Hartjen CA et al. Quality of life of lumbar stenosis-treated patients in whom the X STOP interspinous device was implanted. J Neurosurg Spine 2006;5:500-507.
55. Kondrashov DG, Hannibal M, Hsu KY, Zucherman JF. Interspinous process decompression with the X-STOP device for lumbar spinal stenosis: a 4-year follow-up study. J Spinal Disord Tech 2006;19:323-327.
56. Collignon F, Fransen P. [Treatment of symptomatic degenerative lumbar spinal stenosis by a percutaneous stand-alone lumbar interspinous implant. Preliminary experience with the Aperius device]. Neurochirurgie 2010;56:3-7.
57. Hrabalek L, Machac J, Vaverka M. [The DIAM spinal stabilisation system to treat degenerative disease of the lumbosacral spine]. Acta Chir Orthop Traumatol Cech 2009;76:417-423.
58. Hsu K, Zucherman J, Hartjen C et al. 117. Interspinous process decompression (IPD) significantly improves symptom severity in neurogenic intermittent claudication patients. The Spine Journal 2004;4:S60.
59. Implicito D, Martin M, Ozuna R. 6:12174. Outcomes of neurogenic intermittent claudication patients treated with interspinous decompression as a function of number of levels treated. The Spine Journal 2007;5:S90.
60. Kondrashov D, Hannibal M, Hsu K, Zucherman J. The X STOP interspinous process decompression versus laminectomy for treatment of lumbar spinal stenosis: Economic analysis. Neurosurgery 2006;59:458.
61. Skidmore G, Ackerman S, Bergin C et al. Cost-effectiveness of Interspinous Process Decompression for Lumbar Spinal Stenosis: A Comparison with Conservative Care and Laminectomy. The Spine Journal 2009;7:95S.
62. Zucherman J, Hsu K, Hartjen C et al. A prospective randomized multicenter study of lumbar neurogenic intermittent claudication patients treated conservatively with a laminectomy or with interspinous process decompression. The Spine Journal 2004;4:S33-S34.
63. Zucherman J, Hsu K, Wahlig J, Hartjen C, Shabe P, Tillman J. Five Year Outcomes in Patients Treated with the X-STOP Interspinous Process Device for Neurogenic Intermittent Claudication Due to Lumbar Spinal Stenosis. The Spine Journal 2009;8:153S.
64. Eichholz KM, Fessler RG. Is the X STOP interspinous implant a safe and effective treatment for neurogenic intermittent claudication? Nat Clin Pract Neurol 2006;2:22-23.
65. Weiner BK, Zucherman JF, Hsu KY, Hartjen CA. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: Two-year follow-up results. Spine 2005; 30: 1351-8 [2]. Spine 2005;30:2846-2847.
66. Errico TJ, Kamerlink JR, Quirno M, Samani J, Chomiak RJ. Survivorship of coflex Interlaminar-Interspinous Implant. SAS Journal 2009;3:59-67.
67. Nardi P, Cabezas D, Rea G, Pettorini BL. Aperius PercLID Stand Alone Interspinous System for the Treatment of Degenerative Lumbar Stenosis: Experience on 152 Cases. J Spinal Disord Tech 2010.
68. Senegas J, Vital JM, Pointillart V, Mangione P. Clinical evaluation of a lumbar interspinous dynamic stabilization device (the Wallis system) with a 13-year mean follow-up. Neurosurg Rev 2009;32:335-341.

69. Verhoof OJ, Bron JL, Wapstra FH, van Royen BJ. High failure rate of the interspinous distraction device (X-Stop) for the treatment of lumbar spinal stenosis caused by degenerative spondylolisthesis. Eur Spine J 2008;17:188-192.
70. Bhadra AK, Raman AS, Tucker S, Noordeen HH. Interspinous implant in lumbar spinal stenosis: A prospective cohort. European Journal of Orthopaedic Surgery and Traumatology 2008;18:489-493.
71. Brussee P, Hauth J, Donk RD, Verbeek AL, Bartels RH. Self-rated evaluation of outcome of the implantation of interspinous process distraction (X-Stop) for neurogenic claudication. Eur Spine J 2008;17:200-203.
72. Galarza M, Fabrizi AP, Maina R, Gazzeri R, Martinez-Lage JF. Degenerative lumbar spinal stenosis with neurogenic intermittent claudication and treatment with the Aperius PercLID System: a preliminary report. Neurosurgical Focus 2010;28.
73. Kuchta J, Sobottke R, Eysel P, Simons P. Two-year results of interspinous spacer (X-Stop) implantation in 175 patients with neurologic intermittent claudication due to lumbar spinal stenosis. Eur Spine J 2009;18:823-829.
74. Lee J, Hida K, Seki T, Iwasaki Y, Minoru A. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. J Spinal Disord Tech 2004;17:72-77.
75. Siddiqui M, Smith FW, Wardlaw D. One-year results of X Stop interspinous implant for the treatment of lumbar spinal stenosis. Spine (Phila Pa 1976) 2007;32:1345-1348.
76. Yano S, Hida K, Seki T, Aoyama T, Akino M, Iwasaki Y. A new ceramic interspinous process spacer for lumbar spinal canal stenosis. Neurosurgery 2008;63:ONS108-ONS113.
77. Carragee EJ. The increasing morbidity of elective spinal stenosis surgery: is it necessary? JAMA 2010;303:1309-1310.
78. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. Spine (Phila Pa 1976) 2005;30:1441-1445.
79. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 2010;303:1259-1265.
80. Barbagallo GM, Corbino LA, Olindo G, Foti P, Albanese V, Signorelli F. The "sandwich phenomenon": a rare complication in adjacent, double-level X-stop surgery: report of three cases and review of the literature. Spine (Phila Pa 1976) 2010;35:E96-100.
81. Epstein NE. X-Stop: foot drop. Spine J 2009;9:e6-e9.
82. Phillips B, Ball C, Sackett DL, et al. Levels of evidence and grades of recommendation. Centre for Evidence-Based Medicine: Oxford-Centre for Evidence Based Medicine. 1998.

Chapter 4

The Felix Trial: design and protocol.

Extracted from BMC Muskuloskelet Disord. 2010 May;27(11):100

Moojen WA (A1/A2)
Arts MP (A2)
Brand R (A3)
Koes BW (A4)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
[A4] Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract

Background: Decompressive laminotomy is the standard surgical procedure in the treatment of patients with canal stenosis related intermittent neurogenic claudication. New techniques, such as interspinous process implants, claim a shorter hospital stay, less post-operative pain and equal long-term functional outcome. A comparative (cost-) effectiveness study has not been performed yet. This protocol describes the design of a randomized controlled trial (RCT) on (cost-) effectiveness of the use of interspinous process implants versus conventional decompression surgery in patients with lumbar spinal stenosis.

Methods/Design: Patients (age 40-85) presenting with intermittent neurogenic claudication due to lumbar spinal stenosis lasting more than 3 months refractory to conservative treatment, are included. Randomization into interspinous implant surgery versus bony decompression surgery will take place in the operating room after induction of anesthesia. The primary outcome measure is the functional assessment of the patient measured by the Zurich Claudication Questionnaire (ZCQ), at 8 weeks and 1 year after surgery. Other outcome parameters include perceived recovery, leg and back pain, incidence of re-operations, complications, quality of life, medical consumption, absenteeism and costs. The study is a randomized multi-institutional trial, in which two surgical techniques are compared in a parallel group design. Patients and research nurses are kept blinded of the allocated treatment during the follow-up period of 1 year.

Discussion: Currently decompressive laminotomy is the golden standard in the surgical treatment of lumbar spinal stenosis. Whether surgery with interspinous implants is a reasonable alternative can be determined by this trial.

Dutch Trial register number: NTR1307

Background

Intermittent Neurogenic Claudication (INC) is a complex of symptoms first described by Van Gelderen in 1948 and in 1950 by the Dutch neurosurgeon Verbiest, therefore formerly known as the Verbiest syndrome [1-4]. The characteristic symptom is described as leg pain (frequently in both legs) which can be exacerbated with prolonged walking or lumbar extension. Others, like Evans, describe a cramp, tightness or discomfort of the legs after walking which diminish after a short period of sitting or bending forward [1]. Apart from the leg pain, associated low back pain may occur [5].

Since the description of neurogenic claudication by Verbiest, explanation of the symptoms has been disputed. Verbiest stated in 1954: "*In the writer's humble opinion the ligamentum flavum is most unlikely to contact any spinal root unless this root is distorted from its regular path*"[4]. Evans showed in 1964 a cerebral spinal fluid stop at the low lumbar levels narrowing of the canal by degenerative facet arthrosis resulting in nerve root compression. INC is often seen in patients with lumbar degenerative spinal stenosis [4]. Due to this arthrosis of the facet joints, lumbar nerve root compression will develop. Arnoldi described multiple types of lumbar spinal stenosis. His article published in 1975 was actually a summarization of a symposium on this subject [4,6]. Presently, his classification is still widely used. Like in any acquired disease, INC is usually seen in the elderly [1].

The best treatment of NIC due to lumbal stenosis remains controversial [5,7]. Nonoperative therapy like epidural steroid injections, nonsteroidal anti-inflammatory medication, analgesics, physical therapy, and spinal manipulation, is frequently performed [8]. A 2005 Cochrane review found that the paucity and heterogeneity of evidence limited conclusions regarding surgical efficacy for spinal stenosis [7,9-12]. Indeed, Weinstein et al published in his article the results of a randomized cohort study with relatively poor results in the non-operative group [13-15]. Despite the high level of crossovers in their study, the treatment effect was favouring surgery on the SF-36 scale for bodily pain. Also Malmivaara et al showed a better recovery after surgery versus conservative treatment with a difference of improvement of 11.3 on the ODI disability scale [16]. Furthermore Turner et al published in their attempted meta-analysis a success rate (good to fair outcome) of 64% after surgical bony decompression in patients with INC [17].

Thomé et al prospectively compared the most typically used techniques: laminectomy, unilateral laminotomy and bilateral laminotomy [18]. In the series of Thomé et al, bilateral laminotomy achieved an 80% success rate. It was slightly better compared to laminectomy, which had 70% success rate [18]. Many authors claim that bony decompressive surgery might facilitate spinal fusion in the future [19,20]. Furthermore, local trauma in these surgical strategies should not be underestimated [21]. The above described operations are usually performed under general or local anaesthesia and 2 to 7 days hospitalization maybe required, followed by an 8-weeks recovery period. Furthermore, the clinical outcome seems disappointing, since 35% of the patients documented bad outcome [13-17].

Minimally invasive surgery has gained popularity in recent years, resulting in the development of interspinous implants in the 1980s [22]. One of these models, the Wallis device, was made with a band around the spinous processes. Later in 2003 X-stop, in 2005 Diam, in 2006 Coflex, and afterwards various other kinds of forms were developed to stabilize or distract the interspinous distance [23-33]. These implants are all placed between spinous processes, which will lead to distraction of the interspace with consequent indirect decompression of the nerve roots. Presently, most publications refer to X-stop implants [8,23-26,29-31,34-36]. It is claimed that this indirect decompression will reduce the pressure on the nerves leading to a return to a neutral or slightly tightened position of the vertebral column. Nevertheless, this is a far smaller operation and gives perhaps less destruction to the bony elements of the vertebral column. Therefore, IPD is believed to have better short-term recovery and similar long-term (cost-) effectiveness [8,34,36,37]. Outcomes were reported to be quite favourable in selected series of poor methodological quality. The first randomized multicentre study on interspinous devices compared X-stop with non-surgical treatment [36]. After 2 years, the IPD group shows both clinically and statistically significant improved results in comparison with the conservative treated group [8,36]. However, this trial only compared IPD with conservative treatment. Good evidence on IPD versus other surgical treatment is not yet available. Verhoof et al reported in 2008 a high failure rate in IPD (X-stop), with an average slip on the radiographs of 19.6%, and a high surgical re-intervention rate (seven out of the 12) [35]. Strömqvist reported 13 re-operations in a group of 50 patients [38]. Park et al published one of the few studies with the Coflex implant [39]. However they only placed a Coflex implant after bony decompression [40]. Furthermore long term results, despite from the small retrospective series (twenty patients) of Kondrashov et al, are not yet available [34].

The golden standard in surgical treatment for lumbar spinal stenosis is bony decompression to which all new techniques should be compared. The purpose of our study is to asses whether IPD-surgery is more (cost) effective compared with surgical decompression in patients with INC due to lumbar stenosis. It is hypothesized that IPD gives particularly a favourable short term effect, necessitating a short term evaluation.

Methods/design

An observer and patient blinded randomized (cost-)effectiveness trial in the treatment of lumbar spinal stenosis is presented. In this trial two surgical techniques are compared in a parallel group design. The primary outcome measurement is the Zurich Claudication Questionnaire. The follow-up period will last 1 year. In order to collect enough patients, a multi-center design is necessary. The study protocol was approved in all participating hospitals (see table 1: list of hospitals).

Our primary question is whether IPD-surgery is more (cost-)effective compared with surgical decompression after 8 weeks in people with intermittent neurogenic claudication due to lumbar stenosis. The main advantage of IPD might be a faster recovery after surgery, but after long term follow-up it is unknown if this treatment effect will remain. Therefore, in addition, long-term follow-up (one year) will be compared with short-term follow-up.

Table 1. list of hospitals participating in the Felix Trial

- Leiden University Medical Centre
- Medical Centre Spaarne, Hoofddorp
- Medical Centre Rijnland, Leiderdorp
- Medical Centre Diaconessenhuis, Leiden
- Medical Centre Haaglanden, The Hague
- Medical Centre Bronovo, The Hague
- Medical Centre Groene Hart, Gouda
- Medical Centre Reinier de Graaf, Delft
- Medical Centre Vlietland, Schiedam
- Medical Centre Canisius Wilhelmina, Nijmegen
- Medical Centre Haga, The Hague
- Medical Centre Isala, Zwolle
- Medical Centre Alkmaar
- Medical Centre Tergooier, Hilversum
- University Medical Center Utrecht

Patients

All patients between 40 and 85 years with at least three months of INC due to spinal canal stenosis are eligible for this study. Imaging studies (MRI) must confirm a narrowed lumbar spinal canal, nerve root canal or intervertebral foramen at one or two levels. Patients have received at least three months of conservative therapy. Lumbar discectomie is not possible during IPD surgery. Therefore, patients should be excluded when a surgical relevant herniated disc is present. Additional inclusion and exclusion criteria are listed in table 2. Patients are referred by a neurologist with MRI and conventional imaging of the lumbar spine. During the first visit to the neurosurgical outpatient clinic, the patient's history and a standard neurological examination will be documented. Conform our selection criteria, the neurosurgeon decides whether a patient is eligible for the Felix (Foraminal Enlargement Lumbar Interspinous distraXion) trial and informs the patient about both surgical techniques.

Table 2. inclusion and exclusion criteria

Inclusion criteria:
- signed informed consent
- 40 to 85 years
- has INC, as noted by leg/buttock/groin pain with or without back pain
- at least three months conservative treatment
- has a regular indication for surgical intervention INC
- has a narrowed lumbar spinal canal, nerve root canal or intervertebral foramen at one or two levels confirmed by MRI
- is physically and mentally willing and able to comply with, or has caregiver why is willing and able to comply with, the post-operative evaluations
Exclusion criteria:
- has a cauda equina syndrome
- has a herniated disc at the same level, necessitating lumbar discectomy
- has Paget's disease, severe osteoporosis or metastasis to the vertebrae
- has significant scoliosis (Cobb angle >25 degrees)
- has had previous surgery of the same lumbar level
- has degenerative spondylolisthesis > grade 1 (scale 1 to 4) at the affected level
- has significant instability of the lumbar spine
- has severe co morbid conditions
- has a fused segment at the indicated level

The study, with both treatment options, will be explained to patients and, in case of a positive reaction, appointments are made with research nurses. Because the patient needs sometime to consider participation, the first visit to the research nurse is planned after at least 2 days. After informed consent, the questionnaires, outcome measures and baseline variables are recorded.

Ethical considerations

In concordance to the declaration of Helsinki, the study has been reviewed by an independent ethical committee and approved as being ethically constituted. The design of this study is approved by the Leiden Ethical Medical Committee. Every participating center independently needs an approval before they may include patients for this trial. Freely given informed consent will be obtained from a patient before inclusion in this study. This means that a patient has the right to know that he is being asked to take part, and that he does not have to do so unless he chooses. The patient will also be informed that there will be no financial rewards if he or she agrees to participate.

Randomization procedure

Patients will be randomly allocated to either IPD or conventional decompression. Randomization will take place in the operating room within 4 weeks after inclusion by the research nurse. A randomized block design, stratified by hospital and research nurse, is used to ensure equal distribution of both treatments while ensuring by imposing a variable, random block size that the next treatment is not predictable for the surgeon. The randomization was prepared by the study statistician and the principle data manager at the department of Biostatistics. They were not involved in the selection and allocation of patients and prepared coded, sealed envelopes containing the treatment allocation. In the operating room, after induction of anesthesia, the surgeon will open the envelope and the allocated treatment will be performed. Patients, nursery department and research nurses are kept blinded for the allocated treatment during the follow-up period of 1 year. The operation report will be kept separately and will only be available in case of complications or reoperations.

Interventions

After the induced general anesthesia, randomization in group (A) IPD and (B) surgical decompression will be performed. The patient is positioned in knee-elbow position or prone, dependent by the preference of the surgeon. The affected

spinal level is verified fluoroscopically. The participating surgeons have experience in both techniques and performed at least five implant operations and 15 bony decompression operations.

A) IPD

A median lumbar incision is made over the spinous processes, the laminae of the affected level(s) are exposed subperiosteally, and the supraspinous ligament will be incised. The interspinous ligament of the affected level is removed. No decompression will be performed and the ligamentum flavum will remain intact. A Coflex™ device is placed in the created space between the spinous process with insertion of instrumentation. The wound will be closed in layers with a suction drain. The titanium Coflex™ implant that fits between the spinous processes. The Coflex™ is available in 5 sizes: 8mm, 10mm, 12mm, 14mm and 16mm. The size refers to the minor diameter of the oval spacer assembly of the Coflex™. Patients will be operated with loupe magnification or microscope depending surgeon's preference. When an IPD fails, a standard laminotomy will be performed.

B) Surgical decompression

Similarly as in group A, a median lumbar incision will be made and the paravertebral muscles will be dissected subperiosteally and retracted bilaterally. Decompression will be applied via partial resection of the affected laminae and no complete laminectomy will be performed. The lateral recess will be opened bilaterally and medial facetectomy will be performed in order to maintain stability of the segments. The wound will be closed in layers with a suction drain. Like in the IPD group, patients will be operated with loupe magnification or microscope depending on the surgeon's preference.

The patient will be allowed to leave the bed and walk without aid on the day of surgery. If the patient regains his/her physical function, the patient will be discharged. In both studies, patients and their guided physiotherapists are stimulated to resume home activities and work as soon as possible. The latter are blinded for the allocated treatment arm as well.

Baseline data

The baseline questionnaire assesses demographics, hobbies, sports, work status, smoking status, low back pain history, family history of INC, co-morbidity, weight and length. The patient's satisfaction at work will be registered. The patient's and the

surgeon's treatment preference for IPD or decompression surgery will be assessed on a 5-point scale ranging from "strong preference for IPD" to "strong preference for decompression surgery".

Outcome assessment

The validated outcome parameters described below will be used in this study and assessed by means of questionnaires. Follow-up examinations by the research nurse will take place at 2, 4, 8 weeks, 3, 6, 12, 24 and 60 months after randomization (see table 3: flowchart). Patients will be neurologically examined (at 8 weeks, 6, 12, 24 and 60 months) and the main questionnaires will be filled out at home with a request to complete and return them. The outpatient control by the neurosurgeon will be at 8 weeks and more often if necessary (see table 3: flowchart)

Table 3. Flowchart

Obtained patients' information	V1	V2	V3	V4	V5	V6a	V7	V8	V9	V10
In-patient		x								
Out-patient	x				x	x	x	x	x	x
Demography & diagnosis	x									
Basic physical examination	x									
Neurological examination	x				x		x	x	x	x
Provide study information	x									
Obtain informed consent	x									
X-ray		x						x		
Randomisation		x								
ZCQ	x		x	x	x	x	x	x	x	x
MRDQ	x		x		x	x	x	x	x	x
Shuttle Walking Test	x				x	x	x	x	x	x
SF-36	x			x	x	x	x	x	x	x
McGill Pain Questionnaire	x				x	x	x	x	x	x
VAS for legs and back	x		x	x	x	x	x	x	x	x
Perceived Recovery					x	x	x	x	x	x
Patient Global Impression of change					x	x	x	x	x	x
EuroQol & VAS Quality of Life	x		x	x	x	x	x	x	x	x
Patient diary					x	x	x	x	x	x
Review MRI	x									
Complications		x	x	x	x		x	x	x	x
Re-operation					x	x	x	x	x	x

A, questionnaires will be sent per mail with request to complete and return them; V1, visit 1 - Intake; V2, Visit 2 - surgery; V3, Visit 3 - Follow-up 2 weeks; V4, Visit 4 - Follow-up 4 weeks; V5, Visit 5 - Follow-up 8 weeks; V6, Visit 6 Follow-up 3 months; V7, Visit 7 – Follow-up 6 months; V8, Visit 8 - Follow-up 12 months; V9, Visit 9 - Follow-up 24 months; V10, Visit 10 - Follow-up 60 months

Primary outcome measurement

The disorder-specific functional score will be the primary outcome measure and can be obtained by completing the ZCQ, also known as the Brigham Spinal Stenosis Questionnaire and Swiss Spinal Stenosis Questionnaire [41-43]. The ZCQ scale consists of 3 subscales: symptom severity, physical function and patient satisfaction. Domain scores ranges from 1 to 5, 1 to 4, and 1 to 4 respectively. Like in the study of Tuli in 2006, we chose threshold scores for each scale based on prior work [41-44]. In the symptom severity scale and in the physical function scale the minimal clinically important difference (MCID) is 0.5. A mean patient satisfaction score of less than 2.5 has been shown previously to represent a satisfied patient [42,43]. Despite from the subscale analysis we dichotomize "success" and "failure". When the MCID threshold was achieved in at least two domains, it was described as an overall success [44].

Secondary outcome measurements

1) Modified Roland Disability Questionnaire for Sciatica (MRDQ)
   The 23-points MRDQ is the most widely used patient-assessed measure of health for low back pain and leg pain [45-52]. This questionnaire consist of 23 questions with higher scores indicating increased disability [53]. Patrick et al compared MRDQ to patients satisfactory after from a change of 5 or more, patients feel themselves better. From a change of 12.4 all symptoms are completely gone. Others used a change of 4 or more [7]. The MRDQ will be dichotomized in "good result" (change of 4 or more) and "poor result" (change of 4 or less) [49-51].
2) Shuttle walking test (SWT)
   In this test a distance of ten meters has to be walked by the patients in a certain amount of time. This interval will be shortened until the patient does not finish the ten meters in the prescribed time. The SWT needs to change by 76 meters to ensure that walking distance is changed, but large changes can occur after surgery, and the SWT may thus provide a useful measure on an individual basis [54].
3) SF-36
   The questionnaire consists of 36 items on physical and social status of the patient subdivided in 8 domains: physical function, physical restrictions, emotional restrictions, social functioning, somatic pain, general mental health, vitality, and general health perception. The questions are scored on a scale of 0, "worst health", to 100, "ideal health" [55,56].

4) McGill pain questionnaire
   This score distinguishes three dimensions of pain: sensoric, affective and evaluative dimension [57,58].
5) Visual Analogue Scale (VAS) score of back pain and leg pain
   This parameter will measure the experienced back and leg pain intensity in the week before visiting the research nurse. Pain will be assessed on a horizontal 100 millimetres scale varying from 0 millimetre, "no pain", to 100 millimetres, "the worst pain imaginable" [59].
6) Likert scale
   This 7-point perceived recovery scale varies from "completely recovered" to "worse than ever". Like the patient global impression of change, the scale will be completed by the patient and research nurse. For analysis purposes this test will be dichotomized in "recovered" and "not recovered" [60].
7) Hospital Anxiety Depression Scale (HADS)
   This scale consists of a 7-item depression scale and a 7-anxiety scale. The score range from 0-21 with a high score being indicative for depression/anxiety.

Costs

To estimate utilities the EuroQol is used [61-64]. The EuroQol consists of 5 dimensions: mobility, self-care, daily activities, pain/discomfort, and anxiety/depression. Together with the remaining life expectation, they form QALY's. The QALY is a measure for the number of years someone still may expect, corrected for their quality. The EuroQol will be repeated once every two weeks during the first 8 weeks after surgery. These frequent EuroQol measurements during the first 8 weeks have been chosen in order to record the changes of quality of life. After this first period EuroQol will be recorded on regular basis during the patient's visit to the research nurse (see table 3: flowchart). The patients are also instructed to record a diary in which, for example, work activities will be enlisted. Furthermore direct medical costs will be estimated on basis of the cost centre method.

Complications and re-operation incidence

The research nurse and the neurosurgeon will record complications accurately. This may include infections, post-surgical haematoma, cerebrospinal fluid leakage, an increase in neurological deficit due to surgery, venous thrombosis and other side effects.

Sample Size

The sample size calculation is based on the hypothesis that the short-term results obtained after IPD are superior to the results obtained after surgical decompression. The ZCQ at eight weeks will be used as a primary result measure both to answer the first research question and to calculate the sample size. The sample size of the trial is based on a superiority design and calculated under the alternative null-hypothesis to reach sufficient power to enable a distinction between the two arms in terms of success according to ZCQ if (according to the literature) results obtained after surgical decompression will be 64% and the results obtained after IPD will be at least 84% (20% difference in favour of IPD). A sample size of 98 patients per group ensures 90% power to confirm the null hypothesis when IPD is more than 20% superior to decompression, using a likelihood ratio test in a logistic regression framework (see figure 1: sample size). Accounting for about 10% loss to follow-up, this trial will enrol 216 patients with INC (108 patients in both groups). A sample size of 80 patients per group (including 10% loss to follow-up) will ensure a power of 80%. The feasibility of reaching 216 patients available for analysis will be checked after reaching 160 evaluable patients without deblinding or even analysing the data as a group comparison. This constitutes a methodological valid approach since no multiple testing is involved and stopping further accrual is not based on an intermediate effect estimate. Since the power is based on a dichotomization of the underlying ZCQ scale, an alternative primary analysis of the ZCQ itself will also have sufficient power. The latter analysis will also take the repeated measurements structure into account.

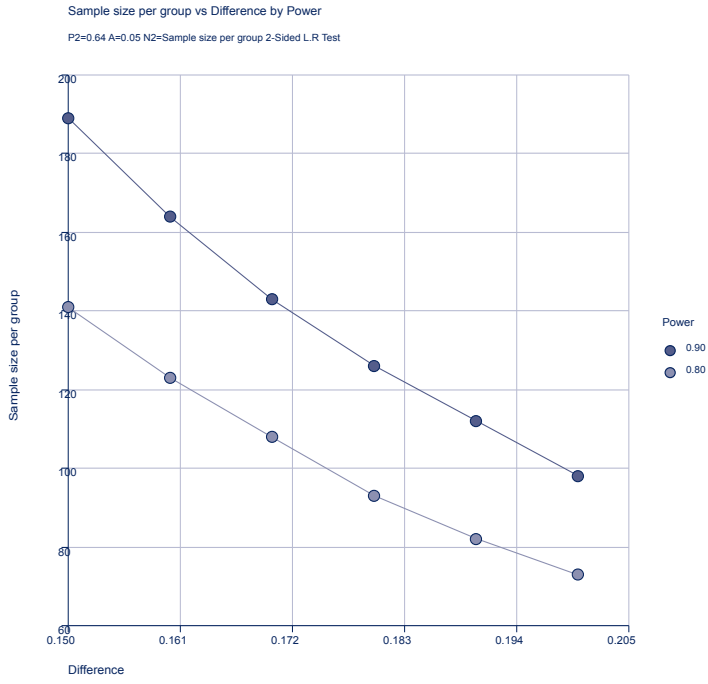
Statistical analysis

Baseline comparability will be assessed by descriptive statistics to determine whether randomization was successful. Differences in outcome between both groups, together with 95% confidence intervals, will be calculated.

Besides a difference in recovery between the two groups at two specified time points (eight weeks and one year), analysis of a difference in time to recovery will be carried out as well, using a survival analysis framework (COX hazards). All data are analysed according to the "intention-to-treat-principle". Furthermore a repeated measurements analysis of variance will be performed on the underlying continuous scales. In all analyses the first assessment of treatment effect will be the estimate of the main effect within the appropriate model, adjusted for the

stratification factors and main covariates. Secondly, an interaction term evaluating a possible effect modification of the treatment effect by the major covariates (see table 5: covariates for sub analysis) is pre-specified as being part of proper statistical modelling of the primary treatment effect. In the presence of severe interaction, the treatment effect will be presented as a function of the effect modifiers. In addition, an explorative subgroup analysis is conducted to investigate whether treatment effect varies over specific subgroups of patients (table 6: subgroups). Data will be stored via the internet-based secure data management system ProMISe of the department of Medical Statistics and Bioinformatics. The analyses will be carried out using appropriate statistical software (e.g. SPSS, version 17).

Figure 1. Sample Size



N1: number of patients needed in the IPD group, P1: the chosen success rate of the IPD group; P2: the success rate of the decompression group (0.64); N2: number of patients needed in the decompression group (equal to N1); A: the alfa is two sided 0.05.

Table 6. subgroups based on the following variables

Demographics
- age < 70 years versus > 70 years
- women versus men
Anamnestic and neurological variables
- short versus long history of back pain
- more leg pain versus more back pain
Radiological variables
- soft versus bony stenosis
- extent of stenosis during MRI examination

Table 5. Covariates for sub analysis

- Age and age banding (< 60 years, > 60 years or similar linked to groups size after recruitment))
- Long medical history of back pain
- Leg pain intensity
- Proportion leg pain/back pain
- Extent of stenosis during MRI examination
- Kind of stenosis (soft or bony)
- Sexe
- Surface area of spinal canal

Discussion

In this article a design of a RCT is presented which evaluates the (cost-) effectiveness of IPD versus decompression surgery in the treatment of intermittent neurogenic claudication. This is the first randomized prospective trial comparing these two surgical techniques. Like the Sciatica-MED trial, the research nurse and the patient are blinded for the allocated treatment [7]. The objective of this trial is to determine whether the IPD is more (cost-) effective after eight weeks compared to the conventional decompression surgery.

Abbreviations

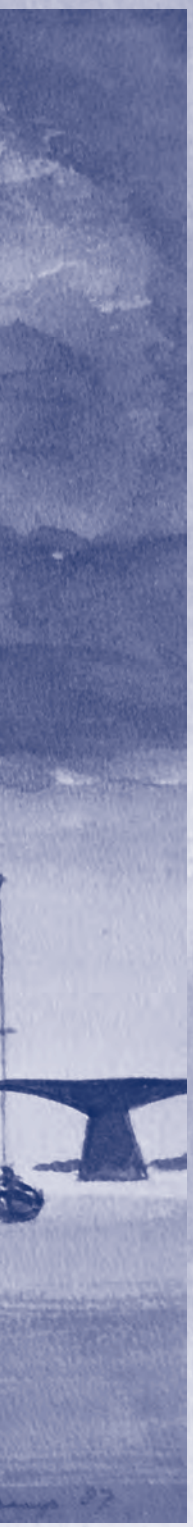
CRF, case record form; INC, intermittent neurogenic claudication; IPD, interspinous process device; MCID, minimal clinically important difference; MRDQ, modified Roland disability questionnaire; MRI, magnetic resonance imaging; QALY, Quality adjusted live years; RCT, Randomized Controlled Trial; SF-36, short form-36; SWT, Shuttle walking test; VAS: visual analogue scale; ZCQ: Zurich claudication questionnaire.

Reference List

1. EVANS JG: Neurogenic Intermittent Claudication. *Br Med J* 1964, 2: 985-987.
2. Verbiest H: [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. *Ned Tijdschr Geneeskd* 1950, 94: 2415-2433.
3. Verbiest H: [Further reports on primary stenosis of the lumbar spinal canal in adults.]. *Ned Tijdschr Geneeskd* 1951, 95: 1965-1970.
4. Verbiest H: A radicular syndrome from developmental narrowing of the lumbar vertebral canal. *J Bone Joint Surg Br* 1954, 36-B: 230-237.
5. Johnsson KE, Uden A, Rosen I: The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. *Spine (Phila Pa 1976)* 1991, 16: 615-619.
6. Arnoldi CC, Brodsky AE, Cauchoix J, Crock HV, Dommisse GF, Edgar MA *et al.*: Lumbar spinal stenosis and nerve root entrapment syndromes. Definition and classification. *Clin Orthop Relat Res* 1976, 4-5.
7. Arts MP, Brand R, van den Akker ME, Koes BW, Bartels RH, Peul WC: Tubular diskectomy vs conventional microdiskectomy for sciatica: a randomized controlled trial. *JAMA* 2009, 302: 149-158.
8. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ *et al.*: A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine (Phila Pa 1976)* 2005, 30: 1351-1358.
9. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev* 2005, CD001352.
10. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis: updated Cochrane Review. *Spine (Phila Pa 1976)* 2005, 30: 2312-2320.
11. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev* 2005, CD001352.
12. Gibson JN, Grant IC, Waddell G: The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine (Phila Pa 1976)* 1999, 24: 1820-1832.
13. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA *et al.*: Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med* 2007, 356: 2257-2270.
14. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B *et al.*: Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008, 358: 794-810.
15. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN *et al.*: Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. *J Bone Joint Surg Am* 2009, 91: 1295-1304.
16. Malmivaara A, Slatis P, Heliovaara M, Sainio P, Kinnunen H, Kankare J *et al.*: Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. *Spine (Phila Pa 1976)* 2007, 32: 1-8.
17. Turner JA, Ersek M, Herron L, Deyo R: Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. *Spine (Phila Pa 1976)* 1992, 17: 1-8.
18. Thome C, Zevgaridis D, Leheta O, Bazner H, Pockler-Schoniger C, Wohrle J *et al.*: Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. *J Neurosurg Spine* 2005, 3: 129-141.
19. Fox MW, Onofrio BM, Onofrio BM, Hanssen AD: Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. *J Neurosurg* 1996, 85: 793-802.
20. Fox MW, Onofrio BM: Indications for fusion following decompression for lumbar spinal stenosis. *Neurosurg Focus* 1997, 3: e2.
21. Airaksinen O, Herno A, Kaukanen E, Saari T, Sihvonen T, Suomalainen O: Density of lumbar muscles 4 years after decompressive spinal surgery. *Eur Spine J* 1996, 5: 193-197.
22. Senegas J: [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. *Acta Orthop Belg* 1991, 57 Suppl 1: 221-226.
23. Bartels RH: The X STOP device. *J Neurosurg Spine* 2007, 6: 620-621.

24. Chiu JC: Interspinous process decompression (IPD) system (X-STOP) for the treatment of lumbar spinal stenosis. *Surg Technol Int* 2006, 15: 265-275.
25. Eichholz KM, Fessler RG: Is the X STOP interspinous implant a safe and effective treatment for neurogenic intermittent claudication? *Nat Clin Pract Neurol* 2006, 2: 22-23.
26. Lee J, Hida K, Seki T, Iwasaki Y, Minoru A: An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J Spinal Disord Tech* 2004, 17: 72-77.
27. Mariottini A, Pieri S, Giachi S, Carangelo B, Zalaffi A, Muzii FV *et al.*: Preliminary results of a soft novel lumbar intervertebral prothesis (DIAM) in the degenerative spinal pathology. *Acta Neurochir Suppl* 2005, 92: 129-131.
28. Oppenheimer JH, DeCastro I, McDonnell DE: Minimally invasive spine technology and minimally invasive spine surgery: a historical review. *Neurosurg Focus* 2009, 27: E9.
29. Siddiqui M, Karadimas E, Nicol M, Smith FW, Wardlaw D: Effects of X-STOP device on sagittal lumbar spine kinematics in spinal stenosis. *J Spinal Disord Tech* 2006, 19: 328-333.
30. Sobottke R, Schluter-Brust K, Kaulhausen T, Rollinghoff M, Joswig B, Stutzer H *et al.*: Interspinous implants (X Stop, Wallis, Diam) for the treatment of LSS: is there a correlation between radiological parameters and clinical outcome? *Eur Spine J* 2009, 18: 1494-1503.
31. Talwar V, Lindsey DP, Fredrick A, Hsu KY, Zucherman JF, Yerby SA: Insertion loads of the X STOP interspinous process distraction system designed to treat neurogenic intermittent claudication. *Eur Spine J* 2006, 15: 908-912.
32. Tsai KJ, Murakami H, Lowery GL, Hutton WC: A biomechanical evaluation of an interspinous device (Coflex) used to stabilize the lumbar spine. *J Surg Orthop Adv* 2006, 15: 167-172.
33. Wilke HJ, Drumm J, Haussler K, Mack C, Steudel WI, Kettler A: Biomechanical effect of different lumbar interspinous implants on flexibility and intradiscal pressure. *Eur Spine J* 2008, 17: 1049-1056.
34. Kondrashov DG, Hannibal M, Hsu KY, Zucherman JF: Interspinous process decompression with the X-STOP device for lumbar spinal stenosis: a 4-year follow-up study. *J Spinal Disord Tech* 2006, 19: 323-327.
35. Verhoof OJ, Bron JL, Wapstra FH, van Royen BJ: High failure rate of the interspinous distraction device (X-Stop) for the treatment of lumbar spinal stenosis caused by degenerative spondylolisthesis. *Eur Spine J* 2008, 17: 188-192.
36. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ *et al.*: A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 2004, 13: 22-31.
37. Simons P, Werner D, Krause P, Mark P, Godde G, Reinhard A: Treatment of neurogenic intermittent claudication (NIC), secondary to lumbar spinal stenosis (LSS), with the COFLEX® Interspinous Process Decompression (IPD) system. *Early ZCQ outcome results from the European registry* 2006.
38. Stromqvist B, Berg S, Gerdhem P, Johnsson R, Sahlstrand T, Tullberg T: Indirect Decompression (X-Stop) versus Conventional Decompressive Surgery for Lumbar Spinal Claudication - A Prospective Randomized Trial. *The Spine Journal* 7 A.D., 7: 70.
39. Park SC, Yoon SH, Hong YP, Kim KJ, Chung SK, Kim HJ: Minimum 2-year follow-up result of degenerative spinal stenosis treated with interspinous u (coflex). *J Korean Neurosurg Soc* 2009, 46: 292-299.
40. Richter A, Schutz C, Hauck M, Halm H: Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. *Eur Spine J* 2009.
41. Pratt RK, Fairbank JC, Virr A: The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2002, 27: 84-91.
42. Stucki G, Liang MH, Fossel AH, Katz JN: Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995, 48: 1369-1378.
43. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN: Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996, 21: 796-803.
44. Tuli SK, Yerby SA, Katz JN: Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. *Spine (Phila Pa 1976)* 2006, 31: 1276-1280.

45. Deyo RA, Diehl AK: Patient satisfaction with medical care for low-back pain. *Spine (Phila Pa 1976)* 1986, 11: 28-30.
46. Deyo RA, Patrick DL: The significance of treatment effects: the clinical perspective. *Med Care* 1995, 33: AS286-AS291.
47. Hutchinson PJ, Laing RJ, Waran V, Hutchinson E, Hollingworth W: Assessing outcome in lumbar disc surgery using patient completed measures. *Br J Neurosurg* 2000, 14: 195-199.
48. Koes BW, van Tulder MW, Thomas S: Diagnosis and treatment of low back pain. *BMJ* 2006, 332: 1430-1434.
49. Patrick DL, Deyo RA: Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989, 27: S217-S232.
50. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB: Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995, 20: 1899-1908.
51. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB: Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995, 20: 1899-1908.
52. Roland M, Morris R: A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care. *Spine (Phila Pa 1976)* 1983, 8: 145-150.
53. Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA: Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther* 1996, 76: 268-275.
54. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE: Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992, 47: 1019-1024.
55. Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T *et al.*: Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992, 305: 160-164.
56. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD: The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994, 32: 40-66.
57. Melzack R: The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975, 1: 277-299.
58. Melzack R: The McGill pain questionnaire: from description to measurement. *Anesthesiology* 2005, 103: 199-202.
59. Carlsson AM: Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983, 16: 87-101.
60. Main CJ, Wood PL, Hollis S, Spanswick CC, Waddell G: The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine (Phila Pa 1976)* 1992, 17: 42-52.
61. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 1990, 16: 199-208.
62. Dolan P: Modeling valuations for EuroQol health states. *Med Care* 1997, 35: 1095-1108.
63. Drummond MF, Stoddart GL: Economic analysis and clinical trials. *Control Clin Trials* 1984, 5: 115-128.
64. Torrance GW: Measurement of health state utilities for economic appraisal. *J Health Econ* 1986, 5: 1-30.



Chapter 5

The Felix Trial: clinical results after one year.

Extracted from BMJ. 2013 Nov 14. 347:f6415

Moojen WA (A1/A2)
Arts MP (A2)
Jacobs WCH (A1)
Van Zwet EW (A3)
Van den Akker – Van Marle ME (A4)
Koes BW (A5)
Vleggeert-Lankamp CLAM (A1)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
[A4] Department of Medical Decision Making, Leiden University Medical Center
[A5] Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract

Objective: To assess whether interspinous process device implantation is more effective in the short term than conventional surgical decompression for patients with intermittent neurogenic claudication due to lumbar spinal stenosis.
Design: Randomized controlled trial.
Setting: Five neurosurgical centers (including one academic and four secondary level care centers) in the Netherlands.
Participants: 203 participants were referred to the Leiden-The Hague Spine Prognostic Study Group between October 2008 and September 2011; 159 participants with intermittent neurogenic claudication due to lumbar spinal stenosis at one or two levels with an indication for surgery were randomized.
Interventions: 80 participants received an interspinous process device and 79 participants underwent spinal bony decompression.
Main outcome measures: The primary outcome at short term (eight weeks) and long term (one year) follow-up was the Zurich Claudication Questionnaire score. Repeated measurements were made to compare outcomes over time.
Results: At eight weeks, the success rate according to the Zurich Claudication Questionnaire for the interspinous process device group (63%, 95% confidence interval 51% to 73%) was not superior to that for standard bony decompression (72%, 60% to 81%). No differences in disability (Zurich Claudication Questionnaire; P=0.44) or other outcomes were observed between groups during the first year. The repeat surgery rate in the interspinous implant group was substantially higher (n=21; 29%) than that in the conventional group (n=6; 8%) in the early post-surgical period (P<0.001).
Conclusions: This double blinded study could not confirm the hypothesized short term advantage of interspinous process device over conventional "simple" decompression and even showed a fairly high reoperation rate after interspinous process device implantation.

Trial registration Dutch Trial Register: NTR1307.

Introduction

Recent developments in spinal surgery implants promise less invasive procedures with superior effectiveness to conventional surgery.[1] [2] Particularly in (older) patients with spinal stenosis due to arthrosis of the facet joints, implantation of an interspinous process device is regularly offered. However, the growing incidence of low back surgery with additional implants for degenerative spine disease has raised questions from the scientific community.[3] [4] Furthermore, the economic burden of management of lumbar spine disorders (lumbar spinal disorder and lumbar disc disease combined) was already worrisome in the 1990s, when they ranked fifth on the basis of cost of hospital care.[5] The increasing use of implants, combined with a growing older population, leads to societal concerns as the cost of the management of spinal stenosis is escalating.[6] [7]

Lumbar degenerative spinal stenosis is caused by arthrosis of the facet joints and development stenosis, which can result in lumbar nerve root compression.[8] As in other acquired diseases, intermittent neurogenic claudication is usually seen in older people.[9] Severe stenosis is common in older people's spines: 30.4% of the Japanese population had a severe stenosis (average age 67.3 (range 40-93) years).[10] Why only 17.5% of these patients have typical symptoms is not yet known. Most of these patients complain of a complex of symptoms, described as leg pain (frequently in both legs), which is exacerbated by walking, prolonged standing, or lumbar extension.[8] [9] [11] [12] Classically, the cramp, tightness, pain, or discomfort in the legs will diminish after a short period of sitting or bending forward.[9] Apart from the leg pain, associated low back pain may occur.[13] The optimum treatment of lumbar spinal stenosis is generally considered to be surgical intervention, as two randomized clinical trials comparing conservative treatment with conventional bony decompression resulted in treatment effects in favor of surgery.[14] [15] The treatment outcome falls short of surgeons' expectations, as surgical decompression yields a modest outcome, being favorable in only 65% of patients.[14-18] This slightly disappointing success rate is said to be due to the destructive nature of bony decompressive surgery of the spinal column.[19] [20] Instability of the lumbar spine follows laminectomy, requiring subsequent instrumental spondylodesis.[21] [22] Spinal surgeons and the medical device industry are therefore looking for a less detrimental alternative in the surgical care for the older population.

Minimally invasive surgery has gained popularity in recent years, resulting in the development of various interspinous process devices.[23] The interspinous process device was developed to stabilize and increase the interspinous distance with indirect decompression of the dural sac and nerve roots.[24-34] Treatment of neurogenic claudication with has been shown to be superior to conservative care.[35-37] The Coflex implant (Paradigm Spine, USA) was developed as a second generation interspinous process device to give indirect decompression and even with the possibility to stabilize the lumbar spine after bony decompression. Although some medical societies in Western countries believe in additional pedicle screw fixation of the lumbar spine, the scientific gold standard of surgical treatment of lumbar spinal stenosis is bony decompression.[38] [39] All new surgical techniques to treat lumbar spinal stenosis should be compared with this technique.[14] [15] [17] [40] Although society might be subjected to media driven medicine and early adoption of surgical implants, the spinal scientific community believes that well designed comparative studies should be conducted before a new implant can replace the gold standard. Interspinous process devices have been suggested to have better short term (eight weeks) recovery than and similar long term (cost) effectiveness to bony decompression.[24-27] [30-32] [34] [36] [37] [41] [42] The purpose of this study was to assess whether interspinous process device implantation is more effective in the short term than conventional surgical decompression for patients with intermittent neurogenic claudication due to lumbar spinal stenosis.

Methods

We did a prospective, randomized, double blind, multicenter trial among patients with intermittent neurogenic claudication due to lumbar spinal stenosis after failed conservative treatment (Foraminal Enlargement Lumbar Interspinosus distraXion: FELIX trial). We compared minimally invasive treatment with interspinous process devices against usual care (conventional bony decompression). The design and study protocol have been published previously.[43]

Eligibility and randomization

Patients aged between 40 and 85 years with at least three months of intermittent neurogenic claudication due to single or two level degenerative lumbar canal stenosis and an indication for surgery were eligible. All patients were diagnosed

as having intermittent neurogenic claudication by a neurologist in one of the participating hospitals. If magnetic resonance imaging showed a lumbar spinal canal stenosis, the consulting neurosurgeon could include patients as surgical candidates for the study. At the time of enrollment, an independent research nurse verified the persistence of the symptoms. We excluded patients with a cauda equina syndrome, a herniated disc needing discectomy, history of lumbar surgery, or significant scoliosis (Cobb angle >25°) or other spinal deformities.

We used a randomized design with variable block sizes, with allocations stratified according to center. Allocations were stored in prepared opaque, coded, and sealed envelopes. The key was accessible only to the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. All patients gave informed consent. After induction of anesthesia, the prepared envelope was opened and the patient allocated to one of the treatment arms. Patients, nurses on the hospital-wards, and research nurses remained blind to the allocated treatment during the follow-up period of one year. The surgical report was kept separately from the patient's regular clinical forms and was available to the neurosurgeon only in case of complications or reoperations.

Interventions

Patients allocated to the experimental group were operated on under general anesthesia in the knee-elbow position; no bony decompression was done, and an interspinous process device was implanted by a posterior midline approach using radiographic data for localization of the appropriate level. Patients in the standard bony decompression group had surgery in the same knee-elbow position done using a similar incision length to the interspinous process device group to keep all caregivers blind to the allocated treatment. A partial resection of the adjacent laminas was executed, followed by a flavectomy with bilateral opening of the lateral recess. If judged necessary, a medial facetectomy was done. Patients in both groups received the same standard postoperative care. Patients and the research nurses who were following them were asked after every visit if they were still blind to the allocated treatment.[43]

Outcomes

The primary outcome measure was a disorder specific functional score, obtained by the Zurich Claudication Questionnaire.[44-46] The primary outcome score was assessed at baseline and at 2, 4, 8, 12, 26, and 52 weeks after surgery. The

questionnaire consists of three domains (symptom severity, physical function, and patients' satisfaction), in which respectively seven, five, and six questions are answered on a five point (symptom severity) or a four point (physical function and patients' satisfaction) scale. The subscale scores were the averages of the points obtained for every question of the subscale, with a maximum score of 5 for symptom severity and 4 for physical function and patients' satisfaction. The score increases with increasing disability. Blinded research nurses obtained the average subscale scores at every follow-up visit.[43] We considered the overall Zurich Claudication Questionnaire score to represent a "successful recovery" when at least two domain subscales were judged as "success."[47] We defined "success" on the symptom severity scale and on the physical function scale as a decrease of at least 0.5 points; a score of less than 2.5 on the patients' satisfaction subscale represented "success."[45] [46]

Secondary outcome measures were the modified Roland Disability Questionnaire for sciatica (scores range from 0 to 23, with higher scores indicating worse functional status),[48-56] a 100 mm visual analogue scale for back and leg pain (with 0 representing no pain and 100 the worst pain ever experienced),[57] the Medical Outcomes Study 36 item Short Form Generated Health Survey (SF-36) scale (based on eight scaled scores, which are the weighted sums in their sections),[58] [59] the McGill Pain Questionnaire (with 0 representing minimum pain score and 78 maximum pain score),[60] [61] and a seven point Likert-type self-rating scale of global perceived recovery as assessed by the question of whether the patient had experienced recovery (dichotomized into 1-2 for recovery and 3-7 for no recovery).[62] Furthermore, patients underwent a shuttle walking test with a predefined maximum distance and timeframe (1200 m or 15 min).[63] Patients were scored as "success" when they walked 1200 m within 15 minutes or showed an increase of more than 80 m compared with their baseline walking distance.[43] [62] [64] [65] Finally, we used a Hospital Anxiety and Depression Scale consisting of a seven item depression scale and a seven item anxiety scale (4 point scale from 0 to 3).[66] The seven items of the depression scale are related (if more than 8 points) to depression, and the seven items of the anxiety scale are related (if more than 8 points) to generalized anxiety disorder.[67] Most studies report a cut-off point at 8 points. We assessed secondary outcome scores at baseline and at two (only visual analogue scale back and leg pain), eight 12, 26, and 52 weeks. The Hospital Anxiety and Depression Scale anxiety and depression scores were obtained at baseline and after 52 weeks.

Sample size

The aim of this study was to assess whether the experimental surgical technique with an interspinous process device would be superior to conventional surgery for patients with intermittent neurogenic claudication due to lumbar spinal stenosis on short term outcome scales. Based on our main outcome score (Zurich Claudication Questionnaire) and an assumed minimal clinically important change of 20% difference in the overall success rate between the two groups at eight weeks and 10% loss to follow-up, we calculated that a sample size of 80 patients in each treatment group would be required to provide a statistical power of 0.80 and a two sided α of 0.05.[44-47] We determined this 20% success rate on the basis of the assumption that superiority would be convincing enough to change the surgical guidelines. Researchers had access to the data only after the full follow-up period of one year.

Statistical analysis

We compared groups on the basis of an intention to treat analysis. We analyzed differences between groups at all the follow-up (2, 4, 8, 12, 26, and 52 weeks) time points with repeated measurement analysis. To account for the correlation between repeated measurements of the same person, we used generalized estimating equations. We present the difference between the results for the two groups as an odds ratio for binary outcome variables and as mean differences for continuous outcome variables.[43] To investigate potential bias due to loss to follow-up, we did a sensitivity analysis for the primary outcome by assigning a poor outcome to all missing cases.

At randomization, the administrative center stratified the study for the purpose of analyzing possible heterogeneity among centers and attempting a clinical interpretation of such heterogeneity. We tested heterogeneity between centers by using center as a covariate in the mixed model. We combined those centers that were referring patients to the same hospital and the same surgeon for treatment. Hence, for the analysis of heterogeneity, a center means the actual location where the treatment (according to random allocation) took place. We used the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center Data for collection of data and checking for quality. We used IBM SPSS software, version 20.0, for all statistical analysis.

Results

Between October 2008 and September 2011, 203 patients with intermittent neurogenic claudication due to spinal stenosis were referred to the Leiden-The Hague Spine Prognostic Study Group. For all patients, the including neurosurgeon confirmed a single or two level, magnetic resonance imaging confirmed, degenerative stenosis and intermittent neurogenic claudication according to the referring neurologists. One hundred and sixty two patients gave informed consent and were enrolled in the FELIX trial (fig 1[f1]). One patient died while waiting for the operation. Two patients were found to have a severe spondylolysis of the L5-S1 segment at the final preoperative check-up and were excluded from the study, because this could cause a detrimental effect in the implant group. The remaining patients were randomly assigned to interspinous process device or decompression, and 159 patients received the allocated treatment. All patients had had intermittent neurogenic claudication for an average period of 23 (intermittent neurogenic claudication group) and 22 (decompression group) months. No significant differences were noted in baseline characteristics between patients in the two treatment arms (table 1[t1]). Seven patients were lost to follow-up in the interspinous process device group and one patient in the bony decompression group.

Successful recovery according to the Zurich Claudication Questionnaire at short term follow-up (eight weeks) was achieved by 63% of the patients in the interspinous process device group compared with 72% in the bony decompression group (odds ratio 0.73; P=0.44). Long term (one year) successful recovery according to the Zurich Claudication Questionnaire was similar in the two groups, resulting in 66% good results in the interspinous process device group and 69% in the bony decompression group (odds ratio 0.90; P= 0.77). Overall, the Zurich Claudication Questionnaire analysis showed no differences between the two treatment arms (table 2[t2]; fig 2[f2]).

Table 1. Characteristics of patients at baseline.

Characteristic	IPD group (n=80)	Decompression group (n=79)
Median (range) age (years)	66 (45-83)	64 (47-83)
Male sex	49 (61)	37 (47)
Median (range) duration of intermittent neurogenic claudication (months)	12 (2-120)	22 (1-204)
Median (range) body mass index*	27 (20-48)	28 (20-37)
Duration of back pain (categorized)	1-3 years	1-3 years
IPD patient's preferred treatment†	39 (49)	36 (46)
Bony decompression patient's preferred treatment†	0 (0)	3 (4)
No preference for specific treatment†	41 (51)	40 (50)
Mild paresis or sensory loss	54 (67)	56 (71)
Localization of stenosis:		
L2-L3	2 (3)	3 (4)
L3-L4	25 (31)	22 (28)
L4-L5	53 (66)	54 (68)
Operated at two levels:	21 (26)	16 (20)
L2-L3-L4	2 (3)	3 (4)
L2-L3 and L4-L5	1 (1)	0 (0)
L3-L4-L5	17 (21)	13 (16)
Zurich Claudication Questionnaire‡:		
Mean (SD) subscale symptom severity 0-5 scale‡	3.1 (0.5)	3.2 (0.5)
Mean (SD) subscale physical function 0-4 scale‡	2.6 (0.5)	2.6 (0.5)
Mean (SD) Roland Disability Questionnaire 23 points	13.0 (5.2)	14.4 (4.5)
Mean (SD) VAS leg pain (mm)§	52 (24)	58 (24)
Mean (SD) VAS back pain (mm)§	60 (44)	49 (25)
Median (range) SWT (m)¶	180 (20-1260) (n=70)	140 (10-1220) (n=70)
Completed SWT¶	8 (10)	13 (16)

*IPD=interspinous process device; SWT=shuttle walking distance; VAS=visual analogue scale.*Weight in kilograms divided by square of height in meters. †Patients were asked if they had any treatment preference (no preference, IPD, or bony decompression). ‡Disease specific outcome score; at baseline, score was reported in two subdomains—symptom severity (range 0-5) and physical function (range 0-4). §Intensity of pain was measured by horizontal 100 mm VAS, with 0 representing no pain and 100 worst pain ever. ¶Obtained before operation; patients were asked to walk until they got symptoms; test was scored "complete" when patients walked 1200 m in 15 min without stopping.*

Figure 1. Flow chart of enrollment and follow-up

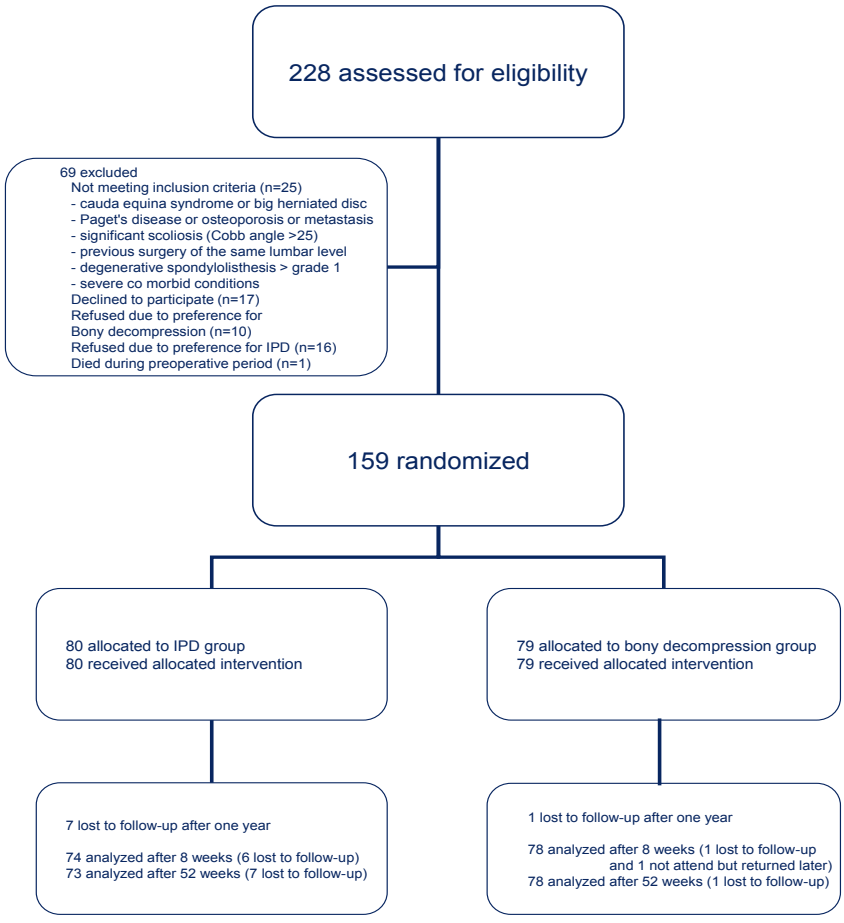


Figure 2. Scores on Zurich Claudication Questionnaire, modified Roland Disability Questionnaire, and visual analogue scale in two groups during follow-up.

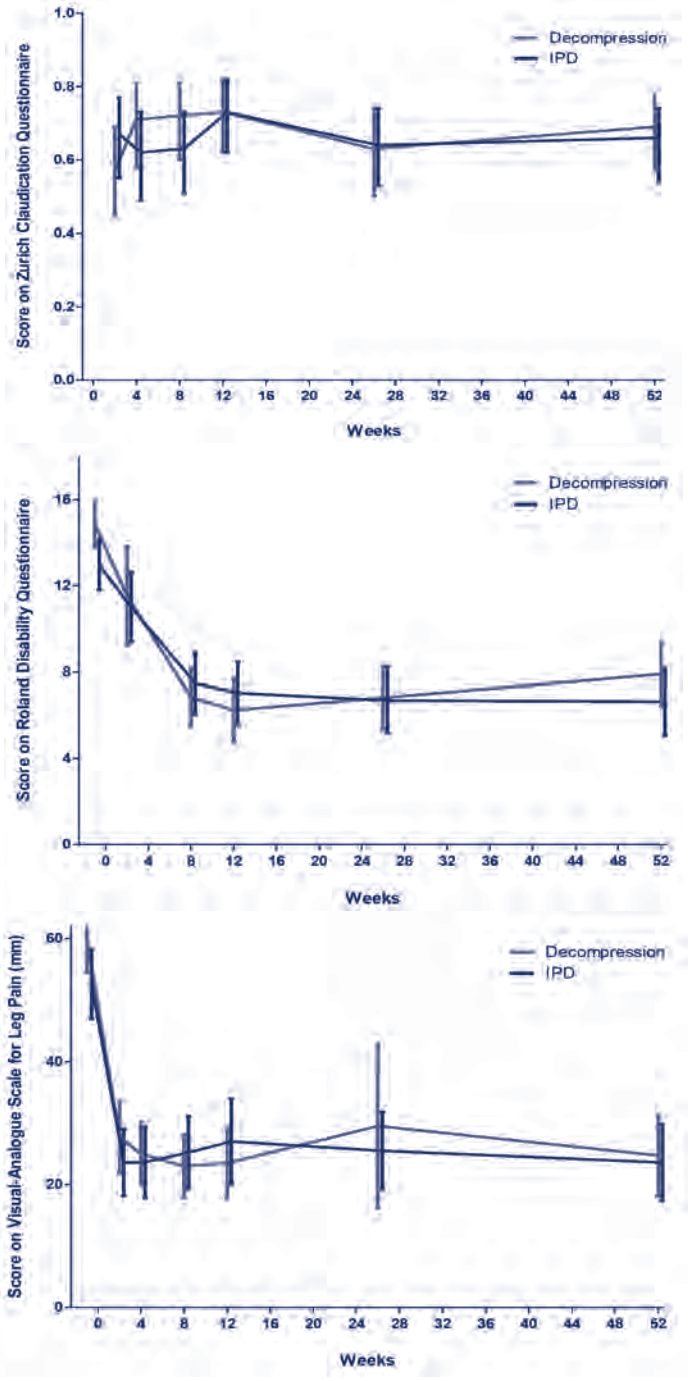


Table 2. Primary and secondary outcomes*

Variable	Baseline (n=159)		2 weeks (n=129)			8 weeks (n=142)			26 weeks (n=141)			52 weeks (n=144)		
	IPD	BD	IPD	BD		IPD	BD		IPD	BD		IPD	BD	
Primary outcome					OR (P value)			OR (P value)			OR (P value)			OR (P value)
% (95% CI) success ZCQ	NA	NA	67 (55 to 77)	57 (45 to 69)	1.64 (0.18)	63 (51 to 73)	72 (60 to 81)	0.73 (0.44)	64 (53 to 74)	63 (50 to 74)	1.20 (0.64)	66 (54 to 74)	69 (57 to 78)	0.90 (0.77)
Secondary outcomes					MD			MD			MD			MD
Mean (95% CI) MRDQ (23 points) score	13.0 (11.7 to 14.2)	14.4 (13.3 to 15.5)	10.4 (9.2 to 11.8)	10.6 (9.3 to 12.0)	0.1	7.5 (6.1 to 9.0)	6.5 (5.3 to 7.8)	1.0	NA	NA	NA	6.9 (5.4 to 8.5)	8.1 (6.6 to 9.7)	1.2 (0.28)†
Mean (95% CI) VAS back pain (0-100 mm)	60 (37 to 83)	49 (44 to56)	32 (27 to38)	33 (28 to39)	1	24 (19 to 30)	23 (17 to 28)	1	NA	NA	NA	23 (17 to 29)	31 (24 to 37)	8 (0.09)†
Mean (95% CI) VAS leg pain (0-100 mm)	52 (47 to 59)	58 (52 to 64)	23 (18 to 28)	26 (20 to 32)	3	26 (20 to 32)	22 (18 to 27)	4	NA	NA	NA	23 (17 to 30)	26 (20 to 33)	3 (0.54)†
								OR (P value)						OR (P value)
Mean (95% CI) Likert % perceived success‡	NA	NA	NA	NA	NA	51 (40 to 63)	53 (41 to 64)	0.94 (0.85)	NA	NA	NA	56 (45 to 67)	49 (38 to 60)	1.37 (0.37)

BD=bony decompression; IPD=interspinous process device; MD=mean difference; MRDQ=modified Roland Disability Questionnaire; NA=test not administered; OR=odds ratio; VAS=visual analogue scale; ZCQ=Zurich Claudication Questionnaire.

*Outcomes were analyzed with generalized estimating equations (GEE). Outcome was reported as odds ratio for better success rate when treated with IPD versus bony decompression and overall P value (based on GEE) of interaction between two groups based on continuous outcome scales with mean differences (MRDQ and VAS).

†Overall score in continuous outcome scales were not significant (MRDQ and VAS).

‡Likert global perceived recovery defined by 7 point scale from "worse" to "complete" recovery; score was dichotomized between good recovery (1-2) and bad recovery (3-7).

Table 3. Secondary outcomes at 8 and 52 weeks*

Outcome	IPD	Decompression	Odds ratio† (P value)
% success (95% CI) on shuttle walking test:	57 (0.45 to 0.68)	59 (0.47 to 0.88)	
8 weeks (increase of 80 m or complete)	(n=73)	(n=72)	0.75 (0.33)
52 weeks (increase of 80 m or complete)	57 (0.43 to 0.69) (n=66)	51 (0.40 to 0.62) (n=70)	1.25 (0.54)
Mean (95% CI) SF-36 score:			
Bodily pain 8 weeks	66 (61 to 71)	63 (58 to 68)	(0.40)‡
Bodily pain 52 weeks	66 (60 to 72)	62 (57 to 68)	
Physical functioning 8 weeks	64 (59 to 70)	67 (62 to 72)	(0.72)‡
Physical functioning 52 weeks	63 (58 to 69)	62 (56 to 68)	
Social functioning 8 weeks	74 (69 to 80)	76 (70 to 82)	(0.95)‡
Social functioning 52 weeks	77 (70 to 83)	77 (72 to 82)	
Physical role 8 weeks	44 (34 to 54)	42 (33 to 51)	(0.96)‡
Physical role 52 weeks	55 (45 to 65)	55 (46 to 65)	
Emotional role 8 weeks	74 (65 to 83)	80 (72 to 88)	(0.46)‡
Emotional role 52 weeks	74 (65 to 83)	79 (71 to 87)	
Mental health index 8 weeks	77 (73 to 81)	76 (72 to 80)	(0.92)‡
Mental health index 52 weeks	75 (70 to 80)	75 (71 to 79)	
Vitality 8 weeks	64 (60 to 68)	62 (58 to 67)	(0.60)‡
Vitality 52 weeks	61 (56 to 66)	59 (55 to 64)	
General health perception 8 weeks	67 (63 to 71)	63 (59 to 67)	(0.34)‡
General health perception 52 weeks	62 (57 to 67)	59 (55 to 63)	
Mean (95% CI) McGill Pain Questionnaire:			
8 weeks	11 (9 to 12)	10 (8 to 12)	
52 weeks	11 (9 to 13)	10 (9 to 12)	(0.70)‡
Median (range) HADS depression score§:			
Baseline	4 (0-9)	3 (1-9)	
52 weeks	3 (0-9)	3 (0-9)	
Median (range) HADS anxiety score§:			
Baseline	7 (2-14)	6 (2-12)	
52 weeks	6 (1-12)	6 (0-14)	
No (%) reoperations	21 (29)	6 (8%)	(<0.001)§
No (%) ZCQ success of reoperated patients	10 (48)	3 (50)	
No (%) operated at two levels	21 (26)	16 (18)	
No of reoperations in patients operated at 2 levels	8	1	(0.03)§
% (95% CI) ZCQ success at 8 weeks	67 (45 to 83)	48 (24 to 73)	2.50 (0.06)
% (95% CI) ZCQ success at 52 weeks	49 (29 to 69)	53 (25 to 79)	0.83 (0.83)
Mean (95% CI) duration of operation (min)	24 (22 to 26)	43 (39 to 47)	(<0.001)
Blood loss (mL)—categorized¶	10-50	50-100	(<0.001)
Complications during hospital stay	5	6	
Spinous process fractures	3	**	**
Mean (SD) hospital stay	1.83 (0.9)	1.89 (1.2)	(0.753)
No (%) blinded to allocated treatment	XX (67)	XX (86)	

HADS=Hospital Anxiety and Depression Scale; IPD=interspinous process device; ZCQ=Zurich Claudication Questionnaire. *Outcomes were analyzed with generalized estimating equations (GEE). †Odds ratio for better success rate when treated with IPD versus bony decompression, based on GEE. ‡Overall P value (based on GEE) of interaction between two groups based on continuous outcome scale (SF-36 and McGill (0-78 points)). §Score consists of sum score of 7 item (0-3 points per item) questionnaire ranging from 0 to 21 points; HADS-anxiety >8 is suspect for generalized anxiety disorder; HADS-depression >8 is suspect for depression. ¶0-10 mL, 10-50 mL, 50-100 mL, 100-200 mL. **Spinous process fractures were not registered as relevant complications in bony decompression group, so no comparison (or P value). P value with Fisher's exact test and Pearson χ^2.

Modified Roland Disability Questionnaire values at eight weeks recovered by a mean score of 7.5 for patients treated with interspinous process device and by a mean score of 6.5 for those treated with bony decompression (P=0.28). Generalized estimating equations analysis showed no differences between the two treatment arms (table 2[t2]; fig 2[t2]). Analysis of all other subscales—visual analogue scale back pain (P=0.09), visual analogue scale leg pain (P=0.54), McGill Pain Questionnaire (P=0.70), and Likert scale for perceived recovery (P=0.37)—showed no differences during the complete follow-up (table 2[t2] and 3[t3]; fig 2[f2]). We found no statistically significant difference in walking distance in the shuttle walking test at eight weeks (odds ratio 0.75; P=0.33) and 52 weeks (1.25; P=0.54) between the two treatment groups. Generalized estimating equations analysis on visual analogue scale back pain and leg pain, SF-36, McGill Pain Questionnaire, Likert score for perceived recovery, shuttle walking test, and Hospital Anxiety and Depression Scale scores also showed no differences (tables 2[t2] and 3[t3]). We did not adjust primary outcome scores for Hospital Anxiety and Depression Scale depression because of the small percentage of participants with a score of 8 or more (indicating depression).

Surgery time (24 min) was shorter in the interspinous process device group than for bony decompression (43 min) (P<0.001). Blood loss was less in the interspinous process device group (10-50 mL) than in the bony decompression group (50-100 mL) (P<0.001). Five direct (during the first initial hospital stay) postoperative complications occurred in the interspinous process device: one patient with short term (48 hours) unexplained visual disturbance, one patient with self-limiting pseudoradicular pain in the other leg, and three patients with interspinous process fractures during interspinous process device placement (table 3[t3]). Direct postoperative complications occurred in six patients in the bony decompression group: two patients with direct epidural hematoma needing reoperation and four patients with dural tears without further consequences. Late

reoperation due to absence of recovery was indicated and performed in 21 (29%) cases in the interspinous process device group compared with 6 (8%) in the bony decompression group (P<0.001). Of patients who initially received an interspinous process device and were reoperated (explantation of the device and subsequent bony decompression), 48% scored successful recovery on the Zurich Claudication Questionnaire; of patients in the bony decompression group who were reoperated, 50% scored successful recovery. The Zurich Claudication Questionnaire outcome of the patients reoperated after interspinous process device placement did not differ significantly from that of the other patients (P=0.08). Average hospital stay was similar in both groups: 1.83 days for the interspinous process device group and 1.89 days for the bony decompression group (P=0.753). Patients were successfully blinded to the treatment chosen in 67% of the IPD group and 86% of the standard decompression group.

We did a sensitivity analysis to assess the effect of the missing values for our primary outcome. Firstly, we replaced all missing values with unfavorable outcomes. This did not affect our results in any substantial way. Next, we replaced all missing values with favorable outcomes. Again, we saw no substantial changes to our results. The results for the primary outcome were therefore not sensitive to loss to follow-up.

Thirty seven patients were operated on at two levels (tables 1[t1] and 3[t3]). The subgroup of patients with lumbar spinal stenosis operated on at two levels with an interspinous process device (21 patients) had a similar outcome on the Zurich Claudication Questionnaire scale at eight weeks (odds ratio 2.5; P=0.06) and at one year (0.83; P=0.83) to those allocated to the bony decompression group (18 patients). Generalized estimating equations analysis showed no difference in Zurich Claudication Questionnaire results between one and two levels of surgery (P=0.44). However, the reoperation rate of 38% (eight patients) in the interspinous process device group at two levels was higher than the reoperation rate in the bony decompression group of 6% (one patient) (P <0.05).

We found no clinically significant heterogeneity in the outcomes between the five centers (supplementary appendix). The small difference supports the contention that the sample of hospitals is a good representation of the Dutch healthcare system.

Discussion

Implantation of an interspinous process device as definite treatment for lumbar spinal stenosis did not show the hypothesized short term superior effect over standard bony decompressive surgery. The one year follow-up results of both surgical procedures did not differ, although the reoperation rate for the interspinous process device was significantly higher than that for conventional bony decompression. Another study started in 2007 was terminated when an interim analysis showed a fourfold higher reoperation rate in the interspinous process device group.[68] The shorter operation time was the only beneficial parameter for patients in the interspinous process device group compared with the bony decompression group, but this did not result in a shorter hospital stay. Furthermore, patients operated on at two levels had an even higher reoperation rate compared with bony decompression. The absence of short term added value of the interspinous process device and the much higher reoperation rate in this study do not allow this new procedure to replace the golden standard of simple bony decompression as treatment for lumbar spinal stenosis.

Strengths and limitations of study

One of the strengths of this study is that this is the first and only blinded randomized study on this subject. Furthermore, by anomyzation during data analysis, we excluded bias as much as possible. However, the study has also features that may limit the generalizability of its findings. Firstly, selection bias could have been introduced through the opinion of the including neurosurgeon that patients with severe spinal stenosis on magnetic resonance imaging should not be offered an interspinous process device and were thus not included in the FELIX trial. However, clinical features of the patients included in this study showed baseline values (mean visual analogue scale (leg and/or back) of 60 mm at baseline) comparable to those of other large trials.[14] [15] Trials in general tend to include standard patients, but, as mentioned earlier, not all patients with stenosis have clinical complaints, which could lead to potential bias that may limit the generalizability. The number of reoperations in the interspinous process device treatment arm is very worrisome, especially because reoperations do not reach the success rate of primary surgeries; use of interspinous process devices might even prevent recovery in 20% of patients. Lastly, shuttle walking tests are believed to be the most objective parameter to classify the disease specific complaints of lumbar spinal stenosis. As

in many other studies, however, using this test for an older population is often difficult.[14] Further research should focus on finding a new objective parameter to evaluate the increasingly older population with lumbar spinal stenosis.

Comparison with other studies

Others researchers have tested the interspinous process device as an alternative for posterior and intercorporal fusion in patients with lumbar spinal stenosis.[69] [70] In Dutch practice, instrumental spondylodesis is not a standard adjuvant in spinal stenosis surgery, and nor is it standard in the modern literature.[2] [36] [37] [41] [43] Nevertheless, two studies compared a wide laminectomy combined with interspinous process device placement against treatment with wide laminectomy combined with posterior and intercorporal fusion.[69] [70] Both studies concluded that adjuvant interspinous process device treatment is as effective as lumbar 360° instrumentation in resolving neurogenic claudication. In addition, a non-randomized study had already shown that patients with intermittent neurogenic claudication treated with bony decompression alone had the same long term satisfactory outcome as did patients treated with bony decompression and adjuvant interspinous process device placement.[71] [72] Furthermore, a recent smaller non-blinded study reported similar results to those presented here. The clinical outcome of patients treated with an interspinous process device was not superior to that in patients treated with bony decompression. As in our study, patients had a higher rate (26% *v* 6%) of reoperation in the interspinous process device group.[73] All studies, including our trial, found no differences between groups with regard to postoperative visual analogue scale leg and back pain.[69] [70]

The results of this study and previous studies lead to the overall conclusion that intermittent neurogenic claudication treated with decompression alone results in a comparable outcome compared with treatment with interspinous process device alone, interspinous process device combined with bony decompression, and 360° instrumented spondylodesis. As instrumented surgery requires more from society and patients, the gold standard for intermittent neurogenic claudication treatment remains the classic bony decompression.

Conclusions

The hypothesized short term superior effect of treatment with interspinous process device over simple standard surgery was not confirmed by this double blind study. In contrast, treatment with interspinous process devices resulted in a higher reoperation rate and thus prevented a better recovery owing to the lower recovery rate after a second operation. As a spinal research group, we would not recommend the interspinous process device, considering the higher reoperation rate without a short term advantage and most likely with higher costs (interspinous process devices cost at least €2000 (£1704; $2756)). We doubt if reimbursement of interspinous process devices by society is appropriate. Furthermore, this study shows that future research in spine surgery should be very critical in the evaluation of a so called favorable outcome and weigh this against the disadvantages in robust double blind randomized trials.

Reference List

1 Park SC, Yoon SH, Hong YP, Kim KJ, Chung SK, Kim HJ. Minimum 2-year follow-up result of degenerative spinal stenosis treated with interspinous u (coflex). *J Korean Neurosurg Soc* 2009;46:292-9.

2 Stromqvist B, Berg S, Gerdhem P, Johnsson R, Sahlstrand T, Tullberg T. Indirect decompression (X-stop) versus conventional decompressive surgery for lumbar spinal claudication—a prospective randomized trial. *Poster Presentation North American Spine Society* 2010.;7(5):70.

3 Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery—the case for restraint. *N Engl J Med* 2004;350:722-6.

4 Deyo RA. Treatment of lumbar spinal stenosis: a balancing act. *Spine J* 2010;10:625-7.

5 Van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in the Netherlands. *Pain* 1995;62:233-40.

6 Carragee EJ. The increasing morbidity of elective spinal stenosis surgery: is it necessary? *JAMA* 2010;303:1309-10.

7 Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *JAMA* 2010;303:1259-65.

8 Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. *J Bone Joint Surg Br* 1954;36-B:230-7.

9 Evans JG. Neurogenic intermittent claudication. *BMJ* 1964;2:985-7.

10 Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. *Osteoarthritis Cartilage* 2012;20:1103-8.

11 Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome]. *Ned Tijdschr Geneeskd* 1950;94:2415-33.

12 Verbiest H. [Further reports on primary stenosis of the lumbar spinal canal in adults]. *Ned Tijdschr Geneeskd* 1951;95:1965-70.

13 Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis: a comparison of surgically treated and untreated patients. *Spine (Phila Pa 1976)* 1991;16:615-9.

14 Malmivaara A, Slatis P, Heliovaara M, Sainio P, Kinnunen H, Kankare J, et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. *Spine (Phila Pa 1976)* 2007;32:1-8.

15 Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008;358:794-810.

16 Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis: attempted meta-analysis of the literature. *Spine (Phila Pa 1976)* 1992;17:1-8.

17 Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med* 2007;356:2257-70.

18 Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, et al. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis: four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. *J Bone Joint Surg Am* 2009;91:1295-304.

19 Airaksinen O, Herno A, Kaukanen E, Saari T, Sihvonen T, Suomalainen O. Density of lumbar muscles 4 years after decompressive spinal surgery. *Eur Spine J* 1996;5:193-7.

20 Hanraets PR. [Radicular syndrome and low-back pain]. *Folia Psychiatr Neurol Neurochir Neerl* 1953;56(1):3-20.

21 Fox MW, Onofrio BM, Onofrio BM, Hanssen AD. Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. *J Neurosurg* 1996;85:793-802.

22 Fox MW, Onofrio BM. Indications for fusion following decompression for lumbar spinal stenosis. *Neurosurg Focus* 1997;3(2):e2.

23 Senegas J. [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. *Acta Orthop Belg* 1991;57(suppl 1):221-6.

24 Bartels RH. The X STOP device. *J Neurosurg Spine* 2007;6:620-1.

25 Chiu JC. Interspinous process decompression (IPD) system (X-STOP) for the treatment of lumbar spinal stenosis. *Surg Technol Int* 2006;15:265-75.

26 Eichholz KM, Fessler RG. Is the X STOP interspinous implant a safe and effective treatment for neurogenic intermittent claudication? *Nat Clin Pract Neurol* 2006;2:22-3.

27 Lee J, Hida K, Seki T, Iwasaki Y, Minoru A. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J Spinal Disord Tech* 2004;17:72-7.

28 Mariottini A, Pieri S, Giachi S, Carangelo B, Zalaffi A, Muzii FV, et al. Preliminary results of a soft novel lumbar intervertebral prothesis (DIAM) in the degenerative spinal pathology. *Acta Neurochir Suppl* 2005;92:129-31.

29 Oppenheimer JH, DeCastro I, McDonnell DE. Minimally invasive spine technology and minimally invasive spine surgery: a historical review. *Neurosurg Focus* 2009;27(3):E9.

30 Siddiqui M, Karadimas E, Nicol M, Smith FW, Wardlaw D. Effects of X-STOP device on sagittal lumbar spine kinematics in spinal stenosis. *J Spinal Disord Tech* 2006;19:328-33.

31 Sobottke R, Schluter-Brust K, Kaulhausen T, Rollinghoff M, Joswig B, Stutzer H, et al. Interspinous implants (X Stop, Wallis, Diam) for the treatment of LSS: is there a correlation between radiological parameters and clinical outcome? *Eur Spine J* 2009;18:1494-503.

32 Talwar V, Lindsey DP, Fredrick A, Hsu KY, Zucherman JF, Yerby SA. Insertion loads of the X STOP interspinous process distraction system designed to treat neurogenic intermittent claudication. *Eur Spine J* 2006;15:908-12.

33 Tsai KJ, Murakami H, Lowery GL, Hutton WC. A biomechanical evaluation of an interspinous device (Coflex) used to stabilize the lumbar spine. *J Surg Orthop Adv* 2006;15:167-72.

34 Wilke HJ, Drumm J, Haussler K, Mack C, Steudel WI, Kettler A. Biomechanical effect of different lumbar interspinous implants on flexibility and intradiscal pressure. *Eur Spine J* 2008;17:1049-56.

35 Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC. Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. *Eur Spine J* 2011;20:1596-606.

36 Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 2004;13:22-31.

37 Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine (Phila Pa 1976)* 2005;30:1351-8.

38 Gibson JN, Waddell G, Grant IC. Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev* 2000;(2):CD001352.

39 Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: updated Cochrane review. *Spine (Phila Pa 1976)* 2005;30:2312-20.

40 Thome C, Zevgaridis D, Leheta O, Bazner H, Pockler-Schoniger C, Wohrle J, et al. Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. *J Neurosurg Spine* 2005;3:129-41.

41 Kondrashov DG, Hannibal M, Hsu KY, Zucherman JF. Interspinous process decompression with the X-STOP device for lumbar spinal stenosis: a 4-year follow-up study. *J Spinal Disord Tech* 2006;19:323-7.

42 Verhoof OJ, Bron JL, Wapstra FH, van Royen BJ. High failure rate of the interspinous distraction device (X-Stop) for the treatment of lumbar spinal stenosis caused by degenerative spondylolisthesis. *Eur Spine J* 2008;17:188-92.

43 Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial: double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. *BMC Musculoskelet Disord* 2010;11:100.

44 Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2002;27:84-91.

45 Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369-78.

46 Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996;21:796-803.

47 Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. *Spine (Phila Pa 1976)* 2006;31:1276-80.

48 Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. *Spine (Phila Pa 1976)* 1986;11:28-30.

49 Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. *Med Care* 1995;33(4 suppl):AS286-91.

50 Hutchinson PJ, Laing RJ, Waran V, Hutchinson E, Hollingworth W. Assessing outcome in lumbar disc surgery using patient completed measures. *Br J Neurosurg* 2000;14:195-9.

51 Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006;332:1430-4.

52 Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27(3 suppl):S217-32.

53 Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995;20:1899-908.

54 Roland M, Morris R. A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care. *Spine (Phila Pa 1976)* 1983;8:145-50.

55 Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)* 1983;8:141-4.

56 Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther* 1996;76:268-75.

57 Carlsson AM. Assessment of chronic pain. I: aspects of the reliability and validity of the visual analogue scale. *Pain* 1983;16:87-101.

58 Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4.

59 McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). III: tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.

60 Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.

61 Melzack R. The McGill pain questionnaire: from description to measurement. *Anesthesiology* 2005;103:199-202.

62 Main CJ, Wood PL, Hollis S, Spanswick CC, Waddell G. The distress and risk assessment method: a simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine (Phila Pa 1976)* 1992;17:42-52.

63 Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992;47:1019-24.

64 Arts MP, Brand R, van den Akker ME, Koes BW, Bartels RH, Peul WC. Tubular diskectomy vs conventional microdiskectomy for sciatica: a randomized controlled trial. *JAMA* 2009;302:149-58.

65 Peul WC, Van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 2007;356:2245-56.

66 Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *J Psychosom Res* 2002;52:69-77.

67 Rabe-Jablonska J, Bienkiewicz W. [Anxiety disorders in the fourth edition of the classification of mental disorders prepared by the American Psychiatric Association: diagnostic and statistical manual of mental disorders (DMS-IV—options book)]. *Psychiatr Pol* 1994;28:255-68.

68 Lønne G. Treatment of lumbar spinal stenosis; comparison of two different surgical methods; mini-invasive decompression to X-stop (LSSS). http://clinicaltrials.gov/show/NCT00546949.

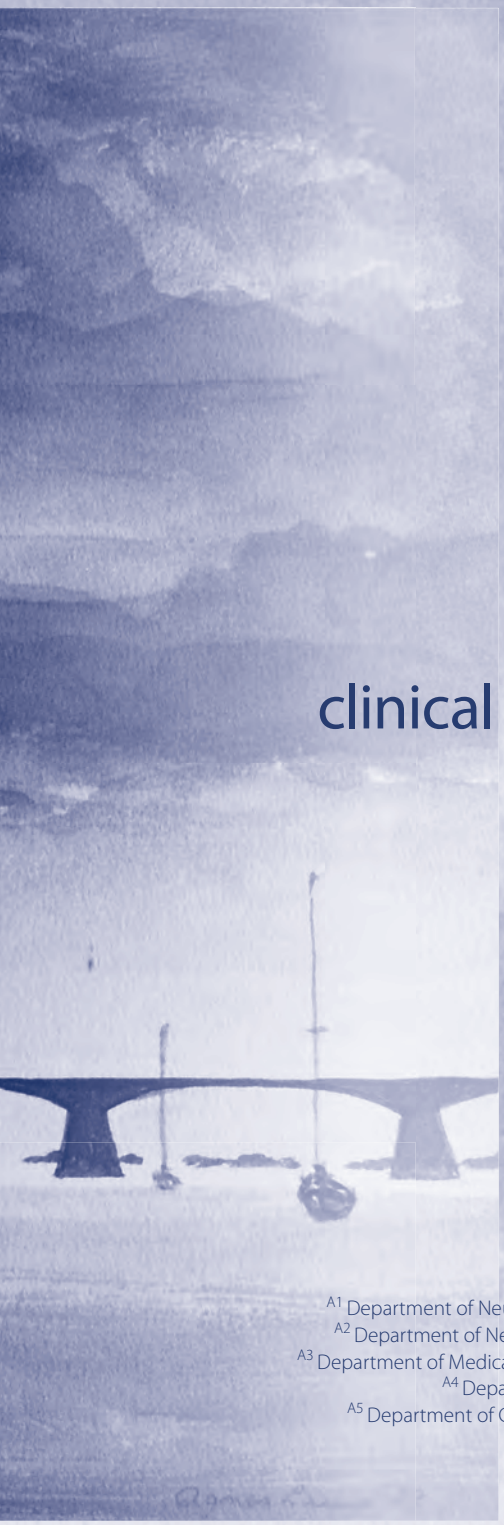
69 Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective randomized, multicenter, Food and Drug administration Investigational Device Exemption trial. Davis RJ, Errico TJ, Bae H, Auerbach JD. Spine (Phila Pa 1976). 2013 Aug 15;38(18):1529-39.

70 Liu J, Liu H, Li T, Zeng J, Song Y, Liu L, et al. [A comparative study between Coflex interspinous dynamic reconstruction and lumbar 360 degrees fusion in treating single-level degenerative lumbar spinal disorders]. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi* 2011;25:693-8.

71 Richter A, Schutz C, Hauck M, Halm H. Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. *Eur Spine J* 2010;19:283-9.

72 Richter A, Halm HF, Hauck M, Quante M. 2-year follow-up after decompressive surgery with and without implantation of an interspinous device for lumbar spinal stenosis: a prospective controlled study. *J Spinal Disord Tech* 2012; published online 4 May.

73 Stromqvist BH, Berg S, Gerdhem P, Johnsson R, Moller A, Sahlstrand T, et al. X-Stop versus decompressive surgery for lumbar neurogenic intermittent claudication: randomized controlled trial with 2-year follow-up. *Spine (Phila Pa 1976)* 2013;38:1436-42.



Chapter 6

The Felix Trial: clinical results after two years.

Extracted from Eur Spine J. 2014.
Selected best paper session EuroSpine 2014.

Moojen WA (A1/A2)
Arts MP (A2)
Jacobs WCH (A1)
Van Zwet EW (A3)
Van den Akker – Van Marle ME (A4)
Koes BW (A5)
Vleggeert-Lankamp CLAM (A1)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
[A4] Department of Medical Decision Making, Leiden University Medical Center
[A5] Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract

Objectives: Interspinous process devices (IPDs) are implanted to treat patients with intermittent neurogenic claudication (INC) based on lumbar spinal stenosis. Although widely implemented by the surgical community, effectiveness has not been compared with the gold standard bony decompression. It is hypothesized that patients with lumbar spinal stenosis treated with IPD have a faster short-term recovery, an equal outcome after two years and less back pain compared with bony decompression.

Design: A randomized design with variable block sizes was used, with allocations stratified according to center. Allocations were stored in prepared opaque, coded and sealed envelopes and patients and research nurses were blind throughout the follow-up.

Setting: 5 neurosurgical centers (including one academic and four secondary level care centers) included participants.

Participants: 211 participants were referred to the Leiden-The Hague Spine Prognostic Study Group. 159 participants with INC based on lumbar spinal stenosis at one or two levels with an indication for surgery were randomized into two groups. Patients and research nurses were blinded for the allocated treatment throughout the study period.

Interventions: 80 participants received an IPD and 79 participants underwent spinal bony decompression.

Main outcome measures: The primary outcome at long-term (two year) follow-up was the score for the Zurich Claudication Questionnaire. Repeated-measurement analyses were applied to compare outcomes over time.

Results: At two years, the success rate according to the Zurich Claudication Questionnaire for the IPD group (69 % (95 % CI 57 to 78 %)) did not show a significant difference compared with standard bony decompression (60 % (95 % CI 48 to 71 %) p-value 0.2). Re-operations, because of absence of recovery, was indicated and performed in 23 cases (33%) of the IPD group versus 6 (8%) patients of the bony decompression group ($p<0.01$). Furthermore, long-term VAS back pain was significantly higher (36 mm on a 100 mm scale (95% CI 24-48)) in the IPD group compared to the bony decompression group (28 mm (95% CI 23-34) p-value 0.04).

Conclusions: This double blinded study could not confirm advantage of IPD without bony decompression over conventional 'simple' decompression, two years after surgery. Moreover, in the IPD treatment arm, the reoperation rate was higher and back pain was even slightly more intense compared to the decompression treatment arm.

Trial registration: Dutch Trial Register Number: NTR1307

Background

Intermittent neurogenic claudication (INC) caused by lumbar spinal stenosis (LSS) is common in the elderly[1-3]. Concomitant with progressive spinal canal narrowing over the years, patients start to develop the typical symptoms due to compression of the roots of the cauda equina: leg pain (frequently both legs) exacerbated by walking, prolonged standing or lumbar extension, and sometimes associated back pain.[3-7] Surgical treatment is considered to be superior to non-surgical treatment.[8;9] Patients' satisfaction after treatment is relatively low.[8-10] This disappointing outcome was believed to be caused by the destructive nature of bony decompression.[11-13] Nowadays some centers even opt for combining bony decompression with instrumented spondylodesis (pedicle screws and/or intercorporal cages) as the golden standard for treatment of patients with INC caused by LSS.[14-16] Many different new treatment options were therefore developed in the 80s and 90s, including less invasive procedures. In particular in (elderly) patients with LSS due to arthrosis of the facet joints, implantation of interspinous process device (IPD) is regularly offered instead of conventional bony decompression.[17;18] Neurogenic claudicatio treatment with IPD has been demonstrated to be superior compared with conservative care.[19-23] The IPD was developed to increase the interspinous distance with indirect decompression of the dural sac and nerve roots due to flexion of the involved segments, and to widen the entry of the spinal root canal at the same time.[17-19;24-30] Additionally, patients are hypothetically expected to have less postoperative pain, a shorter hospital stay, a faster short-term recovery and less back pain at long-term follow-up.

We previously published the short term 1 year results of a double-blind randomized trial comparing treatment with IPDs to bony decompression in patients with intermittent neurogenic claudication due to LSS. [31] Patients that were treated with an IPD without bony decompression showed similar rates of recovery at eight weeks and at 1 year compared to patients treated with bony decompression, although the repeat surgery rate in the interspinous implant group was substantially higher (29 %) in the early post-surgical period compared with the decompression group (8%; p-value <0.001). The 2-year results of the aforementioned trial are presented in the current paper.

Methods

A prospective, randomized double-blind multicenter trial was conducted among patients with INC based on LSS after failed conservative treatment (Foraminal Enlargement Lumbar Interspinosus distraXion: Felix trial). Minimal invasive therapy with placement of an IPD, without any attempt to decompress the spinal canal was compared to the usual care being conventional bony decompression. The medical ethics committees at the five participating hospitals approved the protocol, including an approval for randomization after anesthetic induction. Written informed consent was obtained from all patients. The design and study protocol were published previously.[32] Dutch Trial Register Number: NTR1307.

Eligibility and randomization

Patients between 40 and 85 years with at least three months of INC due to single or two level degenerative lumbar canal stenosis and an indication for surgery were eligible. All patients were diagnosed with INC by a neurologist in one of the participating hospitals. If MRI demonstrated a lumbar spinal canal stenosis, patients could be included as surgical candidates for the study by the consulting neurosurgeon. At the time of enrollment, an independent research nurse verified the persistence of the symptoms. Patients with a cauda equina syndrome, a herniated disc needing discectomy, history of lumbar surgery or those with significant scoliosis (Cobb angle >25 degrees) or other spinal deformities were excluded.

A randomized design with variable block sizes was used, with allocations stratified according to center. Allocations were stored in prepared opaque, coded and sealed envelopes. The key was only accessible to the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. All patients gave informed consent. After induction of anesthesia the prepared envelope was opened and randomized allocation to one of the treatment arms was performed. Patients, nursery department and research nurses remained blind for the allocated treatment during the follow-up period of two years. The surgical report was kept separately from the regular clinical patient forms and was only available for the neurosurgeons in case of complications or reoperations.

Interventions

Patients allocated to the experimental group were operated on general anesthesia in knee-elbow position; no bony decompression was performed and an IPD was

implanted by a posterior midline approach using x-ray data for localization of the proper level.
Patients in the standard bony decompression group underwent surgery in the same knee-elbow position using a similar incision length as the IPD group in order to keep all caregivers blind for the allocated treatment. A partial resection of the adjacent laminae was executed, followed by a flavectomy with bilateral opening of the lateral recess. If judged necessary, a medial facetectomy was performed. Patients of both groups received the same standard postoperative care. Patients and research nurses who were following these patients were asked after every visit if they were still blind for the allocated treatment.[32]

Outcomes

The primary outcome measure was a disorder-specific functional score, obtained by the Zurich Claudication Questionnaire (ZCQ).[33-35] The primary outcome score was assessed at baseline, direct postoperatively (2 weeks) and at 4, 8, 12, 26, 52, and 104 weeks. The ZCQ consists of three domains (symptom severity, physical function and patient satisfaction) in which respectively seven, five and six questions had to be answered on a five point (symptom severity) or a four point (physical function and patient satisfaction) scale. The subscale scores were the averages of the points obtained for every question of the subscale, and were maximized to 5 (symptom severity) or 4 (physical function and patient satisfaction). The score increases with increasing disability. The average subscale scores were obtained at every follow-up moment by blinded research nurses.[32] The overall ZCQ score was considered to be a 'successful recovery' when two domain subscales at least were judged as 'success'.[36] 'Success' on the symptom severity scale and on the physical function scale was defined as a decrease of at least 0.5 points. A score of less than 2.5 on the patient satisfaction subscale was defined as 'success'.[34;35]

Secondary outcome measures were the Modified Roland Disability Questionnaire for sciatica (scores range from 0 to 23, with higher scores indicating worse functional status), [37-45] 100mm visual-analogue scale (VAS) back and leg pain (with 0 representing no pain and 100 the worst pain ever experienced), [46] Medical Outcome Study 36-item short-form Generated Health Survey (SF-36) scale (based on eight scaled scores, which are the weighted sums in their sections),[47;48] McGill pain questionnaire (with 0 representing minimum pain score and 78 maximum pain score),[49;50] and a 7-point Likert self-rating scale of global perceived recovery as

given by the question whether the patient experienced recovery (dichotomized in 1-2 recovery and 3-7 no recovery) compared to the baseline status.[51] Furthermore, a Hospital Anxiety Depression Scale (HADS) consists of a 7-item depression scale and a 7-item anxiety scale (4 point scale from 0-3) were obtained.[52] The seven items of the HADS-depression scale are related (if more than 8 points) to depression and the seven items on the HADS-anxiety scale are related (if more than 8 points) to generalized anxiety disorder.[53] Most studies report a cut-off point at 8 points.[53] Last, patients underwent a Shuttle Walking Test (SWT) with a predefined maximum distance and timeframe (1200 meters or 15 minutes).[54] Patients were scored "success" when they walked 1200 meters within 15 minutes or demonstrated an increase of more than 80 meters compared to baseline walking distance.[32;51;55;56] Secondary outcome scores were assessed at baseline (VAS scores) and at 2 (only VAS back and leg pain), 8, 12, 26, 52 and 104 weeks. The HADS anxiety and depression were obtained at baseline and after 52 and 104 weeks.

Sample size

The aim of this study was to assess whether the experimental surgical technique of IPD without bony decompression would be comparable to conventional surgery for patients with INC due to LSS at the time point of two years after surgery. Based on our primary outcome score (ZCQ) and an assumed minimal clinically important change (MCIC) of 20% difference in the overall success rate between the two groups at 8 weeks and 10% loss to follow-up, it was calculated that a sample size of 80 per treatment group would be required to provide a statistical power of 0.80 and a two-sided alpha of 0.05.[33-36] This difference of 20% in success rate was decided based on the assumption that this level of superiority would be convincing enough to change the surgical guidelines and reimburse the costs of the IPD implant. Data from the 104 weeks follow-up were only accessible for the researchers after completion of the full 2-year follow-up period.

Statistical Analysis

Groups were compared based on an intention-to-treat analysis. Differences between groups at all follow-up (2, 4, 8, 12, 26, 52, 104 weeks) time points were analysed with repeated measurement analysis. To account for the correlation between repeated measurements of the same individual, Generalized Estimating Equations (GEE) were used. The difference between the results for the two groups were presented as an Odds Ratio (OR) for binary outcome variables and as mean

differences for continuous outcome variables.[65] To address potential bias due to loss to follow-up a sensitivity analysis was performed for the primary outcome by assigning a poor outcome to all missing cases and a second analysis was performed for the primary outcome by assigning a favourable outcome to all missing cases. At randomization, the study was stratified by the (administrative) center for the purpose of analysing possible heterogeneity among centers and attempting a clinical interpretation of such heterogeneity. Data collection and checking for quality were performed with the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. IBM SPSS software, version 20.0, was used for all statistical analysis.

Results

Between October 2008 and September 2011 205 patients with INC due to spinal stenosis were referred to the participating hospitals. Patients with a single or two-level, MRI confirmed, degenerative lumbar stenosis and INC according to their referring neurologists were screened for inclusion by the including neurosurgeon. For 162 patients signed informed consent was obtained and the patients were enrolled in the Felix Trial (Figure 1). One patient died during the time waiting for the operation. Two patients revealed a severe spondylolysis of the L5-S1 segment at final preoperative check-up and were excluded from the study. The remaining patients were randomly assigned to IPD without bony decompression or conventional decompression. In effect, 159 patients received the allocated treatment. All patients were suffering from INC for an average period of 23 (IPD group) and 22 (decompression group) months. No significant differences were noted in baseline characteristics between patients in the two treatment arms (Chapter 6, table 1). Ten patients were lost to follow-up in the IPD group and five patients in the bony decompression group at two years after surgery.

Successful recovery according to ZCQ at long-term follow-up (two years) was achieved in 69 % of the patients in the IPD group versus 60 % of the patients in the bony decompression group (OR 0.65; p=0.20). Overall, ZCQ analysis revealed no differences between the two treatment arms (table 2 and figure 2). MRDQ values at long-term (two years) decreased with 5.5 points for patients treated with IPD and with 6.3 points for patients treated with bony decompression (p=0.65). MRDQ

values at 104 weeks were equal compared with the 52 week's value in the bony decompression group and slightly – not significant - higher (0.6 on a 23 point scale) in the IPD group. GEE analysis showed no differences between the two treatment arms (table 2 and figure 2). Analysis of all other subscales, VAS back pain (p= 0.26), VAS leg pain (p= 0.22), (p= 0.52), did not show any differences between treatments at all-time points during the complete follow-up (table 2&3 and figure 2).

Figure 1. Flow diagram of patients through two years of study

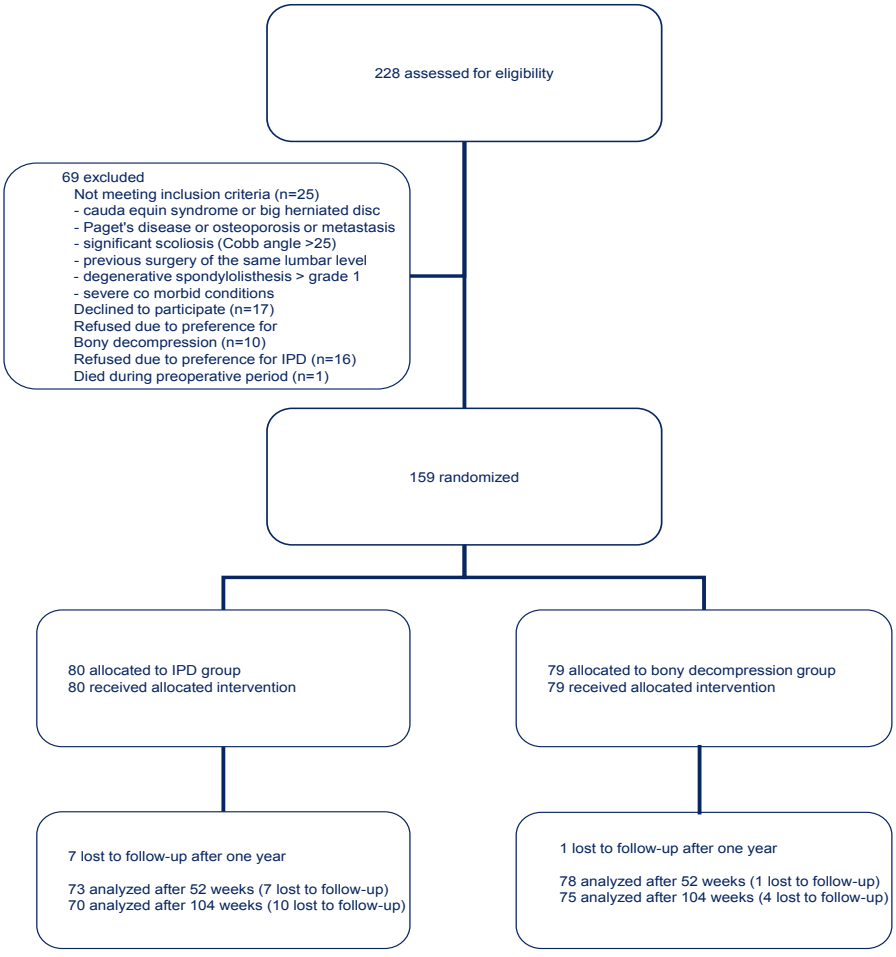
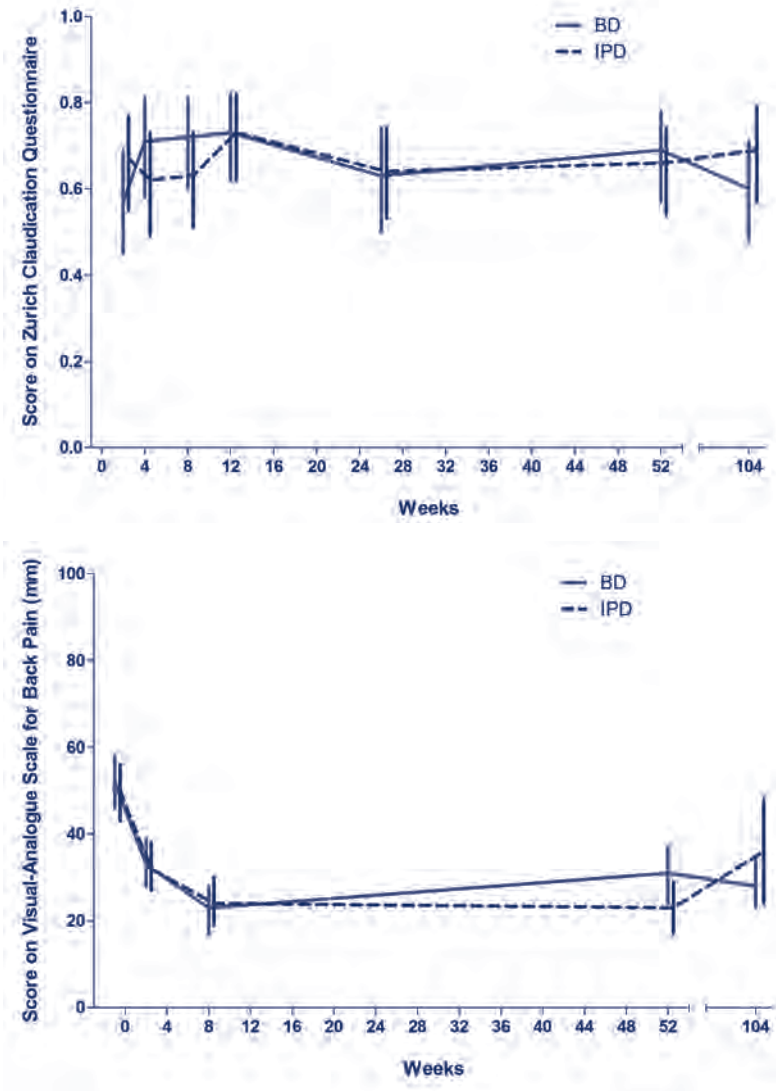


Figure 2. outcome based on the ZCQ and VAS back-pain



BD bony decompression; IPD interspinous process devices

Table 2. Primary and Secondary Outcome‡

Variable	8 weeks			52 weeks			104 weeks		
	IPD	BD	OR (p-value)	IPD	BD	OR (p-value)	IPD	BD	OR (p-value)
Primary outcome									
Success ZCQ - % (CI)	63 N=73 (51-73)	72 N=78 (60-81)	0.73 (0.44)	67 N=73 (54-74)	68 N=78 (57-78)	0.90 (0.77)	69 N=70 (57-79)	60 N=75 (48-71)	0.65 (0.20)
Secondary outcome	IPD	BD	mean difference	IPD	BD	mean difference	IPD	BD	mean difference
Mean MRDQ (23 points) score – mean (CI)	7.5 (6.1-9.0)	6.5 (5.3-7.8)	1.0	6.9 (5.4-8.5)	8.1 (6.6-9.7)	1.2	7.5 (5.6-9.5)	8.1 (6.6-9.6)	0.6 ***(0.65)
VAS back pain (0-100mm) – mean (CI)	24 (19-30)	23 (17-28)	1	23 (17-29)	31 (24-37)	8	36 (24-48)	28 (23-34)	18 ***(0.26)
VAS leg pain (0-100mm) – mean (CI)	26 (20-32)	22 (18-27)	4	23 (17-30)	26 (20-33)	3	21 (15-27)	26 (20-32)	5 ***(0.22)
	IPD	BD	OR p-value	IPD	BD	OR p-value	IPD	BD	OR p-value
Likert percentage of successful perceived – mean (CI)#	51 (40-63)	53 (41-64)	0.94 (0.85)	56 (45-67)	49 (38-60)	1.37 (0.37)	54 45-69	46 32-55	1.21 (0.52)

‡the outcomes were analyzed with Generalized Estimating Equations (GEE). Outcome was reported with an (RC) regression coefficient (beta) on a better success rate when treated with IPD versus bony decompression, and overall p-value (based on GEE) of the interaction between two groups based on a continuous outcome scale (MRDQ and VAS). N stands for number of patients analyzed. CI denotes 95% confidence interval, ZCQ Zurich claudication questionnaire, MRDQ Modified Roland Disability Questionnaire, VAS visual-analogue scale. Dashes denote tests not administrated. *** Overall score in the continuous outcome scales were not significant (MRDQ and VAS). # Likert global perceived recovery is defined by a 7 point scale from "worse" to "complete" recovery. The score was dichotomized between (1-2) good recovery and bad recovery (3-7)

Table 3. Secondary Outcome*

	IPD	decompression	(p-value)
Percentage success on SWT	Success - % (CI)	Success - %(CI)	‡OR (p-value)
	(n=73)	(n=72)	
- 8 weeks (increase of 80 or complete)	57 (45-68)	59 (47-0.88)	0.75 (0.33)
	(n=66)	(n=70)	
- 52 weeks (increase of 80 or complete)	57 (43-69)	51 (40-62)	1.25 (0.54)
	(n=60)	(n=59)	
- 104 weeks (increase of 80 or complete)	63 (52-75)	62 (50-73)	0.90 (0.76)
SF 36	Mean (95% CI.)	Mean (95% CI.)	(overall p-value)***
Bodily pain 8 weeks	66 (61-71)	63 (58-68)	(0.51)
52 weeks	66 (60-72)	62 (57-68)	
104 weeks	64 (54-75)	60 (55-66)	
Physical functioning 8 weeks	64 (59-70)	67 (62-72)	(0.35)
52 weeks	63 (58-69)	62 (56-68)	
104 weeks	63 (56-69)	59 (53-65)	
Social functioning 8 weeks	74 (69-80)	76 (70-82)	(0.05)
52 weeks	77 (70-83)	77 (72-82)	
104 weeks	80 (75-86)	72 (66-78)	
Physical role 8 weeks	44 (34-54)	42 (33-51)	(0.24)
52 weeks	42 (23-60)	55 (46-65)	
104 weeks	42 (23-61)	55 (45-65)	
Emotional role 8 weeks	74 (65-83)	80 (72-88)	(0.30)
52 weeks	74 (65-83)	79 (71-87)	
104 weeks	65 (46-83)	75 (66-85)	
Mental health index 8 weeks	77 (73-81)	76 (72-80)	(0.81)
52 weeks	75 (70-80)	75 (71-79)	
104 weeks	74 (70-79)	73 (69-78)	
Vitality 8 weeks	64 (60-68)	62 (58-67)	(0.36)
52 weeks	61 (56-66)	59 (55-64)	
104 weeks	54 (45-63)	59 (65-63)	
General health perception 8 weeks	67 (63-71)	63 (59-67)	(0.01)
52 weeks	62 (57-67)	59 (55-63)	
104 weeks	64 (59-69)	56 (51-60)	
McGill pain questionnaire (0-78 points)	Mean (95% CI.)	Mean (95% CI.)	(overall p-value)***
- 8 weeks	11 (9-12)	10 (8-12)	
- 52 weeks	11 (9-13)	10 (9-12)	
- 104 weeks	9	11	(0.37)
HADS depression ##	Median (range)	Median (range)	
- Baseline	4 (0-9)	3 (1-9)	
- 52 weeks	3 (0-9)	3 (0-9)	
- 104 weeks	3 (0-11)	3 (0-9)	
HADS anxiety ##	Median (range)	Median (range)	
- Baseline	7 (2-14)	6 (2-12)	
- 52 weeks	6 (1-12)	6 (0-14)	
- 104 weeks	4 (0-12)	4 (0-12)	
	N=70	N=75	
Reoperations (%)	23 (33%)	6 (8%)	(<0.01)#

Table 3. Continued

	IPD	decompression	(p-value)
Operated on two levels – no. (%)	21(26%)	16(18*)	
reoperations in patients operated on two levels	8	1	(0.03)#
Duration of operation – minutes (95% CI)	24 (22-26)	43 (39-47)	(<0.001)
Blood loss – categorized**	10-50mL	50-100mL	(<0.001)
Complications during hospital stay	4	6	
Of which spinous process fractures	3	#	#
Hospital stay after reoperation	1.83 (SD 0,9) 2.41 (SD 2.1) (N=23)	1.89 (SD 1.2) 3.00 (SD 1.9) (N=6)	(0.753)
Blinded to allocated treatment	67%	86%	

*The outcomes were analyzed with Generalized Estimating Equations (GEE).
CI denotes confidence interval, SF-36 the Medical Outcomes Study 36-Item, and McGill pain questionnaire.
N stands for the number of patients analyzed.

However, the back-pain in the IPD group increased in the second year after surgery in comparison to the one year time point (from 23 mm at 52 weeks to 36 mm VAS back pain at 104 weeks). In contrast, the back-pain in the bony decompression group remained equal (31mm at 52 weeks and 28 at 104 weeks), (p-value 0.04). GEE analysis on SF36 and McGill pain scores showed no differences (table 2 & 3). The dichotomized Likert perceived recovery scores showed 54% successful in IPD group and 46% successful in bony decompression group (OR 1.21, p-value 0.52). GEE analysis on HADS scores showed no differences. Primary outcome scores were not adjusted for HADS depression due to the small percentage of participants with an HADS depression of 8 or more (indicating depression). There was no difference in walking distance in the SWT at 104 weeks between the two treatment groups (ORs 0.90 p-values 0.76) and no difference at 104 weeks in both groups compared with the walking distance at 52 weeks (p-value 0.54).

Direct (post)operative complications occurred in six patients in the bony decompression group: two patients with direct epidural hematoma needing reoperation, four patients with dural tears without further consequences. Five patients had complications after IPD treatment: three patients with spinous process fractures, and one patient was explored at the wrong level which was corrected during the same procedure. Re-operations, because of absence of recovery, was indicated and performed in 23 cases (33%) of the IPD group versus 6 (8%) patients of the bony decompression group (p<0.01). This is also, compared with the one year results (17 reoperations in the IPD group and 5 in the bony decompression group),

a bigger increase in the IPD group without bony decompression in the second year of follow-up. Average hospital stay was similar in both groups: 1.83 days for the IPD group per patient (without hospital stay when operated for the second time) and 1.89 days for the bony decompression group (p= 0.753). After reoperations patients were no longer blind for the type of treatment after reoperation.
Sensitivity analysis was performed to assess the impact of the missing values in our primary outcome. First, all missing values were replaced by unfavourable outcomes. This did not affect our results in any substantial way. Next, all missing values were replaced by favourable outcomes. Again, there were no substantial changes to our results. The results concerning the primary outcome were therefore not sensitive to loss to follow-up. There was no clinically significant heterogeneity found in the outcomes between the 5 centers (supplementary appendix). The small difference will support that the sample of hospitals is a good representation of the Dutch Health Care system with high-complex patients' centers and less-complex patients' centers.

Discussion

The long-term follow-up did not show important differences in results (based on the ZCQ) comparing treatment with IPD without bony decompression and conventional bony decompression in patients with INC based on LSS. Previously published short-term results did not show any short-term benefit (based on the ZCQ) of treatment of IPD compared with bony decompression and at long follow-up the ZCQ rate of success was slightly higher for the IPD group, but not significantly.[31] Furthermore, similar to the published one year analysis, the re-operation rate was significantly higher (overall and in the period between 52 and 104 weeks) in the IPD group compared with the bony decompression group. Back pain was hypothesized to be less in the group that underwent an operation with less tissue damage, namely the IPD without bony decompression group. However, this was not the result that was encountered: the long-term back pain in IPD group was significantly – though not clinically relevant – higher compared with the conventional bony decompression group.

The recently published randomized trial comparing wide laminectomy combined with posterior and intercorporal fusion, to bony decompression with IPD showed

comparable back pain in both groups (104 week VAS back pain of 27mm and 24mm respectively)(Table 4).[14]. In addition, it had already been demonstrated in a non-randomized study that patients suffering from INC treated with bony decompression and adjuvant IPD placement (to maintain posterior dynamic stabilization) had the same long-term VAS back-pain as patients treated with bony decompression alone (Table 4).[57;58] In the present study, treatment with IPD without bony decompression (operation with less tissue damage) did not result in less back-pain as well.

Table 4. literature comparison of VAS back-pain

Variable	baseline	104 weeks
	mean mm VAS (SD)	mean mm VAS (SD)
Primary outcome		
IPD without bony decompression Moojen et al.	60 (44) n=79	36 (23) n=72
Bony decompression Moojen et al.	49 (25) n=80	28 (25) n=76
Bony decompression Richter et al.	60* n=30	30* *
IPD with bony decompression Richter et al.	60* n=30	30* *
IPD with bony decompression Davis et al.	80 (15) n=215	24(26) n=162
Bony decompression with fixation Davis et al.	79(14) n=106	27(29) n=86

N stands for number of patients analyzed. SD stands for standard deviation from means. VAS visual-analogue scale. Dashes denote tests not administrated. * no precise values available in abstracts.

The first interspinous device was designed to damp the motion of extension.[17;18] A few years later, implants were hypothesized to achieve indirect decompression.[59] In theory, both properties should lead to less back and leg pain. Furthermore, devices were also designed with more rigidity to achieve a long-lasting effect.[17] In the current study, the VAS leg-pain was comparable in both groups, even after long-term (two years) follow-up. In both groups all success rates (MRDQ, Likert and ZCQ) stabilized, or even increased in the second year of follow-up, without fixation techniques. Indirect decompression with stand-alone device can be achieved and with long-lasting effect. However, the number of reoperations in the IPD treatment arm is worrisome. Especially because re-operated patients do not reach the success rate of primary surgeries, it is suggested that use of IPD prevents recovery in 20% of the patients.[31]

One of the strengths of this study is that this is the first blinded randomized study on this subject. Furthermore, due to blinding of patient information during data-analysis, we excluded as much as possible bias. However, the present study has also features that may limit the generalizability of its findings. First, selection bias could have been introduced by the opinion of the including neurosurgeon that patients with severe spinal stenosis on the MRI should not be offered an IPD and were thus not included in the Felix Trial. However, clinical features of the patients included in this study demonstrated baseline values (mean VAS (leg and/or back) of 60 mm at baseline) comparable to those of other large trials.[8;60] Another limitation might be the fact that, because of lack of power, a difference was not found that might exist. As the 2-year results do not show a significant difference, one cannot say that the outcomes were similar or equal. The intention of this study was to find evidence to present strong superiority in favour of IPD, to create arguments to reimburse the expensive implants. This evidence in favour of IPD, however, was not found and the investigators did not find any suggestion in the data that a larger sample size would lead to a different study result. To the contrary the higher re-operation rate and the higher intensity of LBP in the IPD group do suggest inferiority compared to classical decompression.

Conclusion

This double blinded study could not confirm advantage of IPD without body decompression over conventional 'simple' decompression. Since the introduction thirty years ago there is a lack of proof of the superiority of this expensive implants in the treatment of LSS as a stand-alone decompressive device.

Reference List

1. EVANS JG. Neurogenic Intermittent Claudication. *Br Med J* 1964; 2(5415):985-987.
2. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. *Osteoarthritis Cartilage* 2012; 20(10):1103-1108.
3. Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. *Ned Tijdschr Geneeskd* 1950; 94(33):2415-2433.
4. Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. *Spine (Phila Pa 1976)* 1991; 16(6):615-619.
5. Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. *Acta Orthop Scand Suppl* 1993; 251:67-68.
6. Verbiest H. [Further reports on primary stenosis of the lumbar spinal canal in adults.]. *Ned Tijdschr Geneeskd* 1951; 95(27):1965-1970.
7. Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. *J Bone Joint Surg Br* 1954; 36-B(2):230-237.
8. Malmivaara A, Slatis P, Heliovaara M, Sainio P, Kinnunen H, Kankare J et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. *Spine (Phila Pa 1976)* 2007; 32(1):1-8.
9. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med* 2007; 356(22):2257-2270.
10. Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. *Spine (Phila Pa 1976)* 1992; 17(1):1-8.
11. Airaksinen O, Herno A, Kaukanen E, Saari T, Sihvonen T, Suomalainen O. Density of lumbar muscles 4 years after decompressive spinal surgery. *Eur Spine J* 1996; 5(3):193-197.
12. Hanraets PR. [Radicular syndrome and low-back pain]. *Folia Psychiatr Neurol Neurochir Neerl* 1953; 56(1):3-20.
13. Thome C, Zevgaridis D, Leheta O, Bazner H, Pockler-Schoniger C, Wohrle J et al. Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. *J Neurosurg Spine* 2005; 3(2):129-141.
14. Davis RJ, Errico TJ, Bae H, Auerbach JD. Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective, randomized, multicenter, Food and Drug Administration Investigational Device Exemption trial. *Spine (Phila Pa 1976)* 2013; 38(18):1529-1539.
15. Fox MW, Onofrio BM, Onofrio BM, Hanssen AD. Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. *J Neurosurg* 1996; 85(5):793-802.
16. Fox MW, Onofrio BM. Indications for fusion following decompression for lumbar spinal stenosis. *Neurosurg Focus* 1997; 3(2):e2.
17 Senegas J. [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. *Acta Orthop Belg* 1991; 57 Suppl 1:221-226.
18. Senegas J. Mechanical supplementation by non-rigid fixation in degenerative intervertebral lumbar segments: the Wallis system. *Eur Spine J* 2002; 11 Suppl 2:S164-S169.
19. Chiu JC. Interspinous process decompression (IPD) system (X-STOP) for the treatment of lumbar spinal stenosis. *Surg Technol Int* 2006; 15:265-275.
20. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC. Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. *Eur Spine J* 2011; 20(10):1596-1606.

21. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 2004; 13(1):22-31.
22. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine (Phila Pa 1976)* 2005; 30(12):1351-1358.
23. Anderson PA, Tribus CB, Kitchel SH. Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. *J Neurosurg Spine* 2006; 4(6):463-471.
24. Eichholz KM, Fessler RG. Is the X STOP interspinous implant a safe and effective treatment for neurogenic intermittent claudication? *Nat Clin Pract Neurol* 2006; 2(1):22-23.
25. Lee J, Hida K, Seki T, Iwasaki Y, Minoru A. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J Spinal Disord Tech* 2004; 17(1):72-77.
26. Mariottini A, Pieri S, Giachi S, Carangelo B, Zalaffi A, Muzii FV et al. Preliminary results of a soft novel lumbar intervertebral prothesis (DIAM) in the degenerative spinal pathology. *Acta Neurochir Suppl* 2005; 92:129-131.
27. Oppenheimer JH, DeCastro I, McDonnell DE. Minimally invasive spine technology and minimally invasive spine surgery: a historical review. *Neurosurg Focus* 2009; 27(3):E9.
28. Siddiqui M, Karadimas E, Nicol M, Smith FW, Wardlaw D. Effects of X-STOP device on sagittal lumbar spine kinematics in spinal stenosis. *J Spinal Disord Tech* 2006; 19(5):328-333.
29. Sobottke R, Schluter-Brust K, Kaulhausen T, Rollinghoff M, Joswig B, Stutzer H et al. Interspinous implants (X Stop, Wallis, Diam) for the treatment of LSS: is there a correlation between radiological parameters and clinical outcome? *Eur Spine J* 2009; 18(10):1494-1503.
30. Tsai KJ, Murakami H, Lowery GL, Hutton WC. A biomechanical evaluation of an interspinous device (Coflex) used to stabilize the lumbar spine. *J Surg Orthop Adv* 2006; 15(3):167-172.
31. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW et al. Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. *BMJ* 2013; 347:f6415.
32. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. *BMC Musculoskelet Disord* 2010; 11:100.
33. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2002; 27(1):84-91.
34. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995; 48(11):1369-1378.
35. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996; 21(7):796-803.
36. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. *Spine (Phila Pa 1976)* 2006; 31(11):1276-1280.
37. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. *Spine (Phila Pa 1976)* 1986; 11(1):28-30.
38. Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. *Med Care* 1995; 33(4 Suppl):AS286-AS291.
39. Hutchinson PJ, Laing RJ, Waran V, Hutchinson E, Hollingworth W. Assessing outcome in lumbar disc surgery using patient completed measures. *Br J Neurosurg* 2000; 14(3):195-199.
40. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006; 332(7555):1430-1434.
41. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27(3 Suppl):S217-S232.

42. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995; 20(17):1899-1908.
43. Roland M, Morris R. A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care. *Spine (Phila Pa 1976)* 1983; 8(2):145-150.
44. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)* 1983; 8(2):141-144.
45. Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther* 1996; 76(3):268-275.
46. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983; 16(1):87-101.
47. Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992; 305(6846):160-164.
48. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994; 32(1):40-66.
49. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975; 1(3):277-299.
50. Melzack R. The McGill pain questionnaire: from description to measurement. *Anesthesiology* 2005; 103(1):199-202.
51. Main CJ, Wood PL, Hollis S, Spanswick CC, Waddell G. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine (Phila Pa 1976)* 1992; 17(1):42-52.
52. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002; 52(2):69-77.
53. Rabe-Jablonska J, Bienkiewicz W. [Anxiety disorders in the fourth edition of the classification of mental disorders prepared by the American Psychiatric Association: diagnostic and statistical manual of mental disorders (DMS-IV -- options book]. *Psychiatr Pol* 1994; 28(2):255-268.
54. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47(12):1019-1024.
55. Arts MP, Brand R, van den Akker ME, Koes BW, Bartels RH, Peul WC. Tubular diskectomy vs conventional microdiskectomy for sciatica: a randomized controlled trial. *JAMA* 2009; 302(2):149-158.
56. Peul WC, Van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 2007; 356(22):2245-2256.
57. Richter A, Schutz C, Hauck M, Halm H. Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. *Eur Spine J* 2009.
58. Richter A, Halm HF, Hauck M, Quante M. 2-year Follow-up After Decompressive Surgery With and Without Implantation of an Interspinous Device for Lumbar Spinal Stenosis: A Prospective Controlled Study. *J Spinal Disord Tech* 2012.
59. Stromqvist B, Berg S, Gerdhem P, Johnsson R, Sahlstrand T, Tullberg T. Indirect Decompression (X-Stop) versus Conventional Decompressive Surgery for Lumbar Spinal Claudication - A Prospective Randomized Trial. *The Spine Journal* 7 A.D.; 7(5):70.
60. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008; 358(8):794-810.

Chapter 7

The Felix Trial: cost-utility analysis.

Extracted from Spine J. 2014. Part of the NASS Most Valuable Paper 2014.

Moojen WA (A1/A2) & Van den Akker – Van Marle ME (A3)
Arts MP (A2)
Vleggeert-Lankamp CLAM (A1)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Medical Decision Making, Leiden University Medical Center

Abstract

Background: In the 1980s a new implant was developed to treat patients with intermittent neurogenic claudication (INC) caused by lumbar spinal stenosis (LSS). This implant is now widely used.

Purpose: The objective of this study is to determine whether a favourable cost-effectiveness for interspinous process devices (IPD) compared with conventional bony decompression is attained.

Study design/setting: Cost-utility analysis was performed alongside a double-blind randomized controlled trial. 5 neurosurgical centers (including one academic and four secondary level care centers) included participants for this study.

Patient sample: 159 patients with lumbar spinal stenosis were treated with the implantation of IPD and with bony decompression. 80 participants received an IPD and 79 participants underwent spinal bony decompression.

Outcome measures: Outcome measures were quality-adjusted life-years and societal costs in the first year (estimated per quarter), estimated from patient reported utilities (US and Netherlands EuroQol 5D (EQ 5D) and EuroQol Visual Analogue Scale (EQ VAS)) and diaries on costs (health care costs, patient costs, and productivity costs).

Methods: All analyses followed the intention-to-treat principle. Given the statistical uncertainty of differences between costs and QALYs, cost effectiveness acceptability curves graph the probability that a strategy is cost effective, as a function of willingness to pay. Paradigm Spine funded this trial, but did not had any part in data-analysis or the design and preparation of this manuscript.

Results: According to the EQ-5D, the valuation of quality of life after IPD and decompression was not different. Mean utilities during all four quarters were – not significantly – less favourable after IPD according to the EQ-5D with a decrease in QALYs according to the US EQ 5D of 0.024 (95% confidence interval -0.031 to 0.079). From a health care perspective the costs of IPD treatment were higher (difference €3,030 per patient, 95% interval €561 to €5,498). This significant difference is mainly due to additional cost of implants of €2,350 apiece. From a societal perspective, a non-significant difference of €2,762 (95% confidence interval -€1,572 to €7,095) in favour of conventional bony decompression was found.

Conclusions: Implantation of IPD as indirect decompressing device is highly unlikely to be cost-effective compared with bony decompression for patients with intermittent neurogenic claudication caused by lumbar spinal stenosis.

Background

The average increase in age of the general population results in a growing older population, and thus to an increase in incidence of patients with intermittent neurogenic claudication (INC) caused by lumbar spinal stenosis (LSS).[1-3] INC is a complex of symptoms including pain in – frequently – both legs provoked by prolonged walking and standing and diminishes by flexion of the lumbar spine (such as in sitting position or when cycling). [1-3] Accompanying back pain is also associated with INC.[4] The number of surgical interventions for lumbar stenosis increases– concomitantly with the increase in age of the general population – , and eventually this can lead to an increase of the use of implants.[5-7] In the beginning of the 21ste century, a 15-fold increase of spinal surgeries with fusion techniques was reported.[5-7] One of the possible explanations for this dramatic increase of fusion procedures was the development of new devices in the end of the 20th century.

In 1984 an implant to indirectly decompress the lumbar spinal canal was developed.[8;9] The implant is placed between the spinous processes, and is therefore called interspinous process device (IPD). The operation time was proposed to be shorter with less bony destruction and the technique was meant to accustom day surgery protocols, resulting in a shorter rehabilitation period after surgery. The implant was believed to be ideal for the old and even octogeneric patients with lumbar spinal stenosis. Despite the high costs of the implants and the high rate of implantations of IPDs, clinical trials comparing IPD's with the golden standard (bony decompression) were not performed.[10;11] The scientific evidence published until 2004 showed that the use of IPDs was superior compared – only – with conservative (no intervention) treatment. [12;13]

A lot of different interspinous process devices (IPDs) have been introduced since 1984.[14-17] Since 1984, no good estimation of the total costs for society of these IPDs has been performed. In the systematic review published in 2011, at least 20 different IPDs were identified.[18] The two most studied implants are the Coflex implant (Paradigm Spine) and the X-stop (Kyphon inc and Medtronic).[12-17;19-29] The Coflex implant – like the other IPDs – was used as stand-alone and subjected to our protocol as such. Currently the Coflex is in the US only approved for add-on to decompression. However, little is known about the costs of these regularly used implants. The 2007, worldwide sale of the X-stop implant was 80 million USD.[30] In 2009, 13,128 Coflex implants were used.[31] Usually they are sold for - at least 2,500

USD, resulting in a total worldwide sale of at least 30 million USD in 2009. Other financial information was not accessible for the authors.

In our clinical study we have shown that the use of interspinous implants compared to conventional decompression did not result in the hypothesized better outcome and not even in shorter recovery times. Moreover, it was demonstrated that the reoperation rate was significantly higher in the patients that had received an IPD.[32] The clinical outcome at 8 weeks and 1 year after surgery was comparable in both patient groups. The higher costs of IPDs in combination with the comparable results after one year and the higher re-operation rate do not seem to support the use of IPDs from an economical point of view. However, quality of life estimates and a complete cost overview could lead to another conclusion. In the present study we perform the cost-utility study alongside a clinical trial on this subject, to assess whether the considerable costs on IPDs can be justified from an economical viewpoint.

Methods

Patients and treatment

Patients between 40 and 85 years with at least three months of intermittent neurogenic claudication (INC) due to single or two level, degenerative lumber canal stenosis and an indication for surgery participated in a multicenter double blinded randomised controlled trial that compared treatment with IPD versus standard conventional surgical decompression.[33] Patients with INC based on LSS at L5-S1 were excluded. Institutional Ethical Review Boards of participating hospitals approved the research protocol and participants gave written informed consent. A total sample size of 160 was calculated, based on a minimal clinically important difference of 20% change in the overall success rate according to the Zurich Claudication Questionnaire (ZCQ) between IPD and standard bony decompression in favor of IPD.[34-36]

Between October 2008 and September 2011, 162 patients with INC due to spinal stenosis were enrolled and 159 patients were analysed. A randomized design with variable block sizes was used, with allocations stratified according to center. Allocations were stored in prepared opaque, coded and sealed envelopes. The key was only accessible to the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center.

All patients gave informed consent. After induction of anesthesia the prepared envelope was opened and randomized allocation to one of the treatment arms was performed. 80 patients were allocated to IPD and 79 decompression surgery alone. Three patients were excluded after randomization; one patient died while waiting for the operation and two patients had a severe spondylolysis at the final preoperative check-up and were excluded from the study, because this could cause a detrimental effect in the implant group. Both groups were comparable regarding their baseline characteristics.[32]

Patients with a cauda equina syndrome, a herniated disc needing discectomy, history of lumbar surgery, or significant scoliosis (Cobb angle >25°) or other spinal deformities were excluded. The details of treatment can be found elsewhere.[33] Briefly, in patients in the IPD group an interspinous process device was implanted by a posterior midline approach using radiographic examination for localization of the appropriate level. No bony decompression was performed. In patients in the decompression group a partial resection of the adjacent laminas was executed, followed by a flavectomy with bilateral opening of the lateral recess. Patients and research nurses were kept blinded to the allocated treatment throughout the complete follow-up period of one year.

Utilities and QALYs

Utilities represent the valuation of the quality of life of the patients on a scale from zero (as bad as death) to one (perfect health). Patients described their quality of life using the EuroQol classification system (EQ-5D) from which we calculated utilities for the United States and the Netherlands.[37;38] The EQ-5D provides societal valuation, which is preferred for economic evaluations from a societal perspective. In addition, we obtained valuations by the patients themselves, using a visual analogue scale (EQ VAS) ranging from zero (worst imaginable health) to 100 (perfect health). We transformed these valuations to a utility scale, using the power transformation 1-(1-VAS/100)$^{1.61}$.[39] We obtained measurements for EQ-5D and the VAS at randomisation, and 2, 4, 8, 12, 26 and 52 weeks after randomization. For the EQ-5D and VAS measurements, respectively 3.4% and 3.2% of the items were missing. From the area under the utility curves we calculated the average utility during each separate quarter of the year after randomisation and during the entire year (QALYs).

Costs

We estimated the costs from the societal perspective during the one year of follow-up. Because of this one-year time horizon, costs were not discounted. Costs were converted to 2013 price level using the general Dutch consumer price index.[40] Using cost diaries, patients reported admissions to hospital, visits (specialists, general practitioner and physical therapy), home care, paid domestic help, informal care, drugs and aids, out of pocket expenses as result of intermittent neurogenic claudication, and hours of absenteeism from work. At the follow-up at 4, 8, 26 and 52 weeks after randomization the research nurse went through the diary with the patient. For these follow up moments 0%, 1.9%, 2.5% and 2.5% of the diaries were missing, respectively.

In the costs of surgery we account for costs of depreciation and sterilisation of surgical instruments, costs of the operation room (including personnel) based on the duration of the surgery, and costs of the implant(s) in the IPD group, as these are the costs that differ between both procedures. Costs of use and sterilization of surgical instruments amount to €84.30 and €39.52 for IPD and decompression respectively. Costs of the operation room per minute of use were based on the average costs in 6 Dutch hospitals, resulting in an estimate of €18.74 per minute.[41] Costs of the implant are €2350 apiece; in some patients two implants were used. For other health care costs, we used Dutch standard prices, designed to represent societal costs and to standardize economic evaluations.[42;43] Health care costs are reported including the patients' time and travel costs. We valued the reported hours of absenteeism from work during the one year follow-up period according to the friction cost method using a friction period of 22 weeks, at standard productivity costs of € 28 per hour for women and €35 per hour for men, corrected for elasticity of labour time of 0.8 indicating that a reduction of labour time causes a less than proportional decrease in labour productivity. [42;43]

Analysis

All analyses followed the intention-to-treat principle. Statistical analyses were conducted with Stata 11.2 (StataCorp, College Station, TX, USA). To reduce possible bias due to missing data, we used multiple imputation by chained equations, with 10 iterations for the switching regression model.[44] For each missing utility item or cost item, an imputation regression model was used that included the patient's age, sex, randomisation group, and all (other) utility and cost measures at all (other) moments. Group differences in QALYs and costs were statistically analysed using

standard unequal-variance t-tests. Base case cost-utility analysis compared one-year societal costs to one-year QALYs based on the US EQ-5D. Sensitivity analyses were performed on the use of different utility measures (NL EQ-5D or EQ VAS), and on the perspective (societal or healthcare perspective). Depending on the willingness to pay for obtained effectiveness, a strategy is cost-effective compared with an alternative strategy if it has a better average net benefit (willingness to pay* QALYs – costs). Given the statistical uncertainty of differences between costs and QALYs, cost effectiveness acceptability curves graph the probability that a strategy is cost effective, as a function of willingness to pay.

Results

Utilities and QALYs

According to the EQ-5D, the valuation of quality of life after decompression was equal or better than after IPD (figure 1). Mean utilities during all four quarters were consistently less favourable after IPD (table 1). The difference in QALYs according to the US EQ-5D was 0.024 (95% confidence interval - 0.031 to 0.079), for the NL EQ-5D 0.032 (- 0.036 to 0.100). Thus, the difference in QALYs in the first year after treatment between both groups was not significant.

According to the EQ-VAS only in the first quarter and last quarter of the first year after treatment the quality of life was valued higher in decompression patients, while in the second and third quarter the IPD patients had a higher quality of life. QALYs based on the EQ VAS show a mixed picture during the first year, resulting in a non-significant difference of 0.010 (–0.047 to 0.067) in favour of decompression.

Figure 1. Utilities according to US EQ-5D, NL EQ-5D, and EQ VAS

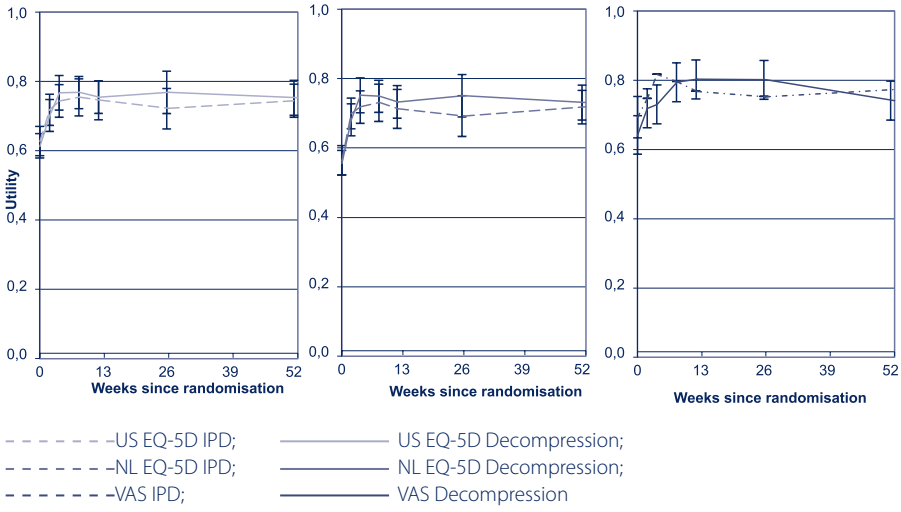


Figure 2. Cost-effectiveness acceptability curves for IPD compared to decompression

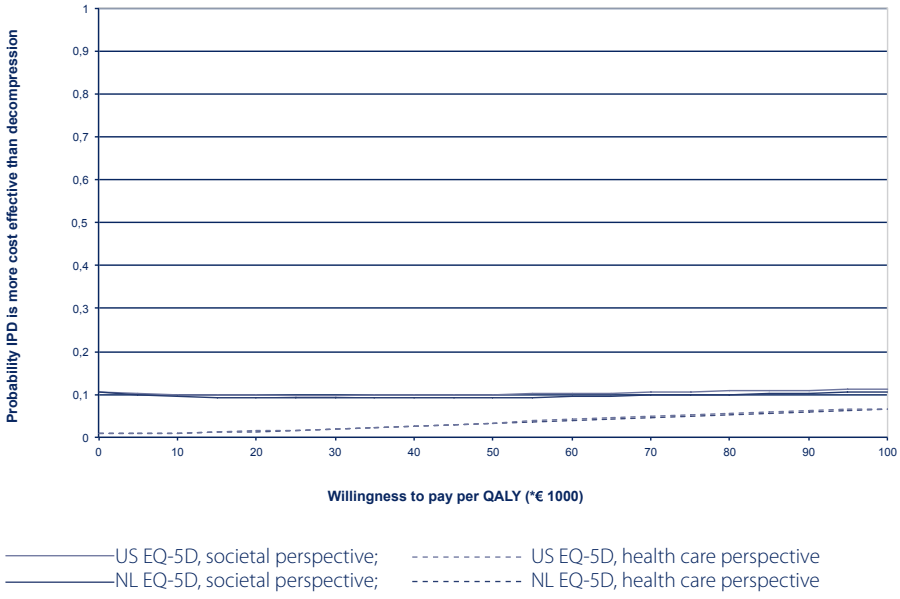


Table 1. Utility and quality of life years (QALYs) after IPD and decompression. Values are means (standard deviations)

Measure	IPD (N=80)	Decompression (N=79)	Difference	P value*
US EQ-5D				
1st quarter	0.731	0.739	-0.008	
2nd quarter	0.729	0.759	-0.030	
3rd quarter	0.724	0.762	-0.038	
4th quarter	0.735	0.754	-0.018	
QALYs	0.730(0.183)	0.753(0.166)	-0.024	0.40
NL EQ-5D				
1st quarter	0.702	0.717	-0.015	
2nd quarter	0.697	0.738	-0.041	
3rd quarter	0.694	0.742	-0.048	
4th quarter	0.707	0.732	-0.025	
QALYs	0.700(0.227)	0.732(0.205)	-0.032	0.35
EQ VAS				
1st quarter	0.776	0.755	0.021	
2nd quarter	0.756	0.799	-0.043	
3rd quarter	0.755	0.783	-0.029	
4th quarter	0.765	0.753	0.012	
QALYs	0.763(0.192)	0.773(0.164)	-0.010	0.74

Healthcare costs

The average costs of IPD treatment including hospital stay in comparison to the decompression were €2,302 higher per patient (95% confidence interval €1,857 to €2,747). This significant difference is mainly due to additional cost of implants of €2,350 apiece. In the first year after treatment 21% of the IPD patients underwent resurgery in comparison with only 6% of the patients in the decompression group, resulting in (non- significant) higher costs of resurgery for IPD patients of €198 per patient (-€112 to €507). In total, mean health care costs per patient in the first year are €10,210 for IPD and €7,180 for decompression patients; the difference of € 3,030 is significant (95% confidence interval €561 to €5,498).

Societal costs

The productivity costs in this population with older patients, did not differ significantly between both study groups (respectively €3,171 for IPD and € 3,374 for decompression). Also the total non-healthcare costs, which are mainly determined by the productivity costs, were not significantly different: the total non-healthcare costs were on average €268 (95% confidence interval -€2,589 to €3,125) lower for IPD

patients. Combining the total healthcare and non-healthcare costs results in total societal costs of €13,858 per IPD patient and €11,096 per decompression patient in the first year after treatment. The difference of €2,762 has a 95% confidence interval of -€1,572 to €7,095, and is therefore not significant.

Cost-utility analysis

The combination of (non-significant) higher societal costs and less favourable QALY outcomes (not significant) after IPD result in a small probability that IPD is more cost-effective compared to decompression (figure 2). For any value of the willingness-to-pay for a QALY the probability that IPD is more cost-effective than decompression is far below 50%. These results hold for all outcome measures (US EQ-5D, NL EQ-5D and VAS). From a healthcare perspective in which IPDs are significantly more costly, the superiority of decompression is even more clear . This leads to the overall result that simple decompression is more cost-effective than IPD.

Discussion

The Felix trial did not show advantage of IPD treatment as indirect decompressing device over conventional decompression.[32;33] In addition the economic evaluation showed less favourable QALY outcomes (not significant) after IPD. Healthcare costs were higher for IPD treatment as indirect decompressing device compared to decompression. From a societal perspective the cost of IPD treatment were also higher, though not significant. In conclusion, implantation of IPD as indirect decompressing device is highly unlikely to be cost-effective compared with bony decompression for patients with intermittent neurogenic claudication caused by lumbar spinal stenosis.

The higher costs of IPD treatment were mainly due to the costs of initial surgery. A major component of the costs of IPD treatment is the cost of implants of €2,350. In 27% of the patients in the IPD group even two implants were used since there were two symptomatic levels to be treated. The operating time in the IPD group was a bit shorter (24 vs 43 minutes). However, this did not result in a significantly shorter overall time in the OR (e.g. including anesthesia and preparations time). Furthermore, the substantially higher repeat surgery rate, leaded to higher costs in the IPD group compared to the conventional group.[32] However, the difference in costs of reoperation per patient was not significant, due to the fact that the

Table 2. Mean healthcare costs and societal cost per patient .

Cost item	IPD (N= 80)		BD (N=79)		Difference	
	Volume	Costs (€)	Volume	Costs (€)	Costs(€)	P-value*
Initial treatment						
- Treatment	100	3327	100	831	2495	
- Hospital stay¶	2.8	1386	3.2	1580	-194	
Total (SD)	100	4713	100	2411	2302	<0.01
Repeated surgery (1 yr)#	21	447	6	250	198	0.21
Physical therapy	91	2431	92	2178	253	0.71
Other hospital admissions	22	868	6	300	569	0.14
Neurologist†	0.9	108	0.6	70	38	0.30
Neurosurgeon†	1.4	164	0.9	100	64	0.06
Orthopaedist†	0.7	82	0.4	42	40	0.19
Other specialists	55	250	59	236	14	0.85
Imaging	68	286	57	245	41	0.41
Pain relief	11	24	11	24	1	0.96
General practitioner†	4.0	173	3.9	166	7	0.86
Home care‡	18	522	31	1000	-477	0.27
Drugs	62	65	70	85	-20	0.55
Aids	25	76	26	77	-2	0.98
Total healthcare costs						
1st quarter		5963		3779	2184	
2nd quarter		2120		1660	460	
3rd quarter		1063		871	193	
4th quarter		1063		871	193	
Total(SD)		10210 (7583)		7180 (7717)	3030	0.02
Paid domestic help‡	13	172	9.4	128	44	0.62
Informal care‡	20	276	30	406	-130	0.53
Out of pocket expenses	6.5	30	4.6	8	21	0.33
Productivity costs (friction)						
1st quarter ‡	31	1339	39	1665	-326	
2nd quarter ‡	30	1294	22	944	350	
3rd quarter ‡	6.5	279	11	475	-196	
4th quarter ‡	6.0	259	6.8	290	31	
Total (SD)	74	3171 (8365)	79	3374 (8678)	-203	0.88
Total non-healthcare costs (SD)		3648 (8919)		3916 (8938)	-268	0.85
Total societal costs						
1st quarter		7437		5676	1760	
2nd quarter		3522		2700	822	
3rd quarter		1460		1453	7	
4th quarter		1439		1268	172	
Total (SD)		13858 (14312)		11096 (12527)	2762	0.21

* t test for unequal variance, correcting for non-response using multiple imputation; ¶ Number of hospital days; † Number of visits; ‡ Number of hours; # percentage of patients actually operated on within twelve months after randomization. BD bony decompression

costs of reoperation were higher for decompression patients and the variance in costs between patients. The operation time of simply removing an implant and performing a standard bony decompression is much shorter compared to a reoperation in a surgical field with scar tissue. The total health care costs in the first year of €10,210 per patient were significantly higher for IPD group compared to the decompression group (€7,180).

Mean productivity costs were low in this study (€ 3,171 per patient in the IPD group and €3,374 in the decompression group), due to the fact that the majority of patients were retired: 71% of patients in the decompression group and 78% of patients in the IPD group did not report productivity costs. Due to the large variance (a minority of patients having productivity costs) the difference of € 203 was not significant. As a result, also the non-healthcare costs in the first year, mainly consisting of productivity costs, were not significantly different in favour of the IPD group (€3,648 per patient versus €3,916 per patient in the decompression group). Taking the healthcare and non-healthcare costs together in the societal costs resulted in non-significant higher costs for IPD patients of €13,858 compared to €11,096 in decompression patients. The combination of (non-significant) higher costs and less favourable QALY outcomes (not significant) after IPD resulted in a small probability that IPD is more cost-effective compared to decompression.

As it is highly unlikely that IPD as indirect decompressing device is more cost-effective than decompression, the high costs for society of these implants in the recent years could not be justified based on the current clinical study. Implants were already widely used for more than twenty years throughout the world for different indications resulting in high worldwide sales. Cost utility analysis of these implants was not performed earlier. The recently published FDA Trial, in which IPDs in combination with bony decompression were more effective compared to bony decompression in combination with spinal fusion, did also not include a societal cost-analysis.[45] In the future, new products and procedures should be subjected to rigid scientific comparison with current gold-standard treatments before introduction.

This study has several limitations. Firstly, we limited the duration of the economic evaluation to one year. However, differences in utility values and costs between the groups decreased in the last quarters, indicating that the most important effects are included in the one year study period. Furthermore, the generalizability of our results may be difficult due to the high amount of different implants and different selling prices of these implants. Using lower prices of the implants did however not

alter the results. Even if we assumed no costs for the implants, the probability that IPD as indirect decompressing device was more cost-effective than decompression was lower than 50% for values of the willingness to pay commonly accepted in The Netherlands (€20,000-€40,000). This can be explained by the fact that IPD did not result in a higher quality of life and also not to significant savings in other costs. In the future, new products and procedures should be subjected to rigid scientific comparison with current gold-standard treatments before introduction. As mentioned earlier, selection bias could have been introduced through the opinion of the including neurosurgeon that patients with relatively high co morbidity should not be offered an IPD. Perhaps, the supposed shorter (hospital) recovery and related costs could have been more clear in this study if we had included older patients or patients with more co-morbidity in a study design under local anesthesia. However, blinding of the patient would not be possible. Therefore we decided to perform a robust double blind comparison.[33]

Conclusion

Implantation of IPDs as indirect decompressing device leads to higher healthcare costs and do not improve quality of life after treatment compared with standard bony decompression. Therefore, implantation of IPD as indirect decompressing device is highly unlikely to be cost-effective compared with bony decompression. The use of these implants over the past three decades could therefore not be justified based on this study.

Reference List

1. EVANS JG. Neurogenic Intermittent Claudication. *Br Med J* 1964; 2(5415):985-987.
2. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. *Osteoarthritis Cartilage* 2012; 20(10):1103-1108.
3. Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. *Ned Tijdschr Geneeskd* 1950; 94(33):2415-2433.
4. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008; 358(8):794-810.
5. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. *N Engl J Med* 2004; 350(7):722-726.
6. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. *Spine (Phila Pa 1976)* 2005; 30(12):1441-1445.
7. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *JAMA* 2010; 303(13):1259-1265.
8. Senegas J. [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. *Acta Orthop Belg* 1991; 57 Suppl 1:221-226.
9. Senegas J. Mechanical supplementation by non-rigid fixation in degenerative intervertebral lumbar segments: the Wallis system. *Eur Spine J* 2002; 11 Suppl 2:S164-S169.
10. Bartels RH. The X STOP device. *J Neurosurg Spine* 2007; 6(6):620-621.
11. Eichholz KM, Fessler RG. Is the X STOP interspinous implant a safe and effective treatment for neurogenic intermittent claudication? *Nat Clin Pract Neurol* 2006; 2(1):22-23.
12. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 2004; 13(1):22-31.
13. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine (Phila Pa 1976)* 2005; 30(12):1351-1358.
14. Hrabalek L, Machac J, Vaverka M. [The DIAM spinal stabilisation system to treat degenerative disease of the lumbosacral spine]. *Acta Chir Orthop Traumatol Cech* 2009; 76(5):417-423.
15. Park SC, Yoon SH, Hong YP, Kim KJ, Chung SK, Kim HJ. Minimum 2-year follow-up result of degenerative spinal stenosis treated with interspinous u (coflex). *J Korean Neurosurg Soc* 2009; 46(4):292-299.
16. Richter A, Schutz C, Hauck M, Halm H. Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. *Eur Spine J* 2009.
17. Zucherman J, Hsu K, Wahlig J, Hartjen C, Shabe P, Tillman J. P108. Five Year Outcomes in Patients Treated with the X-STOP Interspinous Process Device for Neurogenic Intermittent Claudication Due to Lumbar Spinal Stenosis. *The Spine Journal* 2009; 8(5, Supplement 1):153S.
18. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC. Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. *Eur Spine J* 2011; 20(10):1596-1606.
19. Stromqvist BH, Berg S, Gerdhem P, Johnsson R, Moller A, Sahlstrand T et al. X-Stop Versus Decompressive Surgery for Lumbar Neurogenic Intermittent Claudication: Randomized Controlled Trial With 2-Year Follow-up. *Spine (Phila Pa 1976)* 2013; 38(17):1436-1442.
20. Anderson PA, Tribus CB, Kitchel SH. Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. *J Neurosurg Spine* 2006; 4(6):463-471.
21. Bhadra AK, Raman AS, Tucker S, Noordeen HH. Interspinous implant in lumbar spinal stenosis: A prospective cohort. *European Journal of Orthopaedic Surgery and Traumatology* 2008; 18(7):489-493.
22. Brussee P, Hauth J, Donk RD, Verbeek AL, Bartels RH. Self-rated evaluation of outcome of the implantation of interspinous process distraction (X-Stop) for neurogenic claudication. *Eur Spine J* 2008; 17(2):200-203.

23. Galarza M, Fabrizi AP, Maina R, Gazzeri R, Martinez-Lage JF. Degenerative lumbar spinal stenosis with neurogenic intermittent claudication and treatment with the Aperius PercLID System: a preliminary report. *Neurosurgical Focus* 2010; 28(6).
24. Kuchta J, Sobottke R, Eysel P, Simons P. Two-year results of interspinous spacer (X-Stop) implantation in 175 patients with neurologic intermittent claudication due to lumbar spinal stenosis. *Eur Spine J* 2009; 18(6):823-829.
25. Lee J, Hida K, Seki T, Iwasaki Y, Minoru A. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J Spinal Disord Tech* 2004; 17(1):72-77.
26. Richter A, Schutz C, Hauck M, Halm H. Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. *Eur Spine J* 2009.
27. Siddiqui M, Smith FW, Wardlaw D. One-year results of X Stop interspinous implant for the treatment of lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2007; 32(12):1345-1348.
28. Yano S, Hida K, Seki T, Aoyama T, Akino M, Iwasaki Y. A new ceramic interspinous process spacer for lumbar spinal canal stenosis. *Neurosurgery* 2008; 63(1 Suppl 1):ONS108-ONS113.
29. Park SC, Yoon SH, Hong YP, Kim KJ, Chung SK, Kim HJ. Minimum 2-Year Follow-Up Result of Degenerative Spinal Stenosis Treated with Interspinous U (Coflex (TM)). *Journal of Korean Neurosurgical Society* 2009; 46(4):292-299.
30. Sale numbers directly obtained from Kyphon inc sales report 2007. 2014. Ref Type: Generic
31. Sale numbers directly obtained from Kyphon Inc and Paradigm Spine Inc. 2014. Ref Type: Unpublished Work
32. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW et al. Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. *BMJ* 2013; 347:f6415.
33. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. *BMC Musculoskelet Disord* 2010; 11:100.
34. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2002; 27(1):84-91.
35. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995; 48(11):1369-1378.
36. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996; 21(7):796-803.
37. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43:203-220.
38. Lamers LM, Stalmeier PF, McDonnell J, Krabbe PF, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. *Ned Tijdschr Geneeskd* 2005; 149(28):1574-1578.
39. Stiggelbout AM, Eijkemans MJ, Kiebert GM, Kievit J, Leer JW, De Haes HJ. The 'utility' of the visual analog scale in medical decision making and technology assessment. Is it an alternative to the time trade-off? *Int J Technol Assess Health Care* 1996; 12(2):291-298.
40. Statistics Netherlands. Consumer price index. http://statline.cbs.nl/. 5-12-2013. Ref Type: Online Source
41. van den Akker ME, Arts MP, van den Hout WB, Brand R, Koes BW, Peul WC. Tubular diskectomy vs conventional microdiskectomy for the treatment of lumbar disk-related sciatica: cost utility analysis alongside a double-blind randomized controlled trial. *Neurosurgery* 2011; 69(4):829-835.
42. Tan SS, Bouwmans CA, Rutten FF, Hakkaart-van RL. Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care* 2012; 28(2):152-158.
43. Hakkaart-van RL, Tan SS, Bouwmans CA. Manual for costing research: methods and cost prices for economic evaluation in health care, update 2010 [in Dutch]. 2014. Ref Type: Online Source
44. van BS, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999; 18(6):681-694.
45. Davis RJ, Errico TJ, Bae H, Auerbach JD. Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective, randomized, multicenter, Food and Drug Administration Investigational Device Exemption trial. *Spine (Phila Pa 1976)* 2013; 38(18):1529-1539.

Chapter 8

Preoperative MR imaging in patients with intermittent neurogenic claudication: relevance for diagnosis and prognosis

Submitted May 2014.

Moojen WA (A1/A2)
Schenck CD(A1)
Lycklama à Nijeholt GL (A3)
Jacobs WCH (A1)
Van der Kallen BF (A3)
Arts MP (A2)
Peul WC (A1/A2)
Vleggeert-Lankamp CLAM (A1)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Radiology, Medical Center Haaglanden, The Hague, The Netherlands

Abstract

Background: Magnetic resonance imaging (MRI) is an important tool to confirm the diagnosis of lumbar spinal stenosis (LSS) as a cause for intermittent neurogenic claudication. It is believed that the narrowness of the lumbar spinal canal correlates to the severity of complaints, and that it may be a good predictor of clinical outcome if treated. However, this hypothesis has never been (prospectively) tested.

Methods: We studied baseline MRIs of 155 patients with intermittent neurogenic claudication and lumbar spinal stenosis. MRI and patient data were gathered for a randomized trial comparing bony decompression to implantation of an interspinous process device as surgical treatment options for LSS. Three independent specialized readers were asked to evaluate the MRIs to grade narrowing (Schizas scale) and to judge likelihood of cauda compression on the relevant spinal levels. Additionally, several other stenosis-related characteristics were scored. Symptom severity at baseline and at one-year follow-up were quantified using the Modified Roland Disability Questionnaire (MRDQ), the Visual Analogue Scale (VAS) for leg pain, and VAS for back pain. The radiological scores were correlated with clinical baseline and outcome scores to assess diagnostic and prognostic value of MRI findings at baseline.

Results: There was good agreement on the clinically relevant level of lumbar spinal stenosis (kappa range 0.57 to 0.64). MRI assessment of grading of compression (kappa of agreement ranges 0.33-0.46) did not correlate with baseline MRDQ, nor with outcome based on postoperative change in MRDQ (p-value 0.61). However, both absence of epidural fat and presence of tortuous caudal nerves on MR images (kappa of agreement ranges 0.53-0.72 and 0.67-0.70 respectively) in patients with LSS were relatively good predictors for satisfactory recovery after surgery (p-values 0.03 and <0.01 respectively).

Conclusion: The grading of compression on the preoperative MRI is neither ambiguous nor correlating to severity of clinical condition. It does furthermore not have the ability to predict the outcome after one year if surgically treated. MRI should therefore only be used to appoint the pathological level in case surgical decompression is considered.

Background

Since the early publications of Verbiest et al. in 1950 on spinal stenosis, intraspinal diameter is believed to be one of the most important factors influencing severity of symptoms in patients with intermittent neurogenic claudication (INC) caused by lumbar spinal stenosis (LSS).[1] Most patients with INC have a complex of symptoms, dominated by leg pain (frequently in both legs), which may be exacerbated by walking, prolonged standing or lumbar extension.[1-4] The extent of compression was first measured in vivo in 1950s: small antero-postero diameter (smaller than 12 mm at L5) was considered prone for development of INC.[3] Myelography, later computed tomography (CT) imaging and nowadays magnetic resonance imaging (MRI) are used to assess the presence of and the level at which compression exists.[5] It is generally believed that the relative size of cross-sectional area of the dural sac correlate with clinical symptomatology is not known, though this was never (prospectively) evaluated before.[6-10] Narrowness of the lumbar spinal canal is common in the elderly spine: 20% of subjects in the healthy population (median age 66 years) were found to have severe stenosis on MRI.[6,7,10] However, none of these subjects had any complaints. Furthermore, abnormalities related to degenerative spinal disorders, such as lumbar spinal stenosis to any extent, or herniated discs, are found in up to 64% of MRIs made in asymptomatic populations, indicating a generally weak clinico-radiological correlation.[11]

The clinical short-term and long-term follow-up of a double blind multicenter randomized controlled trial comparing interspinous process devices (IPD) and bony decompression to treat patients with INC caused by LSS was recently published.[12] Patients with lumbar spinal stenosis, and failed conservative treatment were included for this trial, and were therefore ideal subjects to select for imaging studies comparing INC complaints with MRI parameters.[12] Prior to both treatment modalities, baseline MRI was obtained for all patients.[13] This provided us with a large set of MRIs in combination with a standardized evaluation of clinical condition of intermittent neurogenic claudication at baseline and during follow up after surgical intervention. The goal of this study was to assess whether stenosis severity as seen on baseline MRI correlates to clinical condition at baseline and whether MRI parameters correlate to the outcome one year after surgery.

Methods

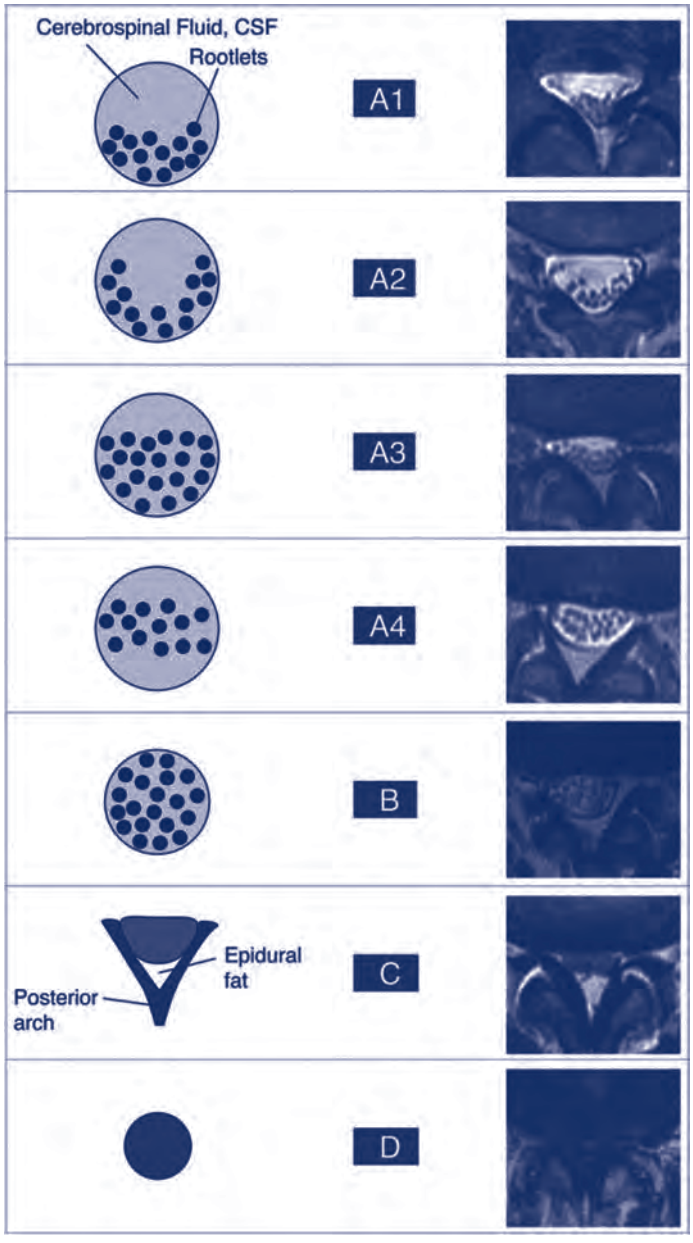
Study population

Patients in this study were participants in the Felix Trial, a double-blind multi center randomized trial among patients with a history of intermittent neurogenic claudication (INC) based on LSS, as seen on MRI. Patients included in the trial were aged 40 to 85 years and were subjected to at least 3 months of conservative therapy. The medical ethics committee at each of the five participating hospitals approved of the Felix Trial protocol (including the use of the MRI for scientific purposes), which is available with the full text of this article at BMJ.com.[12, 13] Written informed consent was obtained from all patients.

MRI protocol and evaluation

All patients included in the trial underwent MRI at time of initial diagnosis. The extent of spinal stenosis was graded according to the grading system for spinal stenosis developed by Schizas et al.[14] The grading system allows for the classification of lumbar spinal stenosis in four grades: grade A, no or minor stenosis; grade B, moderate stenosis; grade C, severe stenosis and grade D, extreme stenosis. Grade A is further divided into four subcategories (Figure 1).[14] In our analysis we used the four category Schizas scale (grades A, B, C and D) and two dichotomized Schizas scales (whereas grade A was scored as 'no to mild stenosis' and grade B, C and D were scored as 'stenosis present' and the alternative whereas grade A and B were scored 'as no to mild stenosis' and grade C and D were scored as 'stenosis present'). Presence of tortuosity of the caudal nerves and presence or absence of epidural fat were also scored. Facet hypertrophy was assessed using the recommended grading system developed by Weishaupt et al., which ranges from grade 0, no facet hypertrophy, to grade 3, severe hypertrophy (Figure 1).[15, 16]

Figure 1.1. Schizas scale



MRIs were performed at each study center using a standardized protocol tailored to a 1.5 Tesla scanner. Standard sagittal T_1 and T_2 and T_2 axial images were obtained, using 3 mm contiguous slices in all directions and an in-plane resolution of $1mm^2$ or less. Two experienced neuroradiologists and one neurosurgeon dedicated to spine surgery independently evaluated all MRIs. The 3 readers were not provided with any clinical information of the included patients. Prior to the evaluation of MRIs the readers met once to evaluate and refine the MRI definitions. The blinded readers were unaware of the spinal level that was operated on and evaluated the images uniformly and independently from each other. First, the readers had to decide which levels showed most compression and were most likely to be clinically relevant (and thus in need of decompression).

Figure 1.2 – Weishaupt scale

Grade	Criteria
0	Normal facet joints space (2-4 mm width)
1	Narrowing of the facet joints space (<2mm) and/or small osteophytes and/or mild hypertrophy of the articular process
2	Narrowing of the facet joints space and/or moderate osteophytes and/or moderate hypertrophy of the articular process and/or mild subarticular bone erosions.
3	Narrowing of the facet joints space and/or large osteophytes and/or severe hypertrophy of the articular process and/or mild subchondral cysts.

Outcome

All assessments obtained from MRI were compared to continuous outcome measures used in the Felix Trial.[12, 13] Continuous outcome measures obtained at baseline and after one year were used in order to evaluate if the MRI assessments correlated to baseline clinical condition and whether they could serve as a predictor for outcome after one year.[12, 13] Consequently, the primary clinical outcome scale for this study was the Modified Roland Disability Questionnaire (MRDQ) for sciatica (scores range from 0 to 23, with higher scores indicating worse functional status).[17-25] This continuous clinical outcome scale was used in all analyses except for the receiver-operating-characteristic ROC analysis where the dichotomized MRDQ (change of 4 or less in poor results and over 4 in good results) was used.[21-23] The 100mm visual-analogue scale (VAS) back pain and VAS leg pain (with 0 representing no pain and 100 the worst pain ever experienced) were also assessed.[26-28]

Interaction of treatment arms on predictive value of MRI measurements

Patients in the Felix trial were treated using an interspinous implant without bony decompression or with bony decompression.[13] Baseline scores and one-year results based on MRDQ and VAS leg and VAS back pain were comparable in both treatment arms.[12] Multivariate analyses with regression techniques were used in order to adjust for potential confounding of effect modifying of the allocated treatment.

Statistical analysis

The majority opinion of the three readers regarding the MRI characteristics (answered independently by at least two of the three) was used in the statistical analysis. If all reviewers had a different opinion, decision was made by a group decision. Interobserver agreement regarding the MRI findings was determined with the use of absolute percentages (e.g. regarding clinical level) and kappa values (weighted in cases of ordered data). Mean scores on MRDQ and VAS back pain and VAS leg pain were correlated to MRI findings stratified according to the extent of compression according to Schizas. An analysis of the receiver-operating-characteristic (ROC) curve is only possible with a dichotomized (MRDQ) outcome scale. Mean scores on VAS back pain were correlated to the extent of facet hypertrophy. In a subgroup analysis, MRI characteristics were compared between patients treated with bony decompression and patients treated with an IPD without bony decompression.
Differences between groups for the clinical continuous data (e.g. MRDQ, VAS leg and back pain) were assessed by one-way ANOVA and regression techniques. Patients with missing MRDQ values at one year and baseline were excluded. Analysis of the receiver-operating-characteristic (ROC) curve was used to assess the diagnostic accuracy of ordinal baseline MRI findings (7 point, 4 point and dichotomized Schizas scale) for a favourable outcome, based on the MRDQ, at 1 year. The area under the ROC curve (AUC) ranges from 0 to 1 and provides a measure of a test's ability to discriminate between participants who have the outcome of interest and those who do not.[29] A test that correctly classifies all participants has an AUC of 1.0, and a test with no discriminatory value has an AUC of 0.5.[29] ROC curves were used to determine a cut-off value for a clinical test; therefore, ROC curves could not be used for the presence of epidural fat or the presence of tortuous caudal nerves. In logistic-regression models, the association between MRI findings and clinical outcome was adjusted for randomized treatment and treatment received. Furthermore, a subgroup analysis was performed by adding the term of 'stenotic level agreed by all readers' to the statistical analysis of dichotomized good outcome

based on the MRDQ. Statistical analysis were performed using SPSS software version 20. The level of statistical significance were set at 0.05.

Results

Patients

159 participants with INC due to lumbar spinal stenosis at one or two levels with an indication for decompressive surgery were blindly randomized into two groups: 80 participants received an interspinous process device (IPD) and 79 participants underwent spinal bony decompression. Results on a baseline MRI were available for 155 patients (97%). Baseline characteristics are shown in Table 1.

Readers

All three reviewers agreed (blinded for the later to be operated level) on the stenotic (symptomatic) lumbar level in 90 out of 159 patients (58%). Moderate-to-substantial agreement was found for the MRI assessment of the relevant stenotic (symptomatic) level (kappa range 0.57 to 0.64). The combined kappa for the fully agreed marked stenotic level compared to the operated level was 0.87. Decision on clinical symptomatic lumbar level by majority was possible for the remaining patients ('majority opinion'). If compared to the level patients were operated upon in the Felix Trial, the majority opinion of the most stenotic (symptomatic) level the kappa for agreement was 0.70 (good agreement).

Table 1. Characteristics of the patients at baseline

Characteristic	(n=155)
Median Age –years (range)	66 (44-83)
Male sex – no. (%)	82 (53)
Median duration of INC –months (range)*	19 (1-204)
Median BMI (range)#	29 (20-48)
Suspected clinical stenotic lumbar level$ (%)	
None	1(1)
L2-L3	2(1)
L3-L4	28(17)
L4-L5	79(49)
L1-L2 and L3-L4	1(1)
L2-L3 and L3-L4	4(3)
L2-L3 and L4-L5	5(3)
L3-L4 and L4-L5	33(20)
Roland Disability Questionnaire 23 points (SD)	13.8 (5.0)
Mean mm VAS leg pain (SD)$$	56 (24)
Mean mm VAS back pain (SD)$$	49 (26)

* Duration of intermittent neurogenic claudication (INC) in months; # Bodily-mass index is the weight in kilograms divided by the square of the heights in meters; $ based on consensus (at least two reviewers agreed on the relevant clinical level; $$ The intensity of pain was measured by a horizontal 100mm visual-analogue scale (VAS), with 0 representing no pain and 100 the worst pain ever.

Table 2a. relation between amount of compression and clinical sign at baseline and clinical outcome after 52 weeks

Schizas scale		Baseline MRDQ (0-23)	Baseline VAS leg pain (0-100)	Baseline VAS back pain (0-100)	Difference MRDQ#	Difference VAS leg pain##	Difference VAS back pain##
A	Mean	16	58	56	-7	-13	-22
	N	7	10	10	7	9	9
	SD	4	20	30	5	39	45
B	Mean	13	57	50	-4	-27	-17
	N	24	24	24	21	23	23
	SD	5	23	25	6	34	27
C	Mean	14	54	44	-8	-37	-23
	N	59	63	63	57	62	62
	SD	5	24	27	6	28	26
D	Mean	14	53	53	-9	-39	-34
	N	31	32	33	30	32	33
	SD	5	26	25	6	31	29
Total	Mean	14	55	48	-7	-34	-25
	N	121	129	130	115	126	127
	SD	5	24	27	6	31	29

Table 2b. relation between facet hypertrophy and clinical sign at baseline and clinical outcome after 52 weeks

Weishaupt scale		Baseline MRDQ (0-23)	Baseline VAS leg pain (0-100)	Baseline VAS back pain (0-100)	Difference MRDQ#	Difference VAS leg pain##	Difference VAS back pain##
1,00	Mean	15	50	61	-7	-11	-7
	N	3	7	7	3	6	6
	SD	2.5	25	33	7	40	40
2,00	Mean	14	53	47	-6	-26	-24
	N	43	45	45	41	43	43
	SD	6	26	29	6	31	24
3,00	Mean	14	58	49	-8	-40	-26
	N	80	83	84	75	83	84
	SD	5	22	24	7	29	29
Total	Mean	14	56	49	-7	-34	-25
	N	126	135	136	119	132	133
	SD	5	24	26	6	31	28

N: number of patients in group; # Difference between baseline and 52 weeks MRDQ values; SD Standard Deviation

Table 3. interaction of MRI characteristics and clinical signs at baseline and one-year follow-up

MRI and clinical signs	Baseline MRDQ (p-value)$	Baseline VAS leg (p-value)$	Baseline VAS back (p-value)$	Difference MRDQ# (p-value)$	Difference VAS leg ## (p-value)$	Difference VAS back ## (p-value)$
Schizas scale	0.61	0.86	0.26	0.03	0.09	0.13
Weishaupt scale	0.77	*	0.43	0.44	*	0.26
Presence of tortuous caudal nerves	0.19	0.27	0.45	<0.01	0.08	0.04
Absence of epidural fat	0.35	0.99	0.11	0.03	0.05	<0.01

$ p-values as interaction of MRI characteristics and clinical outcome estimated by linear or logistic regression models; N: number of patients in group; # Difference between baseline and 52 weeks MRDQ values; ## Difference between baseline and 52 weeks VAS values; * not a relevant correlation and thus not tested.

MRI findings at baseline

The readers judged the grade of compression based on the Schizas scale: 4 patients with level A2, 4 patients with level A3, 2 patients with level A4, 24 patients level B, 64 patients level C, and 36 patients level D (23 missing data for this item). The MRIs of these symptomatic patients were seldom scored with A, therefore the subdividing of grade A was not used in our analysis. The kappa of agreement of the Schizas scale was poor (kappa ranges 0.16-0.46). By comparison, the dichotomized Schizas scale (scale A and B versus scale C and D) kappa was somewhat better (kappa ranges 0.32-0.59). Facet hypertrophy was present in the vast majority of patients: 86 patients

(63%) had severe hypertrophy (level 3) and 43 (31%) patients had moderate facet hypertrophy (level 2) and 8 patients (6%) had mild hypertrophy (level 1) based on the Weishaupt scale (18 missing data for this item). However, the scale's kappa of agreement was very poor (range 0.09-0.34). Tortuous caudal nerves were seen on MRI in 45 patients (29%; kappa of agreement ranges 0.67-0.70). Absence of epidural fat tissue was seen in 51 patients (33%; kappa of agreement ranges 0.53-0.72).

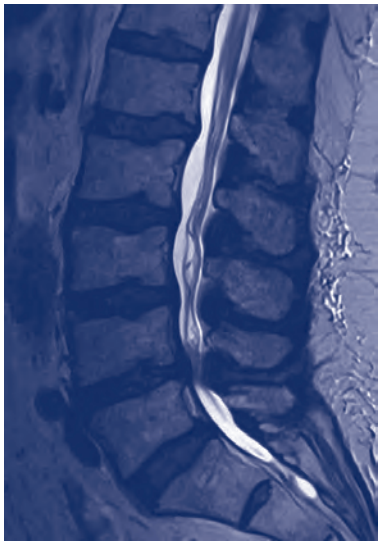
Table 4. agreement among the readers

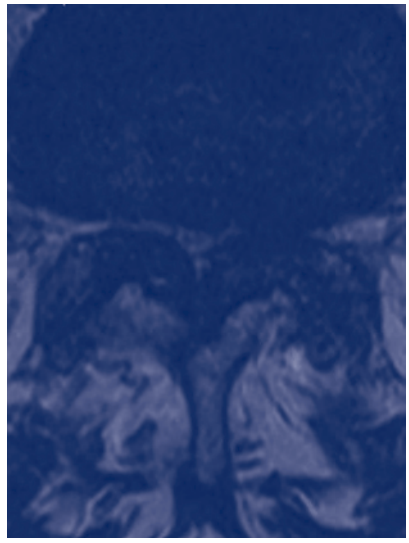
	A versus B	A versus C	B versus C
Agreement on most clinical stenotic level	0.57 (n=155)	0.62 (n=155)	0.64 (n=155)
Amount of stenosis based on Schizas scale$			
- Level L3-L4	0.24 (n=58)	0.33 (n=59)	0.16 (n=67)
- Level L4-L5	0.35 (n=102)	0.46 (n=115)	0.33 (n=115)
Amount of stenosis based on dichotomized Schizas$$			
- Level L3-L4	0.45 (n=58)	0.42 (n=59)	0.32 (n=67)
- Level L4-L5	0.56 (n=102)	0.55 (n=115)	0.59 (n=115)
Facet hypertrophy based on Weishaupt scale#	0.09 (n=106)	0.34 (n=110)	0.11 (n=114)
Presence of tortuous caudal nerves	0.70 (n=155)	0.68 (n=155)	0.67 (n=155)
Absence of epidural fat#	0.59 (n=102)	0.72 (n=106)	0.53 (n=114)

A and C represent the two neuroradiologists, while B represents the neurosurgeon.
$ kappa values of agreement of reviewers' measurements; N: number of patients in group; $ based on levels L3-L4 and L4-L5 whereas other levels were sporadic reports and thus not representative kappa values; $$ based on the dichotomized Schizas scale whereas A and B were scored as 'no stenosis'; scales C and D were scored as 'stenosis present'. # based on level L4-L5 whereas other levels were sporadic reports and thus not representative kappa values.

Association between MRI findings and baseline symptoms (diagnostic value of MRI)

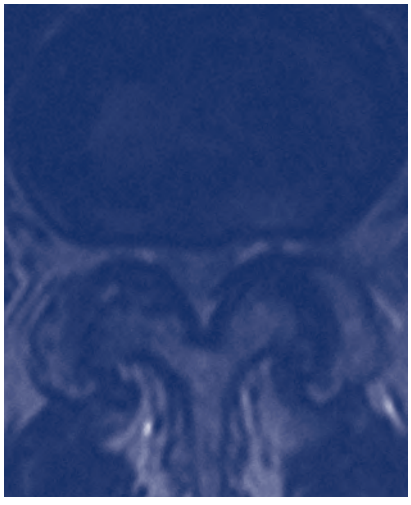
The grade of stenosis based on the different types of Schizas scale did not correlate with functional status (MRDQ) at baseline (p-value 0.61) nor with VAS back pain (p-value 0.26) or VAS leg pain (p-value 0.86)(Tables 2a & 3). The severity of facet hypertrophy according to the Weishaupt scale did not correlate with the clinical parameters MRDQ or VAS back at baseline. Tables 2b & 3). The presence of tortuous caudal nerves and the absence of epidural fat did not correlate to baseline clinical condition (p-values ranging 0.11-0.99).





Figures 2.1 and 2.2 (above). with absence of epidural fat and caudal roots;

Figure 2.3 (below). MRI of a patient with Schizas C



Association between MRI findings and clinical outcome (predictive value of MRI parameters)

There was a significant correlation between preoperative compression (based on the Schizas scale) and clinical outcome after 52 weeks MRDQ (p-value 0.03). However, this was caused by an opposite effect in the relative small group of patients in group A. When the Schizas scale was dichotomized (A and B scales for 'no stenosis' and C and D 'stenosis present') the correlation between the preoperative extent of compression and the MRDQ after 52 weeks disappeared (p-value 0.86). The grade of stenosis on the preoperative MRI (based on the Schizas scale divided in four subscales and dichotomized) did not correlate to the decrease in VAS leg pain over 52 weeks following surgery (p-value 0.09) nor to the decrease in VAS back pain (p-value 0.13)(Table 3). Therefore, the grade of stenosis on the preoperative MRI is judged not to be a good predictor for clinical outcome defined as VAS leg or back pain one year after surgery.

The preoperative observers' ratings on the Schizas scale assessing the extent of stenosis on MRI did not distinguish between patients with a favourable outcome based on MRDQ outcome versus those with an unfavourable outcome (AUC 0.51; CI 0.40 to 0.62). The dichotomized MRDQ was used in order to perform an analysis of the receiver-operating-characteristic (ROC) curve.

Figures 2 show two case-examples of a patient with severe stenosis (tortuosity of caudal roots and absence of epidural fat and a good clinical outcome (Figures 2.1 and 2.2); versus a patient with moderate stenosis (Schizas C) and poor clinical outcome (Figure 2.3).

The severity of preoperative facet hypertrophy on the preoperative MRI is not a good predictor for the decrease in VAS back pain (p-value 0.26) during the 52 weeks following surgery (Table 3). The severity of facet hypertrophy did also not distinguish between patients with favourable outcome versus those with unfavourable clinical outcome based on MRDQ (p-value 0.44; AUC 0.52; CI 0.42 to 0.63). By contrast, absence of epidural fat and/or presence of tortuous caudal nerves on the baseline MRI were good predictors for good long-term clinical outcome (more than 4 points decrease on the MRDQ) in operated patients (p-values 0.03 and <0.01 respectively) (Table 3). 67% of patients (SD 0.47) demonstrated good recovery (MRDQ decrease of more than 4 points) when all readers agreed on the clinical symptomatic level. 54% of patients (SD 0.50) demonstrated good recovery (MRDQ) when not all readers agreed on the clinical symptomatic level (agreed by all reviewers p-value of 0.1).

Interaction of treatment arms on predictive value of MRI measurements

MRDQ decreased from 13.0 at baseline to 6.9 after 52 weeks in the IPD group and from 14.4 at baseline to 8.1 after 52 weeks in the bony decompression group (p-value 0.28). VAS back pain decreased from 60 to 23 (IPD) and from 49 to 31 (bony decompression) between baseline and 52 weeks (p-value 0.09). VAS leg pain decreased from 52 to 23 (IPD) and from 58 to 26 (bony decompression) between baseline and 52 weeks (p-value 0.54). Multivariate analysis on the possible confounding of the randomization groups did not result in any significant interaction with any baseline and one year outcome (predictive value) scales.

Discussion

This is the first prospective study comparing baseline MRI findings with baseline complaints and long-term follow-up after surgical treatment. Inter-rater variability between readers for the MR parameters was relatively high in our study, –which is interesting given the good inter-observer reproducibility reported in the recent study on interpretation of herniated discs, by the same readers.[30] Stenosis severity on baseline MRI correlated poorly with baseline functional measurements and did not distinguish between patients with favourable outcome and unfavourable outcome after decompression. Only presence of extreme compression at baseline (patients with absence of epidural fat and/or the presence of tortuous caudal nerves on their MR images) appeared to be a fairly good predictor of good clinical outcome after operation. Furthermore, these MR parameters showed good interobserver reproducibility, in contrast to the other MR parameters. As suggested by earlier MRI studies assessing lumbar spinal compression in patients without any clinical complaints, compression on MRI is no more than a sign supporting the complaints of the patient.[7] Operating patients without significant complaints, even though they demonstrate compression on baseline MRI, is therefore not wise, and clinical symptoms should guide any decision on operation.

MRI is viable to appoint the pathological level in case surgical decompression is considered. However, the current study demonstrates that the potential of MRI to grade the extent of compression is poor, regarding the low kappa rates. Some studies have been performed before to evaluate the interobserver variability on the topic of radiological images in patients with lumbar spinal stenosis.[31, 32]

These studies could likewise not demonstrate a correlation between the extent of compression on the Schizas scale and baseline or outcome clinical condition: Schizas et al presented an interobserver agreement kappa of only 0.44, similar compared with the kappa range presented in this study[14]. Perhaps myelography alone, myelography combined with MRI, or new diagnostic modalities could help to solve these problem. Morita et al and Ogura et al concluded in retrospective studies (not correlated to baseline clinical condition or long-term outcome after surgery) that CT myelography is still superior or equal compared with MRI.[32, 33] In favour of the combination of MRI and CT myelography are the data in a former retrospective study concluding that the interobserver correlation improved – slightly – by adding myelography to MRI. However, these authors could not demonstrate a correlation with clinical symptoms either.[34] Paine et al presented their results on 95 consecutive INC patients with LSS on myelography in 1976 and reported 80% 'good' recovery and 11% 'fair' recovery 2,5 years after surgery.[35, 36] Nowadays, diagnostics in INC patients with LSS are mainly performed with MRI with agreement of the relevant stenotic level on MRI of only 58% among highly experienced reviewers. Combining this finding with the observation that only 65% of the patients with INC have good recovery after surgery.[12,37-39] suggests that myelography is more accurate to diagnose clinically relevant LSS. Namely, a good diagnosis is highly likely to result in a high success percentage of treatment. An important shortcoming of this conclusion is that in 1976 the surgical decompression of the cauda equina was not as broadly implemented as it is nowadays: it is only logical to expect that only the severe cases made it to the myelographic evaluation and subsequent decompression.

An important limitation of our study is that the included patients had maximal two stenotic lumbar levels. This was due to the design of the Felix Trial, in which it was not possible to implant patients with more than two IPDs.[13] Secondly, outcome based on MRI findings were secondary outcome measurement in the original Felix Trial protocol.[13] Sample size calculation of the trial was based on clinical outcome scales. It is debatable if we had enough power to give a final answer to our research question. Furthermore, this study has a potential selection bias due to the fact that all patients were operated.

Finally, this study shows that future research in spinal stenosis surgery should also focus on diagnostic tools to diagnose stenosis. Further research may focus on the evaluation of the additive value of myelography to MRI evaluation of LSS. For

now, MRI should only be used to confirm the diagnosis of LSS in case the clinical condition is so invalidating that surgical decompression is considered. The patient should know that the success rate of the surgical intervention is only 65%, but that the chance on success is higher in case epidural fat is absent and/or tortuous caudal nerves are present.[37-39]

Reference List

1. Verbiest H. [Primary stenosis of the lumbar spinal canal in adults, a new syndrome.]. Ned Tijdschr Geneeskd 1950;94(33):2415-2433.
2. EVANS JG. Neurogenic Intermittent Claudication. Br Med J 1964;2(5415):985-987.
3. Verbiest H. [Further reports on primary stenosis of the lumbar spinal canal in adults.]. Ned Tijdschr Geneeskd 1951;95(27):1965-1970.
4. Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. J Bone Joint Surg Br 1954;36-B(2):230-237.
5. de Schepper EI, Overdevest GM, Suri P et al. Diagnosis of lumbar spinal stenosis: an updated systematic review of the accuracy of diagnostic tests. Spine (Phila Pa 1976) 2013;38(8):E469-E481.
6. Boden SD, McCowin PR, Davis DO, Dina TS, Mark AS, Wiesel S. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am 1990;72(8):1178-1184.
7. Ishimoto Y, Yoshimura N, Muraki S et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20(10):1103-1108.
8. Kovacs FM, Royuela A, Jensen TS et al. Agreement in the interpretation of magnetic resonance images of the lumbar spine. Acta Radiol 2009;50(5):497-506.
9. Kovacs FM, Martinez C, Arana E et al. Uncertainties in the measurement of lumbar spinal stenosis at MR imaging: are they clinically relevant? Radiology 2012;263(1):310-311.
10. Oba H, Takahashi J, Futatsugi T et al. Study of dural sac cross-sectional area in early and late phases after lumbar decompression surgery. Spine J 2013;13(9):1088-1094.
11. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331(2):69-73.
12. Moojen WA, Arts MP, Jacobs WC et al. Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. BMJ 2013;347:f6415.
13. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord 2010;11:100.
14. Schizas C, Theumann N, Burn A et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphology of the dural sac on magnetic resonance images. Spine (Phila Pa 1976) 2010;35(21):1919-1924.
15. Kettler A, Wilke HJ. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. Eur Spine J 2006;15(6):705-718.
16. Weishaupt D, Zanetti M, Boos N, Hodler J. MR imaging and CT in osteoarthritis of the lumbar facet joints. Skeletal Radiol 1999;28(4):215-219.
17. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. Spine (Phila Pa 1976) 1986;11(1):28-30.
18. Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. Med Care 1995;33(4 Suppl):AS286-AS291.
19. Hutchinson PJ, Laing RJ, Waran V, Hutchinson E, Hollingworth W. Assessing outcome in lumbar disc surgery using patient completed measures. Br J Neurosurg 2000;14(3):195-199.
20. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006;332(7555):1430-1434.
21. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989;27(3 Suppl):S217-S232.
22. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine (Phila Pa 1976) 1995;20(17):1899-1908.

23. Roland M, Morris R. A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care. Spine (Phila Pa 1976) 1983;8(2):145-150.
24. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976) 1983;8(2):141-144.
25. Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. Phys Ther 1996;76(3):268-275.
26. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983;16(1):87-101.
27. Main CJ, Wood PL, Hollis S, Spanswick CC, Waddell G. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. Spine (Phila Pa 1976) 1992;17(1):42-52.
28. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. Spine (Phila Pa 1976) 2006;31(11):1276-1280.
29. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. Radiology 2003;229(1):3-8.
30. el BA, Vleggeert-Lankamp CL, Nijeholt GJ et al. Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med 2013;368(11):999-1007.
31. Cheung JP, Shigematsu H, Cheung KM. Verification of measurements of lumbar spinal dimensions in T1- and T2-weighted magnetic resonance imaging sequences. Spine J 2013.
32. Ogura H, Miyamoto K, Fukuta S, Naganawa T, Shimizu K. Comparison of magnetic resonance imaging and computed tomography-myelography for quantitative evaluation of lumbar intracanalar cross-section. Yonsei Med J 2011;52(1):137-144.
33. Morita M, Miyauchi A, Okuda S, Oda T, Iwasaki M. Comparison between MRI and myelography in lumbar spinal canal stenosis for the decision of levels of decompression surgery. J Spinal Disord Tech 2011;24(1):31-36.
34. Song KS, Jang EC, Jung HJ, Kim KW, Yu H. Observer variability in the evaluation of multiple lumbar stenosis by routine MR--myelography and MRI. J Spinal Disord Tech 2008;21(8):569-574.
35. Paine KW. Results of decompression for lumbar spinal stenosis. Clin Orthop Relat Res 1976;(115):96-100.
36. Paine KW. Myelography in the lumbar disk syndrome. Clin Orthop Relat Res 1976;(115):47-48.
37. Malmivaara A, Slatis P, Heliovaara M et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 2007;32(1):1-8.
38. Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. Spine (Phila Pa 1976) 1992;17(1):1-8.
39. Weinstein JN, Tosteson TD, Lurie JD et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008;358(8):794-810.

Chapter 9

General Discussion
Scientific evaluation of spinal implants; an ethical necessity

Submitted July 2014.

Moojen WA (A1/A2)
Bredenoord AL (A3)
Viergever RF (A4/A5)
Peul WC (A1/A2)

[A1] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands
[A2] Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
[A3] Department of Medical Humanities, University Medical Center Utrecht, Utrecht, The Netherlands
[A4] Department of Primary and Community Care, Radboud University Medical Centre, Nijmegen, The Netherlands
[A5] Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom

Abstract

The clinical introduction of novel medical devices and surgical techniques often occurs without evidence of good methodological quality and with relatively little oversight and regulation. As a consequence, the safety, efficacy and long-term effects of devices are frequently insufficiently known upon device approval. Recent controversies surrounding the PIP breast implants, metal-on-metal hip implants and interspinous implants underscore the need to reconsider how innovation in medical devices and surgical techniques can adhere to sound ethical standards without inhibiting surgical research and development. In this paper the introduction of spinal implants is taken as an example to firstly discuss the scientific and ethical challenges of developing, testing and introducing novel medical devices, and to secondly identify avenues for improving the existing regulatory frameworks for such innovation. Two measures for improvement are most feasible in the short-term: demanding prospective studies before device introduction, and developing registries in order to monitor and evaluate new medical devices.

Introduction

The clinical introduction of novel surgical techniques, particularly those using medical implant devices, often occurs with relatively little oversight and regulation.[1] This is in contrast with the strong regulatory requirements that are in place for the introduction of novel pharmaceuticals. As a consequence the safety, efficacy and long-term effects of medical implant devices are often insufficiently known before they are used in patients.[1,2] Adoption of the novel device is frequently driven by other factors than evidence, such as the enthusiasm of the surgeon or marketing.[3] Recent controversies surrounding the PIP breast implants, metal-on-metal hip implants and spinal implants underscore the need to reconsider how patients can be protected from ineffective, or potentially harmful, medical devices and surgical techniques without inhibiting surgical research and development (R&D).[4–12] In this paper the introduction of spinal implants is taken as an example to firstly discuss the scientific and ethical challenges of developing, testing and introducing novel medical devices, and to secondly identify avenues for improving the existing regulatory frameworks for such innovation. We argue that prospective comparative effectiveness studies should be mandatory before approval of a device, and that post-marketing surveillance for all medical devices, as proposed by the EU, should be introduced as soon as possible.[13,14]

History of spinal implants: the Interspinous Process Devices

Spinal implants are widely used for different indications, ranging from indisputable indications such as reconstruction of the destabilized spine by trauma and reconstruction after surgical resection of vertebral tumours, to less clear reasons such as stabilization for degenerative spinal conditions. Most implants are used for the latter group of degenerative spinal diseases, one of which is lumbar spinal stenosis (LSS). LSS is caused by arthrosis (degeneration) of the facet joints and development stenosis, which can result in lumbar nerve root compression. Removal of the bone and arthrosis around the nerve (bony decompression e.g. laminectomy) is the gold standard to treat LSS in the elderly population. The reported successful clinical outcome after bony decompression is only 64% and many patients remain to have associated low back pain.[15–18] In an effort to improve clinical outcome, a French group introduced a new non-rigid fixation (interspinous process devices (IPDs)) for patients with LSS and associated back pain in 1984: the Wallis system.[19–21]

The Wallis system implants were tested for durability in cadaveric studies and were first implanted in humans in 1986.[20] The results of these first procedures were retrospectively published.[20] Only after this period were patients included in a (non-comparative) prospective study, during which the device was implanted in over 300 patients. The study showed good recovery in 60% of patients.[21] After this study, commercial development of the system was started. While the research group was planning to perform a randomized controlled trial (RCT), such a prospective comparative study of this implant is not available in Pubmed.[20,21] After the introduction of this implant by Senegas, the development of other IPDs followed, such as Minns, X-stop and Coflex.[22–25] Cadaveric studies did not show any biomechanical difference between the various IPDs and they were therefore considered as interchangeable, although differences in clinical effectiveness were not investigated.[24]

After introduction of these devices, various studies were conducted to test the effectiveness and safety of IPD treatment for LSS. However, most of these studies did not compare the results with other interventions, and most did not have prospective study designs.[21,25,26] It took 30 years (from 1984 until 2013) until two prospective studies were published that compared IPD treatment with conventional (surgical) care.[16,27–30] These studies showed that treatment with IPD was not superior compared to bony decompression without implants and that IPD treatment resulted in a higher reoperation rate.[16,30] A third study was terminated because of the high number of reoperations (complications) in the experimental (IPD) group.[28]

The problem of lacking evidence for IPD use extends beyond LSS. Nowadays there are multiple questionable indications for implantation of IPDs: some are used as stand-alone for LSS, others as adjuvant to surgical bony decompression for LSS in the hope to decrease back pain, and yet others to prevent disease at adjacent lumbar segments.[16,28,31] For these indications, IPDs remain in use without any evidence of treatment benefit. Furthermore, the lack of evidence for treatment of LSS extends beyond IPDs. Before the introduction of IPDs, lumbar spines that were "destabilized" after LSS were frequently rigidly stabilized by pedicle screws, and since the mid-nineties of the last century vertebral interbody cages were added to this process.[32–34] However, pedicle screws and discal interbody cages, whose use is widespread for LSS, were introduced without any evidence of added value

compared with conventional surgical decompression without implants, or even any evidence of incidence of spinal instability.[35–37]

Thirty years after the first introduction of IPDs for LSS, it is now clear that there is no justification for treating LSS patients with IPDs.[16] Although precise numbers about the number of implanted IPDs are not available, at least three hundred thousand patients have been implanted with these devices since their introduction.[38] How was it possible that patients were not protected from these harmful devices and society from the use of these costly implants by regulations or any other measurements?

Present regulatory practice in the EU and US

In Europe, what is needed since 1993 for market introduction of a device is the CE ('Conformité Européenne'). The CE approval guarantees that implants do not fall apart or have harmful material in them. However, a CE approval will not guarantee that the medical device will work in patients, or that it does not cause harm in other ways, such as higher re-operation rates as compared to other interventions. Recently, the European Committee (EC) has begun to realize the inadequacy of these regulations and in 2013 released a recommendation for a common framework for a unique device identification system (a monitoring system or registry) of medical devices in the EU.[14] From 1990 till 2013 the EC launched several directives, recommendations and proposals to realize such an identification system for safe, effective and innovative medical devices. The first amending directive which urges for a registry dates from 1993.[13] However, none of these directives were ever implemented.[13] The suggestions made in the Commission Recommendation of 2013 to assure traceability are sound, yet to this day, they remain just a proposal.[14]

The US Food and Drug Administration (FDA) has a stricter system for device approval, in which inventors are required to perform randomized studies before introduction of a device. However, the LSS case shows that this system also shows some shortcomings, since the FDA does not demand that the experimental treatment is compared with the gold standard.[6] IPD treatment with bony decompression is nowadays approved in the US, after the publication of an FDA study on IPD treatment.[27] However, this study did not compare the experimental treatment (IPD) with the gold standard (bony decompression), but with another experimental treatment (bony decompression with fixation techniques). This has happened

before in the spine research field: the FDA study of the CHARITÉ artificial disc for low back-pain compared artificial disc (experimental treatment) with another fusion technique (and not with the standard care).[39]

Ethical lessons

In contrast to pharmaceuticals, where rigorous safety and comparative effectiveness research (multiple phases of RCTs) are required for approval, novel medical devices can be introduced in patients without sound evidence and with relatively little oversight and regulation in patients. The reluctance to set up surgical research and generate systematically collected evidence on the safety and effectiveness of devices is sometimes defended by 'surgical exceptionalism', the view that the somewhat exceptional ethical or regulatory status of surgery is justified by the unique dynamic nature of surgery.[40] There are several reasons why surgeons have taken this view. First, surgical techniques, unlike drugs, do not have chemical compositions, physical properties, routes and rates of excretion, or other qualities that can be measured precisely. Second, surgical procedures are rarely introduced as fully defined, easily reproducible techniques. Rather, they come as principles for solving particular problems, sometimes of an urgent nature.[41] Finally, in situations in which known interventions are of questionable value or where effective interventions do not exist, some state that a rigid regulatory paradigm cannot be applied to the innovative activities at the frontiers of surgical practice without adversely impacting the prospects for advancing the state of the art.[40]

We partially adhere to surgical exceptionalism, accepting that surgery is sufficiently unique that it should not be governed by the same rules/requirements that apply in pharmaceuticals, but resisting that the entire domain of surgery (e.g. non-acute diseases like LSS) would not be suitable for rigorous scientific evaluation of interventions.[40] The need for more rigorous evaluation of novel surgical procedures and medical devices is increasingly acknowledged in this era of evidence based medicine.[42–45] Furthermore, unnecessary and/or unproven treatments can harm patients and can be unnecessarily expensive for societies with growing health care expenditures. The dynamic nature of surgical practice does not preclude rigorous evaluation of new interventions in the surgical domain, and vice versa. It is a wide misunderstanding that if interventions do no good they will at least do no harm and therefore nothing would be lost – this 'no lose philosophy' has already been criticized in the 1970s.[46]

Fifty percent of all new drugs have important side effects discovered only after approval and marketing.[6] Taking into account the statement of FDA officials that "New devices are less likely than drugs to have theirs safety established clinically before they are marketed", the amount of side-effects caused by surgical innovation and devices is potentially even higher.[47] Medical devices are complex assemblies of multiple components, making it impossible to design an implantable device without risks or harms.[48] Since "implanted body parts cannot be recalled as easily as defective auto parts", inadequately tested devices should be prevented from coming on the market, and systems for monitoring safety after a medical device is marketed should be implemented.[2,4] This is true in particular given the lack of informed decision-making for patients, who are commonly operated on without sufficient awareness of the potential harm of experimental implants, and given the substantial commercial interests and aggressive marketing tactics of large international producers of devices.

Way forward for the introduction of spinal implants

The case study set out above gives strong arguments for introducing a stepwise approach to introducing new implant devices.[6,49] We limit our discussion of this approach to implant devices because of the large differences between various (types of) medical devices, and their consequences for regulation in this field.[6,50] Several authors have made suggestions for what needs to be done to avoid harmful and costly mistakes as have occurred in the introduction of IPDs for LSS. In our view, two measures are most feasible in realizing a regulatory system that ensures that medical implant devices are safe and effective. First, prospective controlled trials that compare the experimental device to the present gold standard for that disease should be required for device approval. To be approved, the effectiveness of the new device should be at least equal or compared with the gold standard, and it should be safe. Second, a monitoring system (post-implementation registries) for all medical devices, as suggested by the European Committee, could help to trace these implants and ensure rigorous clinical follow-up. The establishment of registries would allow the collection of reliable data on adverse events and monitoring of long-term safety and efficacy.[2] By early detection of negative results the use of an implant could be stopped or modified in order to avoid further damage. Moreover, it will give a clear overview of the innovations present in the field, so other innovators are not likely to repeat failed surgical procedures with certain implants.

Conclusion

Medical implant-devices are frequently introduced without adequate evidence of safety and efficacy. This results in harmful medical practices for patients. Steps should be taken to strengthen regulation for device development and introduction, without unnecessarily inhibiting R&D. Two measures are most feasible in the short-term: (1) requiring prospective studies before device approval, and (2) developing registries in order to monitor and evaluate new medical devices and all surgical implants.

Reference List

1. (2010) World Health Organization. Medical Devices: Managing the Mismatch.
2. Bredenoord AL, Giesbertz NA, van Delden JJ (2013) Consent for medical device registries. Public Health Ethics 6: 226-229.
3. Campbell B, Stainthorpe AC, Longson CM (2013) How can we get high quality routine data to monitor the safety of devices and procedures? BMJ 346: f2782.
4. Ardaugh BM, Graves SE, Redberg RF (2013) The 510(k) ancestry of a metal-on-metal hip implant. N Engl J Med 368: 97-100. 10.1056/NEJMp1211581 [doi].
5. Benkimoun P (2013) Founder of PIP breast implant company gets four year prison sentence. BMJ 347: f7528.
6. Carragee EJ, Deyo RA, Kovacs FM, Peul WC, Lurie JD, Urrutia G, Corbin TP, Schoene ML (2009) Clinical research: is the spine field a mine field? Spine (Phila Pa 1976) 34: 423-430. 10.1097/BRS.0b013e318198c962 [doi].
7. Carragee EJ (2010) The increasing morbidity of elective spinal stenosis surgery: is it necessary? JAMA 303: 1309-1310. 303/13/1309 [pii];10.1001/jama.2010.402 [doi].
8. Deyo RA, Nachemson A, Mirza SK (2004) Spinal-fusion surgery - the case for restraint. N Engl J Med 350: 722-726. 10.1056/NEJMsb031771 [doi];350/7/722 [pii].
9. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG (2010) Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 303: 1259-1265. 303/13/1259 [pii];10.1001/jama.2010.338 [doi].
10. Deyo RA (2010) Treatment of lumbar spinal stenosis: a balancing act. Spine J 10: 625-627. S1529-9430(10)00380-3 [pii];10.1016/j.spinee.2010.05.006 [doi].
11. Martindale V, Menache A (2013) The PIP scandal: an analysis of the process of quality control that failed to safeguard women from the health risks. J R Soc Med 106: 173-177. 106/5/173 [pii];10.1177/0141076813480994 [doi].
12. Wise J (2013) Review that found PIP implants to be safe was flawed, say campaigners. BMJ 346: f3034.
13. (1993) Council Directive 92/42/EEC.
14. 2013) Commission recommendation on a common framework for a unique device identification system of medical devices in the Union . Official Journal of the European Union .
15. Malmivaara A, Slatis P, Heliovaara M, Sainio P, Kinnunen H, Kankare J, Dalin-Hirvonen N, Seitsalo S, Herno A, Kortekangas P, Niinimaki T, Ronty H, Tallroth K, Turunen V, Knekt P, Harkanen T, Hurri H (2007) Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 32: 1-8. 10.1097/01.brs.0000251014.81875.6d [doi];00007632-200701010-00002 [pii].
16. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW, Vleggeert-Lankamp CL, Peul WC (2013) Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. BMJ 347: f6415.
17. Turner JA, Ersek M, Herron L, Deyo R (1992) Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. Spine (Phila Pa 1976) 17: 1-8.
18. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, Birkmeyer NJ, Hilibrand AS, Herkowitz H, Cammisa FP, Albert TJ, Emery SE, Lenke LG, Abdu WA, Longley M, Errico TJ, Hu SS (2007) Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 356: 2257-2270. 356/22/2257 [pii];10.1056/NEJMoa070302 [doi].
19. Senegas J, Etchevers JP, Vital JM, Baulny D, Grenier F (1988) [Recalibration of the lumbar canal, an alternative to laminectomy in the treatment of lumbar canal stenosis]. Rev Chir Orthop Reparatrice Appar Mot 74: 15-22.
20. Senegas J (1991) [Surgery of the intervertebral ligaments, alternative to arthrodesis in the treatment of degenerative instabilities]. Acta Orthop Belg 57 Suppl 1: 221-226.

21. Senegas J (2002) Mechanical supplementation by non-rigid fixation in degenerative intervertebral lumbar segments: the Wallis system. Eur Spine J 11 Suppl 2: S164-S169. 10.1007/s00586-002-0423-9 [doi].

22. Minns RJ, Walsh WK (1997) Preliminary design and experimental studies of a novel soft implant for correcting sagittal plane instability in the lumbar spine. Spine (Phila Pa 1976) 22: 1819-1825.

23. Richter A, Schutz C, Hauck M, Halm H (2009) Does an interspinous device (Coflex) improve the outcome of decompressive surgery in lumbar spinal stenosis? One-year follow up of a prospective case control study of 60 patients. Eur Spine J . 10.1007/s00586-009-1229-9 [doi].

24. Wilke HJ, Drumm J, Haussler K, Mack C, Steudel WI, Kettler A (2008) Biomechanical effect of different lumbar interspinous implants on flexibility and intradiscal pressure. Eur Spine J 17: 1049-1056. 10.1007/s00586-008-0657-2 [doi].

25. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, Johnson DR, Skidmore GA, Vessa PP, Dwyer JW, Puccio S, Cauthen JC, Ozuna RM (2004) A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. Eur Spine J 13: 22-31. 10.1007/s00586-003-0581-4 [doi].

26. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC (2011) Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. Eur Spine J 20: 1596-1606. 10.1007/s00586-011-1873-8 [doi].

27. Davis RJ, Errico TJ, Bae H, Auerbach JD (2013) Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective, randomized, multicenter, Food and Drug Administration Investigational Device Exemption trial. Spine (Phila Pa 1976) 38: 1529-1539. 10.1097/BRS.0b013e31829a6d0a [doi].

28. Lønne G (2013) Treatment of Lumbar Spinal Stenosis; Comparison of Two Different Surgical Methods; Mini-invasive Decompression to X-stop (LSSS).

29. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC (2010) The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord 11: 100. 1471-2474-11-100 [pii];10.1186/1471-2474-11-100 [doi].

30. Stromqvist BH, Berg S, Gerdhem P, Johnsson R, Moller A, Sahlstrand T, Soliman A, Tullberg T (2013) X-Stop Versus Decompressive Surgery for Lumbar Neurogenic Intermittent Claudication: Randomized Controlled Trial With 2-Year Follow-up. Spine (Phila Pa 1976) 38: 1436-1442. 10.1097/BRS.0b013e31828ba413 [doi].

31. Stromqvist B, Berg S, Gerdhem P, Johnsson R, Sahlstrand T, Tullberg T (7 A.D.) Indirect Decompression (X-Stop) versus Conventional Decompressive Surgery for Lumbar Spinal Claudication - A Prospective Randomized Trial. The Spine Journal 7: 70.

32. Cloward RB (1952) The treatment of ruptured lumbar intervertebral disc by vertebral body fusion. III. Method of use of banked bone. Ann Surg 136: 987-992.

33. Cloward RB (1953) The treatment of ruptured lumbar intervertebral discs by vertebral body fusion. I. Indications, operative technique, after care. J Neurosurg 10: 154-168. 10.3171/jns.1953.10.2.0154 [doi].

34. Ekman P, Moller H, Tullberg T, Neumann P, Hedlund R (2007) Posterior lumbar interbody fusion versus posterolateral fusion in adult isthmic spondylolisthesis. Spine (Phila Pa 1976) 32: 2178-2183. 10.1097/BRS.0b013e31814b1bd8 [doi];00007632-200709150-00006 [pii].

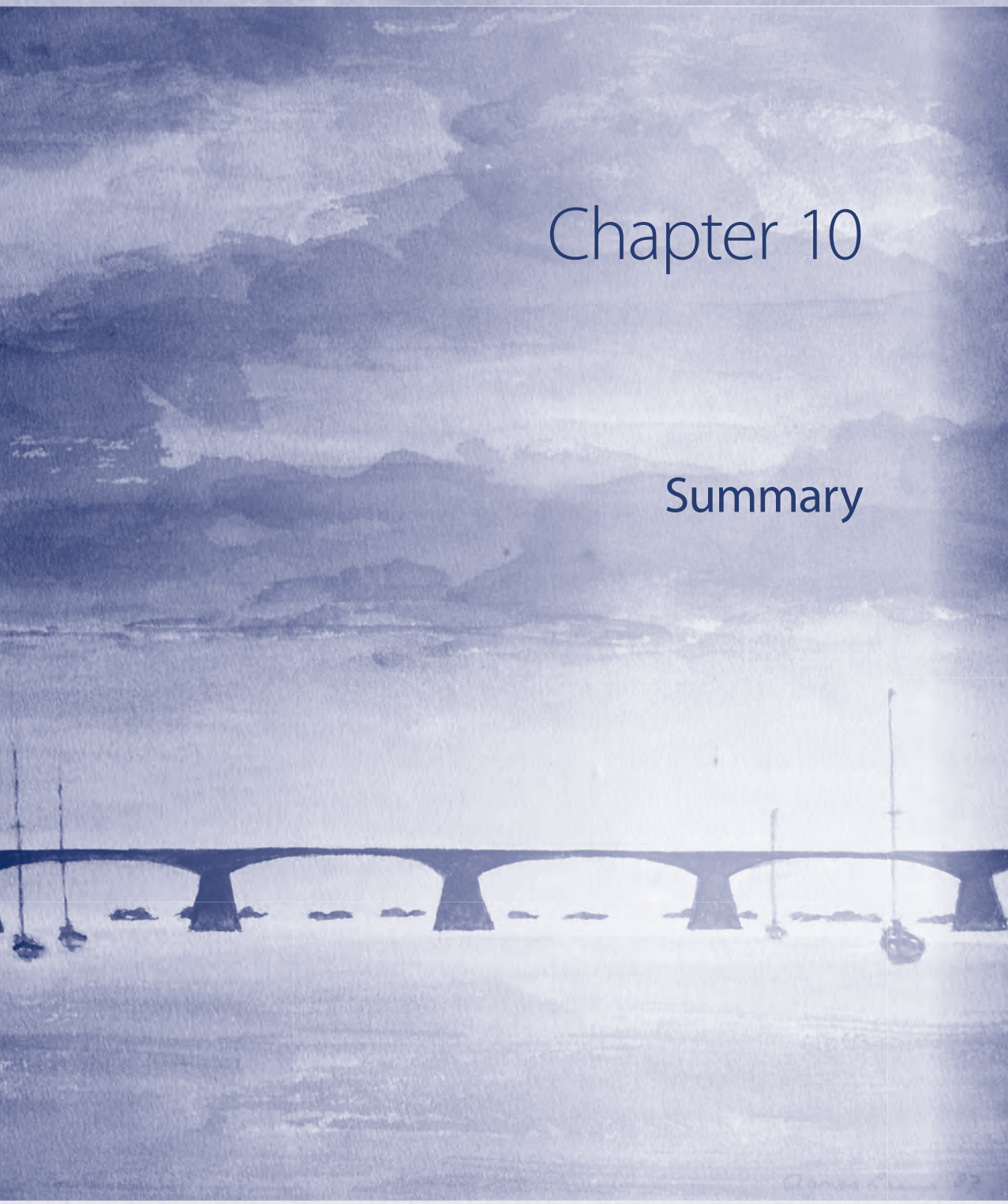
35. Jacobs WC, Vreeling A, de KM (2006) Fusion for low-grade adult isthmic spondylolisthesis: a systematic review of the literature. Eur Spine J 15: 391-402. 10.1007/s00586-005-1021-4 [doi].

36. Jacobs WC, Rubinstein SM, Koes B, van Tulder MW, Peul WC (2013) Evidence for surgery in degenerative lumbar spine disorders. Best Pract Res Clin Rheumatol 27: 673-684. S1521-6942(13)00073-9 [pii];10.1016/j.berh.2013.09.009 [doi].

37. Jacobs WC, Rubinstein SM, Willems PC, Moojen WA, Pellise F, Oner CF, Peul WC, van Tulder MW (2013) The evidence on surgical interventions for low back disorders, an overview of systematic reviews. Eur Spine J 22: 1936-1949. 10.1007/s00586-013-2823-4 [doi].

38. [Anonymous] (2014) Sale numbers directly obtained from Kyphon inc sales report 2007.

39. Blumenthal S, McAfee PC, Guyer RD, Hochschuler SH, Geisler FH, Holt RT, Garcia R, Jr., Regan JJ, Ohnmeiss DD (2005) A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. Spine (Phila Pa 1976) 30: 1565-1575. 00007632-200507150-00003 [pii].
40. London A.J. (2006) Cutting Surgical Practices at the Joints: Individuating and assessing Surgical Procedures. In. pp. 19-52.
41. Love JW (1975) Drugs and operations. Some important differences. JAMA 232: 37-38.
42. Barkun JS, Aronson JK, Feldman LS, Maddern GJ, Strasberg SM, Altman DG, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Flum DR, Glasziou P, Marshall JC, McCulloch P, Nicholl J, Reeves BC, Seiler CM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de LM, Deeks J, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J (2009) Evaluation and stages of surgical innovations. Lancet 374: 1089-1096. S0140-6736(09)61083-7 [pii];10.1016/S0140-6736(09)61083-7 [doi].
43. Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, Seiler CM, Altman DG, Aronson JK, Barkun JS, Campbell WB, Cook JA, Feldman LS, Flum DR, Glasziou P, Maddern GJ, Marshall JC, McCulloch P, Nicholl J, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de LM, Deeks J, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J (2009) Challenges in evaluating surgical innovation. Lancet 374: 1097-1104. S0140-6736(09)61086-2 [pii];10.1016/S0140-6736(09)61086-2 [doi].
44. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, Nicholl J, Aronson JK, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Feldman LS, Flum DR, Maddern GJ, Nicholl J, Reeves BC, Seiler CM, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de LM, Deeks J, Ergina PL, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Meakins J, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J (2009) No surgical innovation without evaluation: the IDEAL recommendations. Lancet 374: 1105-1112. S0140-6736(09)61116-8 [pii];10.1016/S0140-6736(09)61116-8 [doi].
45. Stirrat GM (2004) Ethics and evidence based surgery. J Med Ethics 30: 160-165.
46. Galbraith S (1978) The 'no lose' philosophy in medicine. J Med Ethics 4: 61-63.
47. Feigal DW, Gardner SN, McClellan M (2003) Ensuring safe and effective medical devices. N Engl J Med 348: 191-192. 10.1056/NEJMp020170 [doi];348/3/191 [pii].
48. Resnic FS, Normand SL (2012) Postmarketing surveillance of medical devices--filling in the gaps. N Engl J Med 366: 875-877. 10.1056/NEJMp1114865 [doi].
49. Malchau H (2000) Introducing new technology: a stepwise algorithm. Spine (Phila Pa 1976) 25: 285.
50. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI (2005) United States trends in lumbar fusion surgery for degenerative conditions. Spine (Phila Pa 1976) 30: 1441-1445. 00007632-200506150-00016 [pii].



Chapter 10

Summary

Intermittent neurogenic claudication (INC) is a complex of symptoms caused by lumbar spinal stenosis (LSS). Symptoms include leg pain, numbness of (frequently both) the legs and associated back pain. Usually patients will have a few months of conservative treatment before they are offered surgery. For over a century, bony decompression has been the gold standard. However, the present success rate of this gold standard is only 65%. New imaging techniques to improve patient selection and new treatment options to improve direct outcome have been developed. The main goal of new treatment options is to reduce the surgical extent in order to maintain the balance of the spine and to lower the risk of instability after the decompression. (**Chapter 1**) In 1986, a research group proposed a new technique in which an implant was placed between the spinous processes in order to indirectly decompress the spinal nerve roots and to unload the facet joints. Many similar devices have been developed over the past decades, but no biomechanical differences were demonstrated. Thousands of patients have since been implanted with these devices. Efficiency (not cost-efficiency) has only been demonstrated in comparison to conservative treatment. This thesis outlines the results of a double blind multicentre trial (the FELIX trial) in which an interspinous process device (IPDs) was compared with conventional bony decompression in the treatment of INC. Furthermore, we evaluated the diagnostic and prognostic value of MR images in patients with lumbar stenosis and INC.

Chapter 2 describes a national survey held among Dutch spine surgeons of their daily practice of patients with INC caused by LSS. The survey demonstrated a wide variety of Dutch spine surgeons' preferred treatment of symptomatic LSS. To minimize variety, national and international protocols based on high-quality randomized controlled trials (RCT) and systematic reviews are necessary to give surgeons more tools to support everyday decision-making. Furthermore, over 30 percent of Dutch spine surgeons used IPDs (outside study protocol) in treatment for patients with INC.

The existing evidence on interspinous implant surgery was systematically reviewed in **chapter 3**. This systematic review did not include the results of the Felix Trial, because the review was done in 2011 before the publication of the Felix Trial (2013). Two independent RCTs and eight prospective cohort studies on patients with INC caused by LSS were available. In total 563 patients treated with IPDs were included. All studies showed improvement in validated outcome scores (patient satisfaction, disease specific questionnaires and leg pain) after six weeks and one year. Pooled data of the RCTs were more in favor of IPD treatment compared with

conservative treatment. Overall complication rate was 7%. The evidence presented in this systematic review was relatively low and the costs of the IPDs were high, which at that time were already widely used, therefore we needed more thorough (cost-) effectiveness data.

The design of the Felix Trial is presented in **Chapter 4**. The manufacturers of these new implants claimed that treatment with IPDs would give a faster recovery. Therefore, IPDs had to be superior at short-term (eight weeks) compared with (the gold standard) bony decompression. At least 160 patients had to be randomized. To enroll enough patients, fifteen hospitals participated in the study after approval of the protocol by the medical ethics committees.

The 1-year clinical results (presented in **Chapter 5**) did not show any benefit of treatment with IPD over bony decompression. This double-blinded study could not confirm the hypothesized short-term advantage of interspinous process device over conventional "simple" decompression. Furthermore, after 1-year significantly more patients were re-operated in the IPD group (29%) compared with the bony decompression group (8%). The subgroup analysis (**Chapter 5**) could neither assign any subgroup (such as elderly or obese people) which would benefit from treatment with IPDs.

The long-term (2-years) results were similar to the 1-year results (**Chapter 6**). The long-term follow-up did not show important differences in results (based on the ZCQ) comparing treatment with IPD and conventional bony decompression in patients with INC based on LSS. Furthermore, similar to the one-year analysis, the re-operation rate was significantly higher (overall and in the period between 52 and 104 weeks) in the IPD group compared with the bony decompression group. Back pain was hypothesized to be less in the group that underwent an operation with less tissue damage, namely the IPD without bony decompression group. However, to the contrary, the long-term back pain in IPD group was significantly – though not clinically relevant – higher compared with the conventional bony decompression group.

Chapter 7 describes the cost-benefit analysis of the Felix Trial. The economic evaluation showed less favourable QALY outcomes (not significant) after IPD. Healthcare costs were higher for IPD treatment compared with decompression. From a societal perspective the cost of IPD treatment were also higher, though not significant. Implantation of IPD is highly unlikely to be cost-effective compared with bony decompression for patients with INC caused by LSS. The use of these implants over the past three decades could therefore not be justified based on this study.

MRI evaluation (**Chapter 8**) showed good agreement on the clinically relevant level of lumbar spinal stenosis between experienced readers. However, MRI assessment of grading of compression did not correlate with baseline complaints of INC, nor with outcome after one year. However, both absence of epidural fat and presence of tortuous caudal nerves on MR images in patients with LSS were relatively good predictors for satisfactory recovery after surgery. MRI should therefore only be used to appoint the pathological level in case surgical decompression is considered.

Conclusions

The clinical introduction of novel medical devices, such as the introduction of IPDs, often occurs without high quality evidence and with little regulation compared to pharmaceutical interventions. As a consequence the safety, efficacy and long-term effects for patients are insufficiently guaranteed. IPDs heave been used for over thirty years, while good clinical data are becoming available only now. Treatment with IPD did not show any societal benefit though it is responsible for higher health care costs. The controversies surrounding the implementation of IPDs underscore the need to reconsider how innovation in medical devices and surgical techniques should adhere to sound ethical standards without inhibiting surgical research and development (**Chapter 9**). Ethical lessons should be learned from the history in the past thirty years of the introduction of various devices. There is no such thing as the 'no lose' philosophy: patients are harmed by new developments. We suggest that this might be prevented by demanding prospective studies upon device introduction, and by developing (European) registries in order to monitor and evaluate new medical devices.

Hoofdstuk 11

Samenvatting

Neurogene claudicatio klachten (syndroom van Verbiest) is een verzameling van symptomen die veroorzaakt worden door degeneratieve lumbale kanaalstenose (door 'slijtage' versmalling van het kanaal waar de zenuwen doorheen lopen). Met de toenemende levensverwachting komt deze degeneratieve ziekte ook navenant meer voor: het is nu in West-Europa de ziekte met de meeste zorgbelasting voor de maatschappij. Het syndroom bestaat meestal uit beenpijn (meestal beide benen), doofheid (onstabiel gevoel) in beide benen en regelmatig ook geassocieerde rugpijn. vanaf de eerdere publicaties van de vorige eeuw wordt benige decompressie van het bot als gouden standaard gezien: verwijderen van de dorsale zijde (lamen/ boog) van het wervellichaam. Echter, deze gouden standaard heeft slechts een succespercentage van 65% na behandeling. In de afgelopen decennia zijn er voor de patiënt minder ingrijpende behandelingsmethoden ontwikkeld voor de patiënt. In 1986 (zie **hoofdstuk 1**) werd er door een onderzoeksgroep een nieuwe behandeling beschreven. Tussen de twee processus spinosi (achterste gedeelte van de wervel) van wervels plaatsten zij een implantaat dat de wervels 'uit elkaar duwt'. Hierdoor wordt er op een gerichte plek indirect meer ruimte gegeven aan de zenuwen die door het wervelkanaal lopen. Daarnaast wordt met dit implantaat volgens de hypothese ook de gewrichten (facetten) ontlast. Daarna zijn er meerdere implantaten (in samenwerking met industrie) ontwikkeld. Deze implantaten blijken in biomechanische studies precies dezelfde eigenschappen te hebben. Mondiaal hebben honderdduizenden patiënten een dergelijk implantaat gekregen. Voor dit proefschrift waren bedoelde implantaten niet getest op (kosten)effectiviteit ten opzichte van de eerder beschreven gouden standaard. Dit proefschrift beschrijft de resultaten van onderzoek (FELIX studie) waarin behandeling met een implantaat (IPD) werd vergeleken met benige decompressie, waarbij patiënten via loting een behandeling toegewezen kregen (randomisatie). Patiënten en gespecialiseerde verpleegkundigen die de patiënten vergeleken en onderzochten wisten niet welke therapie, IPD of benige decompressie, de patiënt gekregen had (dubbel blind onderzoek). Daarnaast worden in dit proefschrift de resultaten van de MRI (waarop de mate van versmalling van het kanaal bepaald werd) gepresenteerd.
In **hoofdstuk 2** worden de resultaten beschreven van een enquête die in 2011 is ingevuld door leden van de Nederlandse vereniging van de wervelkolom geïnteresseerde chirurgen met hun ervaringen met het Verbiest syndroom. In 2011 werden er in Nederland zeer veel verschillende behandelingstechnieken gebruikt. Bovendien bleek dat 30% van de wervelkolom centra deze IPDs gebruikte buiten studies in 2011 (dus voor de publicatie van de FELIX studie). Deze aangetoonde variëteit

(praktijkvariatie) vraagt om meer protocollen en meer kwalitatief goede vergelijkende studies van de verschillende behandelingstechnieken die gebruikt worden.

In **hoofdstuk 3** worden de resultaten van een systematische literatuurstudie gepresenteerd (van voor deze studie). In de periode tot 2011 waren er twee vergelijkende studies waarin IPD behandelingen werden vergeleken met conservatieve behandeling (geen operatie of injecties) en er was een achttal beschrijvende studies die een groep patiënten (behandeld met een IPD) volgden door de tijd. Er was over het geheel een 7% kans op complicaties. Behandeling met IPD was inderdaad beter dan geen behandeling, maar vergelijkende studies met de gouden standaard waren er tot dusverre niet.

In **hoofdstuk 4** wordt het protocol van de FELIX studie besproken. De veronderstelling bij de IPD behandelmethode is dat zij vooral sneller een beter resultaat geven, want een IPD implantatie zou een minder ingrijpende operatie zijn. Daarom gaat het onderzoek er vanuit dat de patiënten die loten voor IPD een beter herstel hebben na acht weken in vergelijking met de patiënten die loten voor de standaard benige decompressie. Om een verschil tussen beide behandelingsmethodes betrouwbaar te kunnen aantonen hebben we een onderzoekspopulatie van 160 patiënten nodig met klachten van neurogene claudicatio en aangetoonde lumbale kanaalstenose op de MRI. De studiepatiënten werden twee jaar vervolgd. De studie werd goedgekeurd door de ethische commissie van de vijftien deelnemende ziekenhuizen.

De resultaten na één jaar laten geen positiever effect zien van behandeling met een IPD ten opzichte van de behandeling met de benige decompressie (**hoofdstuk 5**). Evenmin werd met deze dubbelblinde studie het aangenomen voordeel van behandeling met een IPD ten opzichte van behandeling met benige decompressie aangetoond. Bovendien was het heroperatiepercentage veel hoger in de IPD groep (29% werd opnieuw geopereerd) ten opzichte van de benige decompressie groep (slechts 8% kreeg een tweede operatie). Ook analyse op vooraf bepaalde factoren als leeftijd gaf geen subgroep waarbij behandeling met een IPD een betere resultaat opleverde dan de gouden standaard (benige decompressie).

Op langere termijn ware de resultaten vergelijkbaar (**hoofdstuk 6**). Er was geen verschil tussen beide groepen. Bovendien was het aantal heroperaties in de IPD groep toegenomen. De veronderstelling dat de implantaten minder druk op de facetten zouden geven en daarmee ook minder rugpijn op de langere termijn ten opzichte van benige decompressie, bleek niet op te gaan. In de IPD groep was de rugpijn (na twee jaar) zelfs heftiger dan in de benige decompressiegroep.

Ook het verschil in kosteneffectiviteit is in de analyse beoordeeld (**hoofdstuk 7**). De IPD groep had een lagere QALY (maat voor kwaliteit uitgedrukt in geld per gezond levensjaar) in vergelijking tot de benige decompressiegroep. De kosten voor de gezondheidszorg zijn ook groter in de IPD groep en van een maatschappelijk perspectief waren de kosten (niet significant) hoger. Daarom valt niet aan te nemen dat het implanteren van IPD kosteneffectief is, dit nog los van de prijs die de industrie voor een dergelijk implantaat vraagt. Het gebruik van implantaten is daarom niet te verantwoorden.
De uitslagen van de MRIs (**hoofdstuk 8)** toonde een vergelijkbare uitkomst tussen de beoordelaars (goede correlatie). Echter, de mate van compressie vertoonde geen correlatie met de mate van stenose op het moment dat de MRI gemaakt werd. Bovendien is de mate van stenose ook geen voorspeller voor de mate van herstel na een eventuele operatie. Alleen als er bij patiënten op de MRI een zeer nauw kanaal (de zenuwen liggen 'opgekruld' of al het normaliter aanwezige vet om de zenuwen is verdwenen) zichtbaar is, dan kan deze bevinding een goede uitkomst na chirurgische behandeling voorspellen.

Conclusie:

Sinds de tweede helft van de jaren '80 zijn medische implantaten, zoals IPDs en PIP borstimplantaten, als behandelmethode geïntroduceerd, echter, zonder wetenschappelijk overtuigende onderbouwing dat deze behandelmethode betere resultaten opleverde. Er is overigens weinig (internationaal) toezicht op de introductie van implantaten in het algemeen. Hierdoor is vaak onduidelijk voor de clinicus en burger hoe veilig en/of effectief een implantaat is. IPDs zijn meer dan 30 jaar zonder goed wetenschappelijk bewijs geïmplanteerd bij patiënten over de gehele wereld. Het onderzoek in dit proefschrift levert geen wetenschappelijk onderbouwing voor de veronderstelling dat gebruik van IPDs een superieur behandelresultaat oplevert ten opzichte van de 'benige decompressie' behandelmethode. Verder maakt dit proefschrift duidelijk dat gebruik van IPDs als behandelmethode alleen maar kostbaar is geweest voor de gezondheidszorg en de maatschappij. Als beroepsgroep moeten de medische stand zorgen dat de implementatie van nieuwe implantaten en technieken veiliger wordt, zonder dat vooruitgang wordt belemmerd (**hoofdstuk 9)**. Ethisch moeten we ook hier ook bij stilstaan, want bij gebruik van behandelmethodes geldt niet: "baat het

niet, schaadt het niet". Uitkomst van dit proefschrift is om vóórafgaand aan de introductie van een implantaat te eisen dat prospectieve studies aantonen dat het te introduceren implantaat positieve resultaten oplevert. Daarnaast adviseren wij het instellen van (Europese) registraties om na introductie een implantaat de lange termijneffecten te volgen. Daarnaast moet de wetenschappelijke gemeenschap ook van onderzoeksgroepen eisen dat deze eerst vergelijkende studies doen, voordat experimentele therapie wordt toegepast op patiënten.

List of publications

1. Moojen WA, Schenck CD, Lycklama à Nijeholt GJ, Jacobs WCH, Van der Kallen BF, Arts MP, Peul WC, Vleggeert-Lankamp CLAM. Preoperative MR imaging in patients with intermittent neurogenic claudication: relevance for diagnosis and prognosis. 2014. Submitted.
2. Moojen WA, Bredenoord AL, Viergever RF, Peul WC. Scientific evaluation of spinal implants; an ethical necessity. 2014. Submitted.
3. Van den Akker-van Marle ME, Moojen WA, Arts MP, Vleggeert-Lankamp CL, Peul WC. Interspinous Process Devices without bony decompression versus Standard Conventional Surgical Decompression for Lumbar Spinal Stenosis: Cost Utility Analysis. Spine J. 2014.
4. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW, Vleggeert-Lankamp CL, Peul WC. IPD without bony decompression versus conventional surgical decompression for lumbar spinal stenosis: 2-year results of a double-blind randomized controlled trial. Eur Spine J. 2014.
5. Overdevest GM, Moojen WA, Arts MP, Vleggeert-Lankamp CL, Jacobs WC, Peul WC. Management of lumbar spinal stenosis: a survey among Dutch spine surgeons. Acta Neurochir (Wien). 2014 Aug 7.
6. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW, Vleggeert-Lankamp CL, Peul WC; Leiden-The Hague Spine Intervention Prognostic Study Group. BMJ. 2013 Nov 14;347:f6415. Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial.
7. Jacobs WC, Rubinstein SM, Willems PC, Moojen WA, Pellisé F, Öner FC, Peul WC, van Tulder MW. Eur Spine J. 2014 Feb;23(2):481-2. Answer to the letter to the editor of F. M. Kovacs et al. entitled "overviews should meet the methodological standards of systematic reviews" concerning "the evidence on surgical interventions for low back disorders, an overview of systematic reviews" by Wilco C.H. Jacobs et al.; Eur Spine J, doi:10.1007/s00586-013-2823-4.
8. Jacobs WC, Kruyt MC, Moojen WA, Verbout AJ, Oner FC. J Clin Epidemiol. 2013 Dec;66(12):1347-55.e3. No evidence for intervention-dependent influence of methodological features on treatment effect.
9. Jacobs WC, Rubinstein SM, Willems PC, Moojen WA, Pellisé F, Oner CF, Peul WC, van Tulder MW. The evidence on surgical interventions for low back disorders, an overview of systematic reviews. Eur Spine J. 2013 Sep;22(9):1936-49.

10. Bartels RH, van Tulder MW, Moojen WA, Arts MP, Peul WC. Laminoplasty and laminectomy for cervical sponydylotic myelopathy: a systematic review. Eur Spine J. 2013 Apr 11.
11. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC. Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. Eur Spine J. 2011 Oct;20(10):1596-606. doi: 10.1007/s00586-011-1873-8. Epub 2011 Jun 11. Review.
12. Moojen WA, Vleggeert-Lankamp CL, Krol AD, Dijkstra SP. Spine (Phila Pa 1976). Long-term results: adjuvant radiotherapy in en bloc resection of sacrococcygeal chordoma is advisable. 2011 May 1;36(10):E656-61.
13. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord. 2010 May 27;11:100. doi: 10.1186/1471-2474-11-100.
14. Boone J, Schipper ME, Moojen WA, Borel Rinkes IH, Cromheecke GJ, van Hillegersberg R. Robot-assisted thoracoscopic oesophagectomy for cancer. Br J Surg. 2009 Aug;96(8):878-86. doi: 10.1002/bjs.6647.

Curriculum vitae

Wouter Anton Moojen was born on the 27th of June, 1983 in Utrecht, The Netherlands. He grew up in the village Groenekan. He obtained his Athenaeum diploma (pre-university diploma) cum laude in Utrecht 2001. In 2001 he started medical school at the Utrecht University and received his medical doctor degree in 2007. In the same year he started working as a house officer at the University Medical Center Utrecht (dr. J.W. Berkelbach van der Sprenkel). His interests in vascular/skull base surgery and spine surgery (dr. G.J. Amelink and Professor F.C. Öner) started. After six months he started at the Leiden University Medical Center. In the summer of 2008 he visited Toronto Western Hospital (Professor Fehlings) where his interest for spine research was triggered. In December 2008 his neurosurgery residency program (J.HC. Voormolen and Professor W.C. Peul) commenced. The patient inclusion for Felix Trial in Leiden University Medical Center and the Medical Center Haaglanden started in 2009. In 2009 he was trained in neurology at the Leiden University Medical Center (Professor R.A.C. Roos), in radiology at the Leiden University Medical Center (Professor M.A. van Buchem), in intensive care medicine (Professor E. de Jonge), pathology (dr. S.G. van Duinen), and in thoracic surgery (Professor R.J.M. Klautz). In 2011 and in 2013 he worked in total for 15 months at the neurosurgical departments in the Medical Center Haaglanden (R. Walchenbach) in The Hague and for three months in Erasmus Medical Center, Rotterdam (Professor C.M.F. Dirven). In 2012 he finished his epidemiology training (MSc) at the VU University Amsterdam. In 2014 he visited two neurosurgical centers abroad as part of a travelling fellowship: Helsinki (Professor J. Hernesniemi) in July and August, and Berlin (Professor P. Vajkoczy) in September and October. He will finish his neurosurgery residency at the end of 2014. His areas of interest in neurosurgery are spine, vascular and skull-base surgery.

Acknowledgements

The following persons made it possible for me to write this thesis:

Elske van den Akker-van Marle, Sandra Annokkee, Hans Amelink, Annelien Bredenoord, Mark Arts, Ronald Bartels, Elmy van Beelen, Irene van Beelen, Janneke van Beijnum, Nanette van den Bersselaar, Jolanda Blom-de Vreugd, Ralph de Boer, Jannie van Duijn, Saskia Dukker, Nathalie den Dulk, Niels van der Gaag, Anneke Gaasbeek, Mirjam de Gooijer, Els den Hollander-Duym, Wilco Jacobs, Bas van de Kallen, Ingrid van Kessel, Fred Kloet, Bart Koes, Brenda Kromhout-Hogewoning, Jolanda Lavue, Geert Lycklama à Nijeholt, Martijn Malessy, Claudine Nogarede, Marjon Nuijten, Cumhur Öner, Marina Oosterhuis, Wilma Ouwehand, Wilco Peul, Willem Pondaag, Michelle de Raaf, Petra Robbers, Godard de Ruiter, Catharina Schenck, Lidwien Smakman-Hageman, Jan Schoones, Brenda Schouten-Le Maître, Pieter Schutte, Anouk Streede-Lamme, Adriaan Taselaar, Maurits van Tulder, Jos Twisk, Ditte Varkevisser-Miedema, Hanneke Verheijde, Marco Verstegen, Rik Viergever, Joice Videler-Pitlo, Tjitske Visser, Jacqueline Vogelaar-van Welie, Johan Voormolen, Yvon van der Voort-Warmerdam, Christie Waanders, Rob Walchenbach, Bonny Wennekers, Peter Willems, Jasper Wolfs, Hans Wurzer, Robin van Zijl.

The following organisations made it possible for me to write and print this thesis:
- Spine Intervention Prognostic Study Group
- Dutch Spine Society
- DePuySynthes companies of Johnson&Johnson
- InSpine
- Paradigm Spine
- BIOMET
- ChipSoft
- Implantcast Benelux

This PhD Thesis is also digital viewable at:
http://www.my-thesis.nl/Moojen/, password: PDFMoojen

http://www.my-thesis.nl/Moojen/
password: PDFMoojen