Thesis for doctoral degree (Ph.D.) 2008

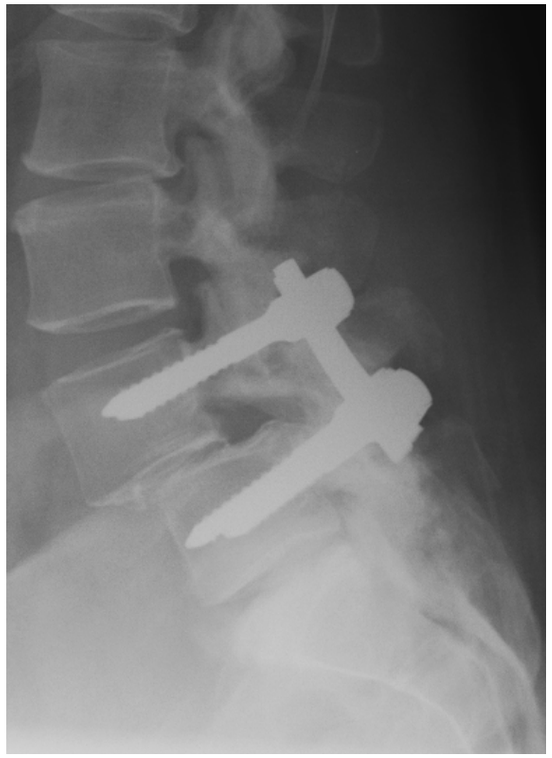
Lumbar Fusion For Chronic Low-back Pain In Isthmic Spondylolisthesis

Per Ekman

Karolinska Institutet

Thesis for doctoral degree (Ph.D.)
2008

LUMBAR FUSION FOR CHRONIC LOW-BACK PAIN IN ISTHMIC SPONDYLOLISTHESIS



Per Ekman

Karolinska Institutet

From
Department of Clinical Science, Intervention and Technology (CLINTEC), Division of Orthopedics, Karolinska University Hospital Huddinge, Karolinska Institutet
And
Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet
Stockholm, Sweden

LUMBAR FUSION FOR CHRONIC LOW-BACK PAIN IN ISTHMIC SPONDYLOLISTHESIS

Per Ekman

Karolinska Institutet

Stockholm 2008

Published by Karolinska Institutet. Printed by [larserics]

ISBN 978-91-7409-259-2

ABSTRACT

Manifestation of isthmic spondylolisthesis (IS), with a prevalence of 5% in the general population, varies from totally asymptomatic to severe disability. Although fusion has positive short-term effects, the long-term outcome of such treatment, as well as possible accelerated degeneration of adjacent segment discs (ASD) and its clinical significance, are unknown. The primary objective of the present investigation was to provide this missing information. Additional aims were to identify factors of value in predicting the outcome of fusion with the ultimate goal of improving patient selection, and to determine the medium-term outcome of more extensive fusion including the anterior vertebral body, i.e., Posterior Lumbar Interbody Fusion (PLIF).

For all patients, the inclusion criteria were absence of previous spinal surgery, an age of 18-55 years, and at least 12 months of disabling symptoms due to IS. In a long-term randomized controlled trial (RCT), 111 patients were treated with a one-year exercise program (EX, n=34) or posterolateral fusion (PLF, n=77), with (n=37) or without (n=40) pedicle screw instrumentation. In a prospective investigation the outcome for 86 patients subjected to PLIF was assessed after two-year follow-up on the basis of the Pain Index (VAS), Disability Rating Index (DRI), Oswestry Disability Index (ODI), work status, and global self-assessment. In addition, quality of life was assessed with the SF-36 questionnaire; preoperative pain drawings (PD) obtained; and the long-term development of ASD quantified. In the prospective PLIF study, the two-year follow-up rate was 98%, while in the RCT, the 9- and 13-year rates, including radiographs, were 91% and 72%, respectively.

The two-year outcome for the PLIF group was virtually identical to the short-term outcome for the PLF patients in our previous RCT, although a larger number of major complications occurred following PLIF (12 (14%) versus 4 (5%); p=0.06). 54% of the patients whose PD were "organic"(O) and 33% of those with a "non-organic"(NO) PD rated themselves as “much better” (p=0.038) and, moreover, the mean pain index and DRI and ODI values for the O-group were also significantly better. Multivariate analysis revealed that working actively, male gender and regular exercise were associated with a more favourable outcome, but these factors could account for no more than 20% of the observed variability in outcome.

In the RCT involving patients treated by fusion or with exercise, 76% of the PLF group but only 50% of the EX patients rated their overall outcome as “much better” or “better”(p=0.015) after a mean of 9 years. All other assessments of outcome favoured the surgical group, but the differences were not statistically significant.

After a mean follow-up of 13 years, disc height was reduced by 2% in EX and 15% in PLF patients (p=0.0016). According to the UCLA grading scale, the discs of all patients in the EX group, but only 62% of the PLF group were normal (p=0.026). The only significant difference in outcome for PLF patients with and without ASD was the more favourable self-reported global outcome for the latter. No significant differences in long-term outcome or frequency of radiologically verifiable ASD in instrumented and non-instrumented patients

were observed. Finally, laminectomised patients developed ASD more often than those not laminectomised ($p=0.015$).

We conclude that treatment of IS patients with either PLF or PLIF results in similar short-term outcomes. PLF, with or without pedicle screw fixation, is associated with a long-term outcome that is modestly better than that obtained by natural healing processes. This study also shows that fusion, particularly in combination with laminectomy, accelerates degenerative changes at the adjacent level, with a consequent minor negative effect on outcome.

LIST OF PUBLICATIONS

I. The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study

Per Ekman, Hans Möller, Rune Hedlund

The Spine Journal 5 (2005) 36–44

II. Posterior lumbar interbody fusion versus posterolateral fusion in adult isthmic spondylolisthesis

Per Ekman, Hans Möller, Tycho Tullberg, Pavel Neumann, Rune Hedlund

Spine 32, (2007), 2178–2183

III. Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis

Per Ekman, Hans Möller, Rune Hedlund

Submitted for publication

IV. A prospective randomized study on the long-term effect of lumbar fusion on adjacent disc degeneration

Per Ekman, Hans Möller, Adel Shalabi, Yiang Xiao Yu, Rune Hedlund

Submitted for publication

TABLE OF CONTENTS

1.1 Introduction 1
1.1.1 General background 1
1.1.2 Classification of spondylolisthesis 2
1.1.3 The etiology of isthmic spondylolisthesis 3
1.1.4 Clinical characteristics 3
1.1.5 Treatment options for patients with isthmic spondylolisthesis/CLBP 4
1.2 Aims of the thesis 6
1.3 Assessment of Outcome 7
1.3.1 Subjective symptoms 7
1.3.2 Fusion 8
1.3.3 Degeneration 9
1.4 The Study populations and methods of approach 10
1.4.1 Criteria for inclusion and exclusion 10
1.4.2 Study population 1 10
1.4.3 Study population 2 12
1.4.4 Surgical treatment 14
1.4.5 Conservative treatment 14
1.4.6 Asessments 15
1.5 Statistical Analyses 17
1.6 Ethical Considerations 17
1.7 Summary of the findings 18
1.7.1 Paper 1 18
The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study 18
1.7.2 Paper 2 22
Posterior Lumbar Interbody Fusion versus Posterolateral Fusion in Adult Isthmic Spondylolisthesis 22
1.7.3 Paper 3 24
Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis 24
1.7.4 Paper 4 29
A Prospective randomized study on the long-term effect of lumbar fusion on adjacent disc degeneration. 29
1.8 General discussion 34
1.9 Conclusions 40
2 Acknowledgements 41
3 Summary in Swedish 42
3.1 Steloperation mot kronisk ländryggsvärk vid kotförskjutning av Isthmisk typ 42
4 References 44
5 Papers 1 – 4 54
5.1 ERRATA concerning Papers 1 and 2 54

LIST OF ABBREVIATIONS

"360"	Global or combined posterior and anterior fusion
ALIF	Anterior lumbar interbody fusion
ASD	Adjacent segment degeneration/disease
CDI	Cotrel-Dubousset instrumentation
CLBP	Chronic low-back pain
CT	Computer tomography
DDD	Degenerative disc degeneration/disease
DRI	Disability rating index
FDA	Food and Drug Administration
ICC	The Shrout-Fleiss intraclass correlation coefficient
IS	Isthmic spondylolisthesis
MRI	Magnetic resonance imaging
ns	Not statistically significant
ODI	Oswestry disability index
PI	Pain Index
PLIF	Posterior lumbar interbody fusion
PLF	Posterolateral fusion
QMA™	Quantitative motion Analysis
RCT	Randomized controlled trial
SD	Standard deviation
SF-12	Medical outcomes study: Short Form survey, 12 items
SF-36	Medical outcomes study: Short Form survey, 36 items
VAS	Visual analogue scale

1.1 INTRODUCTION

1.1.1 General background

Chronic low back pain (CLBP) affects the physical function and quality of life of a large number of individuals. Indeed, "back-problems" constitute one of the largest medical and socioeconomic problems in the world today[4], with the costs connected with Chronic LBP alone representing 1.7% of the gross national product of European countries during the mid-1990´s[115,173]. Fusion of the spine has become a more and more common approach to treatment of a wide variety of back disorders during the past 50 years[30,37,83,160] not only for treatment of fractures, infections and deformities, but, in particular, for degenerative disorders such as degenerative disc disease (DDD).

One cause of back pain, Spondylolisthesis, or the slipping of a vertebra, occurs in approximately 5% of the population[9,45,71,122,132,185] and more commonly in men[45,185]. Back pain associated with this condition is sometimes accompanied by sciatica resulting, in restrictions in everyday life, including an inability to, e.g., exercise or work. Isthmic sponylolisthesis (IS) a vertebral slip due to a defect in the *pars interarticularis* which connects the lamina with the vertebral body, develops most often during childhood and can give rise to considerable problems, albeit of a different nature, in both children and adults. In the case of adult patients, who have been studied here, this disorder can be treated successfully with fusion surgery, at least in the short term[80,108]. The Swedish lumbar spine study revealed that such fusion also produces similar positive results in patients with DDD[48]. However, in both groups improvement was modest; many patients reported only a limited reduction in pain and disability, indicating a requirement for improved treatment strategies, including better fusion procedures, and/or more careful selection of patients for treatment by fusion.

At present, posterolateral fusion (PLF) is still considered to be "the gold standard" for fusion of the degenerative spine. However, a wide variety of other fusion procedures including pedicle screw fixation, anterior lumbar interbody fusion (ALIF) or "360°-fusion" e.g., PLF in combination with ALIF or PLIF offer at least in theory, advantages over PLF alone. To date, no significant differences in the outcomes obtained with different fusion procedures has been documented[80,89], with the exception of a single recent Danish study which reported a more favourable outcome with "360°-fusion"[177].

At the same time, slightly higher rate of successful fusion, although without a better clinical outcome, has been reported in connection with instrument-guided fusion[44,56,57]. In fact the randomised controlled trials on CLBP documented so far provide only limited evidence that surgical intervention in the spine is more favourable than various conservative treatments[18,19,41,48,105,147]. Nonetheless, spurred on by new techniques and commercial considerations, the frequency of spinal fusions has increased dramatically[130]. Moreover, complex fusion procedures with potentially more complications have been introduced in hope of improving fusion rates and attaining a better outcome with respect to pain and function. However, the evidence in favour of more sophisticated approaches such as interbody fusion is, weak and fusion of the degenerative lumbar spine remains controversial[112,113,130].

None of the few prospective, randomized studies on the long-term effect of lumbar fusion[2,177] has included control patients who have not undergone operation. In addition,

the natural development of chronic low-back pain upon ageing is poorly documented. Consequently, the long-term effect of fusion is difficult to assess, and evaluation of surgical treatment without appropriate controls cannot provide evidence for the efficacy of such treatments[61,130,171].

Theoretically, fusion of a portion of the spine will increase the mechanical load on adjacent spinal segments and, indeed, such an effect has been demonstrated in numerous studies[6,21,27,31,38,59,62,93,120,127,175]. Such an elevated non-physiological load might in principle, induce degenerative changes adjacent to the fusion and thereby worsen the long-term outcome. Although accelerated adjacent segment degeneration (ASD) is occasionally observed, it is not known whether this is a consequence of the fusion and/or normal aging processes. Furthermore, the extent to which such a radiological finding reflects a patient's actual symptoms is unclear. In spite of these uncertainties, a wide variety of spinal procedures or so-called motion preservation techniques-- most notably the disc prosthesis, but also semi-rigid or dynamic instrumentation techniques-- are being employed because of their theoretical advantage in reducing the risk for long-term ASD.

In 1989, in attempt to determine the outcome of fusion following operation, we began a randomized controlled trial designed to compare surgical and conservative treatment of patients suffering from adult isthmic spondylolisthesis, a well-defined and homogenous study population[108]. Almost 10 years later, we considered such patients to represent an appropriate group with which to address the controversies described above. In this connection we have generally followed the guidelines for high-quality clinical studies, established by the Cochrane review group in 2003[81,171]. As will be described here 20 years after the first long-term investigation was initiated, we now have the ability to determine not only the natural long-term course of adult isthmic spondylolisthesis, but also to evaluate the influence of surgery on both the clinical and radiological outcomes, including ASD. The primary null hypothesis tested by the present study is that in the long-term, fusion neither affects this natural course nor accelerates adjacent segment degeneration.

1.1.2 Classification of spondylolisthesis

Most commonly, spondylolisthesis is classified into five types on the basis of anatomy and etiology: dysplastic (due to dysplasia), isthmic (due to a lesion in the *pars interarticularis*) , degenerative (degeneration of disc and facet joints), traumatic (caused by fractures in structures other than pars interarticularis) and pathological (resulting from bone disease)[179]. The isthmic variety is further subdivided into three subgroups, i.e., due to lytic-fatigue fracture of the pars (type A), elongated, but intact pars (B) or acute pars fracture (C). An additional subgroup, i.e., iatrogenic (due to surgical trauma) has subsequently been proposed[182] . The patients involved in the present investigations suffered from isthmic spondylolisthesis of the lytic (A) or dysplastic (B) type.

An alternative classification developed by Marchetti- Bartolozzi consists of a developmental group (with high- and low- dysplastic subgroups) and an acquired group (with traumatic, post-surgery, pathological and degenerative subgroups)[102]. Moreover, in 2005 Herman and Pizzutillo suggested a new classification for children and

adolescents involving four major groups, i.e., dysplastic, developmental, traumatic and pathologic[73].

The degree of slippage is commonly quantified according to Meyerding, with grade 1 being less than 25%, grade 2 25-50%, grade 3 50-75% and grade 4 75-100% slippage of the sagittal vertebral diameter. Spondyloptosis, the complete dislocation of a vertebra, is referred to as grade 5 slippage[104].

The condition involving a defect in the pars interarticularis is reffered to as spondylolysis and can be present without slippage, in which case it is often classified as Meyerding grade 0. In the present study patients with all grades of slippage were included.

1.1.3 The etiology of isthmic spondylolisthesis

The causes of IS are considered to be multifactorial and include trauma, mechanical stress and hereditary factors. The high incidence (30%) of isthmic or dysplastic spondylolisthesis among first-degree relatives, along with its association with *spina bifida occulta* indicate that there is a hereditary pre-disposition for this defect[45,69,187]. At the same time, this condition is, not found in stillborn babies[134] or in patients who have been non-ambulatory or bedridden for their entire lives, strongly suggesting that mechanical factors are also involved[133]. Furthermore, the incidences among athletes and individuals experiencing repeated stress to the lumbar back are elevated, lending support to the proposal that spondylolysis and isthmic spondylolisthesis result from a fatigue fracture caused by activities associated with ambulation[159,183]. Thus, most likely and not unexpectedly, both hereditary and mechanical parameters are major risk factors for developing IS.

1.1.4 Clinical characteristics

Approximately 50% of patients with spondylolysis (pars defects) develop vertebral body subluxation[53,184], most often with low grade slippage (Meyerding grade 1). In 90% of such cases, the slippage is located at L5 in the lower lumbar spine and in 5% of the cases at L4[45,53]. Slippage at more cranial levels or more than one level is uncommon and reports on such conditions are therefore rare[39,108].

Surprisingly, no clear-cut relationship between IS and LBP has been established. Fredrickson and co-workers and, later, Beutler and collegues described 30 individuals, among 500 children in their first year of school who had pars lesions, but SF-36 scores identical to those age-matched unaffected individuals 45 years later[9,45,109]. These observations are supported by the finding of neurological findings in these patients similar to those in individuals with non-specific CLBP, as well as the lack of an elevated prevalence of IS amongt 952 subjects receiving disability pension[46,109].

In patients with a slippage that is symptomatic, the most common symptoms include low-back and radicular pain, both of which are more common in adults than in children[25,33,66,72]. In severe cases of high-grade spondylolisthesis, other symptoms may include disturbance in gait, tightness of the hamstrings, hyperlordosis of the lumbar spine and neurological dysfunction[16] and usually debut during the adolescence. However certain individuals with high-grade slippage, remain asymptomatic[68].

The symptoms, signs and clinically relevant problems experienced by adult patients with low-grade spondylolisthesis are otherwise indistinguishable from those patients with CLBP of non-specific origin or caused by DDD. This is consistent with the well known fact that IS accelerates DDD at the level of slippage[138,139,158]. Therefore, adult IS can be viewed as a radiologically verifiable disorder associated with low-back pain and accelerated disc degeneration[109].

1.1.5 Treatment options for patients with isthmic spondylolisthesis/CLBP

Since the symptoms associated with IS are usually mild and spontaneous improvement may occur, initial treatment is usually conservative. Such treatment is highly similar to that prescribed for non-specific CLBP, e.g., education concerning how one's back functions and can be used without undue strain, medication, exercise therapy, epidural injections of corticosteroids and electrical stimulation of transcutaneous nerves [147]. Certain reports have shown that use of a brace can be beneficial in treating IS[8,10,155]. Exercise were not found to have beneficial mid-term effects on IS in our earlier randomised controlled study[108]. O´Sullivan, however, observed a positive effect of stabilization training[116].

A wide variety of surgical approaches have been described including techniques for direct repair of the lytic lesion[20,110,114,151]. The rationale for such direct repair is preservation of motion and has been proposed for younger patients with no disc degeneration. However, in a long term study Schlenzka and co-workers reported that direct repair attenuates the mobility of the lumbar spine, does not prevent disc degeneration, and, most importantly, leads to a worse outcome than that obtained with fusion[140]. In the case of adult patients with degenerative changes, direct repair of the pars is not recommended.

Decompression alone (i.e., Gill procedure) have been reported to result in suboptimal results[53,89,121]. Moreover, Carragee and collegues have reported that in cases of IS, fusion in combination with decompression can enhance pseudartrosis and lead to results that are less satisfactory than with fusion alone[23]. With the exception of disc or nucleus prostheses, which are only used with patients suffering from DDD, and direct repair for IS, the same surgical approaches are generally employed for the treatment of DDD and IS.

According to the Swedish lumbar spine registry, the frequency with which different surgical procedures were used to treat spondylolisthesis in our country during 2003 were as follows: decompression together with instrumented fusion was used in 35% of cases; posterolateral instrumented fusion in 18%; anterior lumbar interbody fusion (ALIF), with or without pedicular instrumentation in 14%; decompression and posterolateral uninstrumented fusion in 9%; uninstrumented fusion alone in 9%; PLIF, with or without titanium implants, in 9%; and in a few cases, decompression alone[156].

Despite the large number of different fusion techniques that are available, several studies indicate that the simpliest procedure, posterolateral fusion (PLF) in situ, is as effective as more complicated methods, as well as producing fewer undesirable side-effects[56,57,122].

As has also been reported to be the case for DDD, most investigations on treatment of IS with instrumented fusion, anterior fusion or 360°-fusion have shown no

improvement in outcome over that obtained with PLF[89]. A recent Danish study on patients with DDD, who randomly recieved circumferential fusion or instrumented PLF revealed no statistically significant difference two years after operation. However, more long-term, the outcome for the patients treated by circumferential fusion was significantly better although subgroup analysis demonstrated that this was not the case for the patients with IS[177].

The theoretical advantages that anterior fusion and the PLIF technique were proposed earlier to have over PLF include support of the anterior column, indirect foraminal decompression, restoration of lordosis and reduction of the slippage via ligamentotaxis, and removal of the degenerated disc supposed to generate pain[7,94,111,157]. In cases of high-grade spondylolisthesis (Meyerding grades 3-5) 360°-fusion is recommended by many investigators since performance of PLF alone may result in progressive kyphosis[34,70,143]. Although certain authors recommend reduction of the slipped vertebra in patients with high-grade spondylolisthesis, this recommendation remains controversial, since no clear improvement in outcome has been demonstrated and complications are significantly more frequent[16,33,163]. For spondyloptosis (Meyerding grade 5), vertebrectomy of L5 has been reported to result in an acceptable outcome despite a high rate of complications[52]. In general, because relatively few patients have high-grade slippage and investigations of high quality are lacking, there is at present no general agreement regarding treatment of this group of patients[164].

1.2 AIMS OF THE THESIS

The ultimate goal of the present investigation was to improve the outcome of treatment for adult patients with IS. This goal was approached by specifically examining the following:

1. the long-term outcome of treating adult patients with isthmic spondylolisthesis, accompanied by low-back pain with or without sciatica, with the "gold standard" procedure, posterolateral fusion, with or without instrumentation;

2. the mid-term outcome of treating such patients with a more extensive fusion procedure, the PLIF;

3. factors that might be of value in predicting the outcome following fusion in adults with isthmic spondylolisthesis; and

4. the possibility that fusion accelerates long-term degenerative changes at the adjacent level and, if so, whether acceleration influences the long-term outcome.

1.3 ASSESSMENT OF OUTCOME

1.3.1 Subjective symptoms

Treatment outcomes are often difficult to assess, especially in the case of disorders such as IS, where subjective symptoms constitute the major clinical problem. It is well known that subjective assessments produce results that are apparently better than those obtained by objective measurement. Furthermore, the design of a questionnaire may influence the outcome obtained[75].

This situation is reflected in the multitude of different of questionnaires employed to study of LBP. The visual analogue scale (VAS) has been in use in medical sciences since the 1960's and was shown to provide a reliable assessment of pain by Huskisson a decade later[76]. A variety of different scoring systems are applied to the locomotor apparatus, including many specifically designed for low-back disorders, e.g., more than 20 health-related quality-of-life rating scales that have been validated in accordance with current methodological standards[188].

The questionnaires used in the present study inlude the Oswestry Disability Index (ODI)[42] and the Short Form 36 (SF-36)[174], the presently most generally accepted and widely used in connection with clinical trials involving the spine today. The ODI, a validated disease-specific instrument that assesses disability associated with spinal disorders. It consists of 10-items with six response alternatives for each item. The items evaluated are the intensity of pain; personal care; the ability to lift; walk, sit, stand and sleep; quality of sex life; social interactions; and the ability to travel. Normal functioning is rated as 0 and the most severe disability for each item is scored as 5. The sum of the scores for all 10 items is multiplied by 2 to give an ODI of 0–100, with 100 representing the highest level of disability.

In the study we also included the disability rating index (DRI), a validated disease-specific or so-called domain-specific instrument for assessment of physical function[136], uses a VAS to evaluate 12 parameters related to daily activities: dressing, walking outdoors, climbing stairs, sitting for longer periods, standing bent over a sink, carrying a bag, making a bed, running, performing light and heavy work, lifting heavy objects and participation in exercise or sports. For each item, the patient indicates on a 100-mm scale his/her ability to perform this particular activity (0=without difficulty and 100=impossible) and the mean score for these items then represents the DRI. This form is self-administered after recieving oral or written instructions.

Composite measures such as the SF-36 and EuroQol, so-called generic instruments, encompass a broader view of health and can be employed to compare different populations or groups of patients. The 36 questions of which SF-36 is composed assess outcome in terms of four physical (physical function, role physical, bodily pain, general health) and four mental domains (social function, role emotional, mental health and vitality). The findings can be presented as a profile that includes an assessment of health-related quality of life.

In an attempt to standardize assessment of outcome in clinical research, in 1998 an international group of researchers concerned with back pain recommended the use of measures of the severity and frequency of symptoms (e.g., ODI), measures of general health status (e.g., SF-12 which is a more concise variant of SF-36), and a measure of overall satisfaction (such as global outcome)[35]. The importance of this latter

recomendation is also emphasized by the findings of Hägg and collegues that the global outcome provides a valid and responsive descriptor for the overall effect obtained in randomized controlled trials conserning treatment of CLBP [64].

Pain drawing has been employed as a tool for assessing this symptom and, moreover, proposed as an aid in selecting the optimal treatment strategy for individual patients. The Pain drawing can be interpreted either quantitatively[12,103,118,162] or qualitatively by characterization as "organic" or "non-organic" according to a system of penalty points[129]or the general impression[142,168], respectively. Certain studies reveal that PD is correlated to psychological distress[26,29,63,101,123,129,135,142,149,162,170], which other studies have shown to worsen outcome[1,32,60,65,98,148,150,152,154,165-167,181,186]. At the same time, several authors question the value of PD as an indicator of emotional or psychological disturbance[22]. In spine surgery the PD has been shown to be predictive in connection with surgery for disc hernia[152] or nerve root compression[186]. In contrast, the Swedish lumbar spine study observed no correlation between the pain drawing and the outcome for patients with unspecified chronic low-back pain treated with fusion[63].

Here we employed Spangfort's[153], modified version of a PD developed by Ransford and co-workers[129]. On front and back outlines of a person, the patient indicates the location of his/her pain, as well as its character using six different symbols, for dull, burning, numbing, stabbing or cutting, pins-and-needles and cramping (Figs. 1-4 in Paper 3). We classified the pain drawings as organic, possibly organic, possibly non-organic and non-organic according to the criteria of Uden and colleagues[168].

Work status-- including sick leave, disability pension etc.-- is also commonly used to assess outcome and also documented in the study, particularly because of its impact on the societal costs of LBP. However, this status is dependent on many factors other than subjective symptoms, e.g., the nature of one´s employment, the national system of social security, and possible litigation, as well as other personal factors such as motivation and ambition, limiting the possibilities of comparisons between different patient populations.

With respect to possible improvement of outcome, the treating surgeon more often than an unbiased observer tends to overestimate the improvement [131]. Consequently it is advocated that the patient not fill out questionnaires under the supervision of the treating surgeon, a proposal that has been adhered to all of the studies described here, in which all assessment of the clinical outcome is based on self-reporting by the patients. Unbiased observers that had not been involved in operating on the patient were used or, alternatively, the questionnaires were mailed to the patients.

1.3.2 Fusion

Fusion is generally recognised as being problematic to demonstrate[11,17,74,82,91]. In this connection plain radiographs are often found to be less accurate than thin-slice helical CT scan[24,137,145], with some exceptions[43]. The presence of stainless steel in the instrumentation for CT and, in particular, MRI creates distortion that renders

interpretation difficult. Even when titanium implants are used, some distortion is unavoidable.

A common way to assess posterolateral fusion employing plain radiographs is to classify the fused masses into four categories according to Lenke et al: (A) definitely solid, i.e., bilateral fusion masses that are solid large and trabeculated bilateral; (B) possibly solid = a large unilateral fusion mass accompanied by a small contralateral fusion mass; (C) probably not solid =small, thin fusion masses located bilaterally; and (D) definitely not solid =graft resorption bilaterally or a fusion mass with obvious bilateral pseudarthrosis[96].

1.3.3 Degeneration

Histological findings associated with degenerative disc disease (DDD) have been classified by Boos and coworkers[14], among others, and in vitro histological degenerative findings have been shown to be correlated to radiological findings in human cadavers[176,178] and animal models[127]. In addition, correlation between radiological degenerative findings and the results of the more sensitive MRI procedure has also been documented. For example, Frobin and colleagues demonstrated that the reductions in disc height are in general comparable to those assessed by MRI, although MRI also detects earlier signs of disc degeneration. However, same investigators showed that classification of height loss in individual discs on the basis of MR imaging was imprecise[50].

Straightforward A-P and lateral radiographs also possess the considerable advantages of being simple, inexpencive, non-invasive and usually readily available[178]. At the time when the present long-term study was initiated, assess to MRI was not commonly available and, therefore, was not performed on many of our patients, prior to inclusion.

Assessment of degenerative changes adjacent to a fusion or slippage remains controversial because of the difficulties involved in measuring such degeneration accurately. Another approach is to utilize semi-quantitative scales which take factors other than disc height into consideration. For example, the University of California at Los Angeles grading scale for disc degeneration also includes evaluation of osteophyte formation and sclerosis of the end-plate[55]. Degree of slippage and disc angle are additional parameters that may be influenced by disc degeneration.

1.4 THE STUDY POPULATIONS AND METHODS OF APPROACH

1.4.1 Criteria for inclusion and exclusion

The criteria for inclusion and exclusion were identical during all parts of the present study. The inclusion criteria were a diagnosis of isthmic spondylolisthesis (of any grade and levels, experience of low-back pain with or without sciatica, severe restriction of functional capabilities for at least one year and an age of 18 – 55 years. The exclusion criteria were having previously undergone spinal surgery or the presence of a psychiatric disorder and/or drug and/or alcohol abuse.

1.4.2 Study population 1

Between the years of 1990 and 1995 111 patients radiographically diagnosed with isthmic spondylolisthesis or spondylolysis (81 at Huddinge University Hospital and 30 at Linköpings University Hospital) were randomly assigned to receive one of three different treatments, i.e., posterolateral fusion (PLF) without instrumentation (n=40), PLF with pedicle screws (CDI) (n=37), and conservative treatment, i.e., a one-year exercise program (EX) (n=34). The procedure for randomization without stratification was as follows: After the patient had given his/her informed consent, a nurse at the outpatient ward randomly chose one of three different lots indicating the treatment to be provided. In this way, the nature of treatment was unknown to both the patient and the doctor until after consent.

Initially 5 other patients were also included, but two of these refused to accept random assignment of treatment; two improved sufficiently during the waiting period to render surgery unnecessary; and the remaining patient arranged for surgery at another hospital because of our long waiting period. Of the patients finally studied, 54 were women and 57 men, with an overall mean age of 39. 94 patients had lysis at L5; 14 at L4; and 3 at both L5 and L4. 67 patients had slippage of Meyerding grade 1; 42 of grade 2 and 2 of grade 3. The three groups exhibited similar distributions with respect to the level and grade of slippage, age, pre-treatment pain and disability, and life-style factors (Table 1).

Their pre-treatment pain drawings indicated that 32 of the patients experienced LBP only, 67 had low back pain and sciatica (i.e., draw symbols of pain below the knee); while 8 suffered only from sciatica. When the patients with sciatica and assigned to undergo PLF, with or without CDI, were examined by CT-myelography and/or MRI, none demonstrated any radiological signs of disc prolapse or central spinal stenosis.

In study 4 this entire population of 111 patients was again included; while the patients in study population 1 treated surgically (by PLF with or without instrumentation) were also included in studies 2 and 3.

Table 1
The demographic characteristics, symptoms, level and grade of slippage and life-style factors prior to treatment.

		Patients randomly assigned to treatment by		
	All patients (n=111)	**PLF (n=40)**	**PLF+CDI (n=37)**	**EX (n=34)**
Mean age (years)	39	39	39	37
Mean age at onset of symptoms (years)	26	25	29	25
Gender distribution, % (n)				
Women	49 (54)	45 (18)	57 (21)	44 (15)
Men	51 (57)	55 (22)	43 (16)	56 (19)
Symptoms, % (n)				
Low-back pain only	31 (33)	25 (10)	30 (11)	39 (12)
Low-back pain+sciatica	62 (67)	68 (27)	62 (23)	55 (17)
Sciatica only	7 (8)	8 (3)	8 (3)	6 (2)
Level of lysis, % (n)				
L5	85 (94)	83 (33)	84 (31)	88 (30)
L4	13 (14)	15 (6)	14 (5)	9 (3)
L4 and L5	3 (3)	3 (1)	3 (1)	3 (1)
Degree of slippage, % (n)				
Grade 1	60 (67)	68 (27)	54 (20)	59 (20)
Grade 2	38 (42)	30 (12)	43 (16)	41 (14)
Grade 3	2 (2)	3 (1)	3 (1)	0
Life style factors				
On sick-leave or with a disability pension , % (n)	71 (79)	68 (27)	84 (31)	62 (21)
Mean period of sick-leave prior to treatment, months	16	15	14	18
Blue collar worker, % (n)	80 (87)	90 (36)	75 (27)	73 (24)
Immigrants, % (n)	32 (35)	30 (12)	27 (10)	38 (13)
Married or co-habiting, %(n)	74 (82)	75 (30)	76 (28)	71 (24)
Smokers, % (n)	54 (60)	63 (25)	57 (21)	41 (14)
Medication for other than back pain , % (n)	21 (23)	30 (12)	16 (6)	15 (5)

1.4.3 Study population 2

The second investigation described here involved 86 patients included in accordance with the same criteria as in study 1 and treated surgically by posterior lumbar interbody fusion (PLIF) between the years 1997 – 2003 at Huddinge University Hospital (n=50), the Stockholm Spine Center (n=19), Stockholm Soder Hospital (n=11) or at Ryhovs hospital in Jönköping (n=6). Three other patients initially operated on were eventually excluded because one had a conjoined nerve root that rendered the PLIF operation impossible, and preoperative data were missing for two others, leaving 86 patients in study 2.

The patients experiencing sciatica were examined by CTmyelography and/or MRI and as in the case of the PLF group (see above), none exhibited any radiological signs of disc prolapse or central spinal stenosis. This study population was composed of 53 women and 33 men, with an overall mean age of 40. 66 patients had lysis at the level of L5; 17 at L4; and 1 at L3. Two patients had a slippage at two different levels. 75 had slippage of Meyerding grade 1; 9 of grade 2; and 2 of grade 3 slip. This PLIF group was similar to Study population 1 with regards to their distributions of age and gender, level of pre-treatment pain and disability and life-style factors such as work status, nature of their domicile and marital status.
In contrast grade 1 slippage and smoking were more frequent and alcohol use less frequent in the PLF group (Compare Tables 2 and 1).

Study population 2 was also used in study 3.

Table 2
The pre-operative demographic characteristics, symptoms, level and grade of slippage and life-style factors for our patients treated by PLF with or without CDI (from study 1) or PLIF (Study population 2).

Pre-operative characteristic	**All patients (n=163)**	**Patients treated by PLF with or without CDI (n=77)**	**Patients treated by PLIF (n=86)**	**p-value (PLF versus PLIF)**
Mean age (years)	40	39	40	0.78
Gender distribution Women, % (n)	56% (92)	51% (39)	66% (53)	0.16
Height (cm)	170	170	169	0.34
Weight (kg)	76	76	77	0.49
Experiencing Sciatica, % (n)	72% (118)	73% (56)	72% (62)	0.93
Location of slippage, % (n)				
L5	80% (130)	83% (64)	77% (66)	
L4	17% (28)	14% (11)	20% (17)	0.53
L3	1% (1)	0% (0)	1% (1)	
L4 + L5 or L3 + L4	2% (4)	3% (2)	2% (2)	
Degree of slippage, % (n)				
Grade 1	75% (122)	61% (47)	87% (75)	0.0004
Grade 2	23% (37)	36% (28)	11% (9)	
Grade 3	2% (4)	3% (2)	2% (2)	
On sick-leave or receiving disability pension , % (n)	69% (113)	75% (58)	64% (55)	0.12
Blue-collar worker, % (n)	80% (130)	84% (64)	77% (66)	0.23
Immigrants, % (n)	26% (42)	29% (22)	23% (20)	0.44
Married or co-habiting, % (n)	70% (114)	75% (58)	65% (56)	0.16
Smokers, % (n)	39% (63)	60% (46)	20% (17)	<0.0001
Teetotaller (Consuming no alcohol), % (n)	24% (38)	33% (25)	16% (13)	0.014
Exercise regularly (≥ once a week), % (n)	49% (78)	45% (33)	53% (45)	0.29
Living in a house (not an apartment), %(n)	40% (64)	38% (29)	41% (35)	0.74
DRI	48	49	47	0.45
Pain Index	65	64	66	0.36

1.4.4 Surgical treatment

All of the 77 patients subjected to PLF with or without CDI underwent a posterolateral fusion in situ, with transfer of autologous bone graft from the right iliac crest to a posterolateral location including in between the intertransverse processes. In all patients experiencing sciatica (n=56) a Gills procedure involving removal of the loose lamina and relief of nerve root compression was performed. Fusion at two or three different levels was performed in 16 and one of these 77 patients, respectively, due to pronounced degeneration of the adjacent disc, retrolisthesis, or pronounced slippage. Pedicle screws (CDI) were added to the 37 patients assigned randomly to receive such treatment.

In the case of the PLIF group, following clearance of the disc space with shavers two wedged (7°) carbon fibre "ramps" (DePuy Spine Inc., Raynham, MA) were introduced. Thereafter, autologous bone graft taken from the loose lamina excised was placed between and lateral to the "ramps". In addition, posterolateral fusion with laminar and/or iliac bone graft was performed in all except 19 of these patients. When fusion at more than one level was required the PLIF procedure was performed only at the level of slippage. Pedicle screw fixation was utilised in all of the PLIF procedures.

In two of the patients treated with PLIF, it was impossible to introduce two ramps for technical reasons, therefore in one of these cases the PLIF was performed using only a pulverized autologous bone graft and in the other only a single ramp was implanted. Due to degenerative changes in the adjacent level or pronounced slippage, fusion was performed at two levels in 20 patients and three levels in one patient.

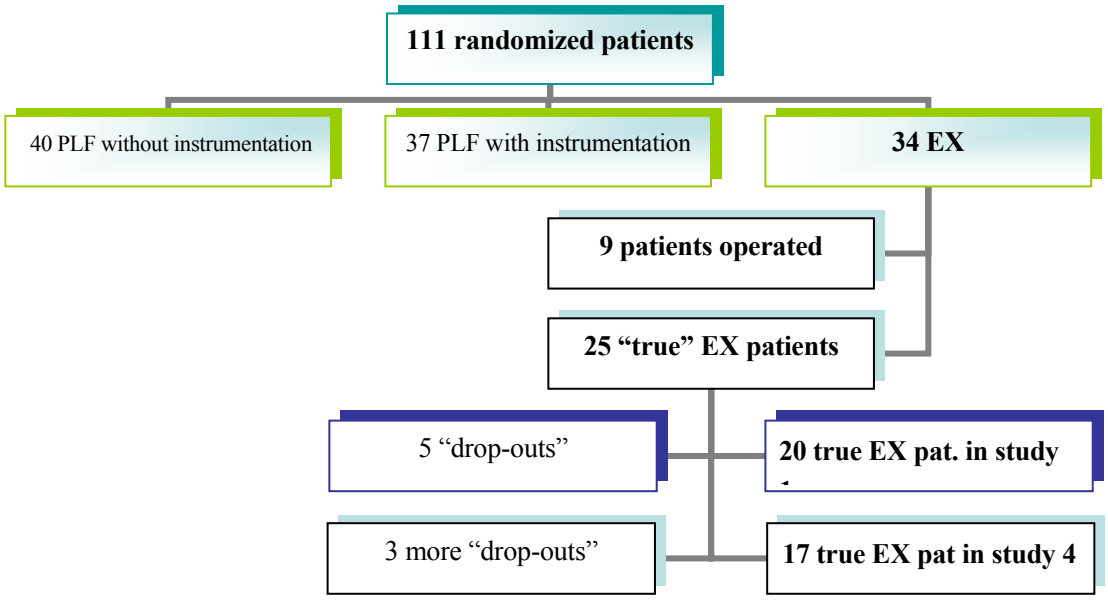
1.4.5 Conservative treatment

The 34 patients assigned randomly to receive conservative treatment underwent a one-year exercise program supervised by a physiotherapist, involving three sessions a week for the first half-year and two a week for the second 6-months period. The 12 exercises involved focused on improving the strength of back and abdominal muscles, in combination with postural training. Eight of these exercises did not require any special equipment and after the first year, the patients were instructed to continue doing these at home on their own. 67% of the patients complied with the entire first year program, but it is not known to what extent this group continued with the recommended exercises thereafter.

Nine of these patients were eventually operated on for their original symptoms at, for most of them, a relatively early time-point following inclusion in our study (i.e., after 20, 21, 25, 26, 27, 28, 32, 34 and 54 months (range=20-54, median=27 months). Seven were operated on only at the level of slippage, while in two cases an extra segment (L4 to S1) was fused because of a high degree of slippage, for ease of instrumentation and to increase fusion area.

"Drop-outs" and the remaining "true" EX patients are illustrated in Figure 1.

Figure 1
Flow-chart illustrating the number of patients assigned to receive conservative treatment who remained non-operated 9 (study 1) and 13 years (study 4) after inclusion.



1.4.6 Asessments

In all of our studies the patients filled out self-administered questionnaires prior to treatment as well as one and two years after treatment, and in the case of studies 1 and 4, also at long-term follow up. The following parameters were assessed on the basis of these questionnaires:

Sociodemographic characteristics: Age, sex, height, weight, civil status, parentage and the presence of children at home, nature of domicile (house- or apartment), nationality, smoking, drinking and exercise habits, dominant hand, work status (blue/white-collar work, sick leave or disability pension).

Pain drawing according to Spangfort (Figures 1-4 in Paper 3)[153].

Pain index, as the mean of two VAS for immediate pain right now and the worst pain experienced during the past week.

Degree of Disability according to the ODI, DRI and SF-36.

Quality of life according to SF-36.

Global assessment: the patient's rating of his/her condition as "much better", "better", "unchanged" or "worse".

Radiographic assessments

1. **The degree of fusion** was assessed according to Lenke et al.[96]. In Study 4, bony bridging between vertebral bodies (indicating fusion), and resorbtion around screws (indicating pseudarthrosis) were employed as additional criteria.
2. **Degenerative Disc Degeneration (DDD)** was assessed in Study 4 by measuring disc height and retro- or anterolisthesis utilizing two different digital radiographic procedures:
    1. The standard software system employed in our hospitals for viewing and measuring distances in radiographs (IDS5, Sectra PACS™); and
    2. the FDA-approved computer-assisted QMA™ measurement, which was performed by Medical Metrics, Inc (Tx, USA)[126,189] and also employed to determine disc angle.

    A third approach, the semi-quantitative UCLA grading scale for disc degeneration [55] was used in slightly modified form to compare pre-operative radiographs with those obtained during long-term follow-up.

Complications and re-operations during the follow-up period were recorded. In addition, questions concerning re-operations were included in the questionnaires. Finally, the patients' medical journals were reviewed in order to minimise the risk of missing complications.

1.5 STATISTICAL ANALYSES

The "intention-to-treat principle" was applied in the analyses within and between the different groups receiving surgical treatment or exercise (Studies 1-3).

For ordinal data, such as the VAS scores for pain index and DRI as well as for other scores; e.g., from the ODI and SF-36, non-parametric tests were used for statistical analysis. The Mann-Whitney U-test was applied for comparison of two sets of unpaired data and the Wilcoxon signed rank test was for paired data, (i.e., longitudinal comparisons within each individual group). The 95% confidence intervals, based on the standard error of the mean, were also calculated.

In order to test for differences between treatment groups with respect to normally distributed, continuous variables such as reduction in disc height and vertebral slippage the unpaired t-test was employed; while the paired t-test was used to analyze longitudinal differences within each treatment group. To determine differences in categorical data (such as the ability to work, global outcome, and pre-treatment demographics), the χ^2-test was applied. For ordered categorical data (e.g., the patient's global outcome), the χ^2-test for trend was also used. In all cases a p-value of <0.05 was considered to be statistically significant.

For assessment of appropriate sample sizes, a power analysis was performed. Setting the risk for a Type-I error at 5% (significance level 0.05) and the risk for a Type-II error at 10% (90% power), with a standard deviation for the DRI of 18 and the minimal clinical difference of interest equal to 15 DRI points, the necessary size of each group was calculated to be 30 patients.

To evaluate the validity of our digital procedure for x-ray measurement (study 4), inter- and intraobserver reliability were estimated to be 0.87-0.99 and 0.93-0.97, respectively, on the basis of the Shrout-Fleiss Intraclass Correlation Coefficients (ICC)[36]. Interobserver reliability regarding of the pain drawings classification into O and NO groups was assessed to be 0.78 by kappa statistics.

In Study 3 possible correlations between prognostic factors and the major outcome variable PI after 2 years of follow-up were evaluated with the MannWhitney U-test for non-parametric data and all factors exhibiting $p<0.10$ analyzed further. To examine the crude association between the outcome variable and each of these correlated factors, we applied univariable linear regression. Thereafter, adjusted associations between these factors and the outcome, were analyzed by forward linear multivariate stepwise regression, with entry testing based on the F-test (the two sided p-values for entry being set to 0.05). Finally, the crude and adjusted association for the "pain drawing" factor were compared.

1.6 ETHICAL CONSIDERATIONS

The studies involving the 111 patients assigned randomly to receive treatment by PLF or physiotherapy were pre-approved by the Medical Ethical Committee of Huddinge University Hospital while the multicenter study also involving 86 patients subjected to PLIF was pre-approved by Karolinska Institute's Regional Medical Ethical Committee.

1.7 SUMMARY OF THE FINDINGS

1.7.1 Paper 1

The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study

Methodological considerations

This prospective randomized controlled study was designed to compare on the long-term outcome of treatment by fusion or a one-year program of exercise. The patients were assigned randomly into three treatment groups, i.e., posterolateral fusion without instrumentation (PLF; n =37), PLF with pedicle screw instrumentation (CDI; n=40) and physiotherapy/exercise (EX; n=34) which exhibited similar demographic characteristics, symptoms, level and grade of slippage, and life-style factors (Table 1). The average period and rate of follow-up time were 9 years (range 5-13) and 91% (101/111 patients) respectively. During long term follow-up, 9 patients in the EX group were eventually operated on, leaving 19 non-operated individuals in this group. The "intention-to-treat" principle was used, but statistical analysis was also performed after exclusion of the 9 EX patients who underwent operation. The limited number of patients in the conservative group and the drop out of 9 patients renders interpretation more difficult with a risk of a type 2 error and selection bias.

Findings

Longitudinal analysis

All the three treatments reduced the patients' level of pain, but only the groups treated surgically demonstrated functional improvement. However, following surgery, functional disability (as assessed by the DRI, but not the ODI) was significantly worse after long-term than two-years follow up; whereas no such difference between these two periods of follow-up was observed in case of the EX group. Among the patients treated surgically, 25% were working prior to their operation and 51% during long-term follow-up ($p<0.0001$). The corresponding values for the EX group were 38% and 46% (ns) (Table 3; Figures 2 and 3).

Exclusion of the 9 EX patients that had been operated on exerted no significant impact on any assessment of outcome in connection with long term follow-up.
In addition, no significant differences between the instrumented (CDI) and non-instrumented PLF patients were observed for any of the variables examined (Tables 4 and 5).

Cross-sectional analysis

Upon long-term follow-up, the only significant difference in any assessment of outcome between the patients treated surgically and with exercise was a better global assessment for the former. 76% of the surgical patients rated their overall outcome as much better or better compared, with 50% of those treated conservatively ($p=0.015$; Table 5). As evaluated by the SF-36 long-term quality of life was the same for all three groups and considerably lower (with respect to all of the 8 domains) than for the general population (according to the SF-36 manual) (Figure 4).

Table 3

The mean pain index (PI), DRI, and ODI for the individuals treated surgically (PLF) or with exercise (EX) and the percentages of those working prior to treatment after vario us periods of follow-up.

Tre atm ent	Para-meter	Prior to treatment (n=106)	After follow-up for			p [b]
			1 year (n=98)	2 year (n=106)	Long-term (n=101)	
PLF	PI [a]	63 (58.5-67.7)	35 (28.7-42.2)	37 (29.6-43.8)	40 (34,0-47,1)	<0,0001
	DRI [a]	48 (43.9-52.3)	29 (23.0-34.6)	29 (23.5-34.9)	33 (27.8-38.8)	<0,0001
	ODI [a]	nd	nd	26 (18.1-31.6)	28 (23.0-33.0)	0.223*
	Working	25%	46%	54%	51%	<0.001
EX	PI [a]	65 (57.3-71.9)	54 (44.7-63.7)	56 (48.7-63.8)	49 (38,4-58,8)	0.013
	DRI [a]	44 (38.2-50.3)	45 (36.4-53.7)	44 (36.5-50.9)	38 (29,1-47,7)	0.131
	ODI [a]	nd	nd	28 (20.5-35.0)	31 (24.0-37.4)	0.887*
	Working	38%	48%	55%	46%	n.s.

[a] The best and worse possible status are scored as 0 and 100, respectively. The values within parentheses are 95% confidence intervals.

[b] These p-values are for the comparison between the pretreatment scores and the scores after long-term follow-up, with the exception of the ODI parameter, where the p-values (marked with an asterisk) refer to the comparison between the values obtained after 2-year versus long-term follow-up.

n.s. = not significant

nd = not determined

Fig 2

The mean Pain index for the patients treated surgically or with exercise prior to treatment and after various periods of follow-up. The bars indicate 95% confidence intervals.

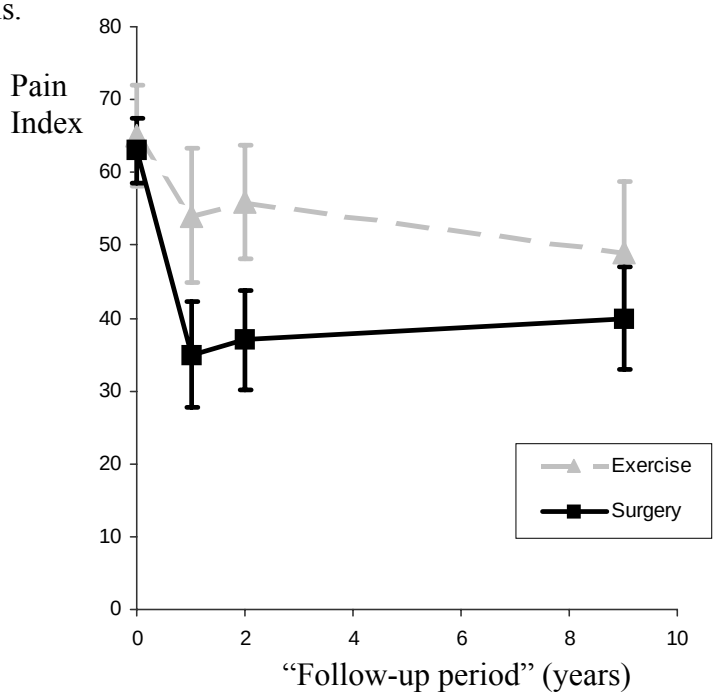


Fig 3
The mean DRI value for the patients treated surgically or with exercise prior to treatment and after various periods of follow-up. The bars indicate 95% confidence intervals.

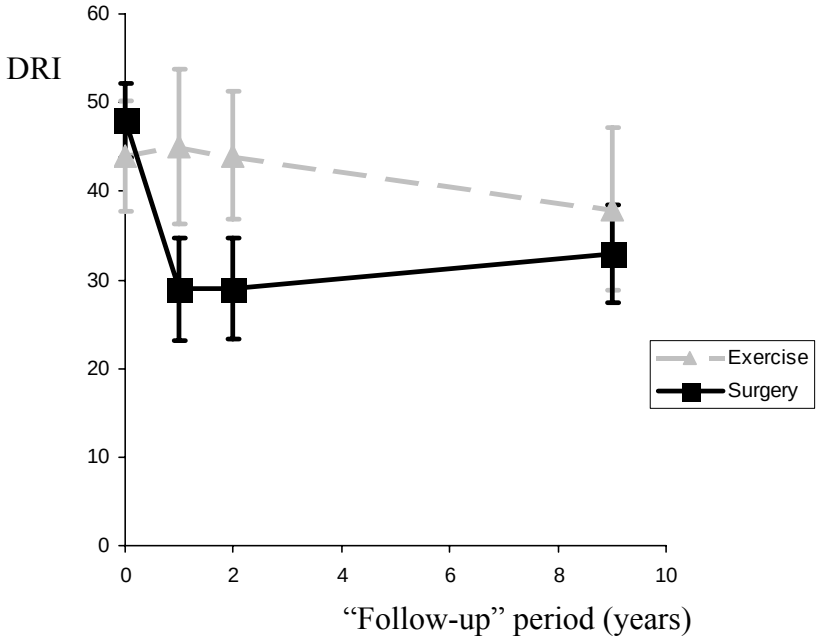


Table 4
The mean pain index and DRI and ODI values for patients subjected to PLF with and without instrumentation and the percentages working after long-term follow-up.

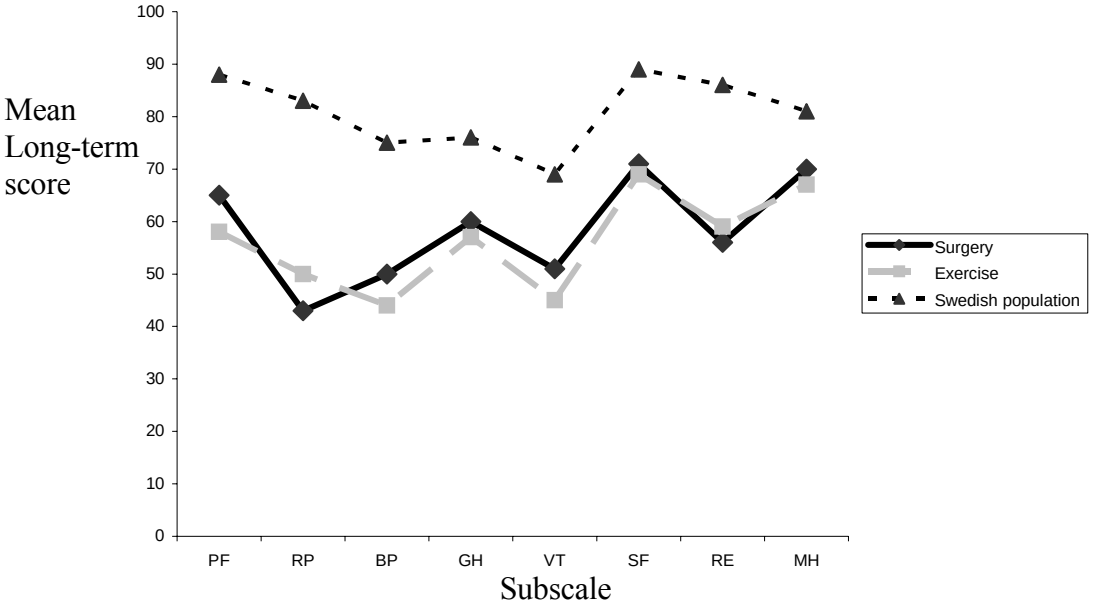
	Non-instrumented PLF	Instrumented PLF	p
Pain index [a]	45	36	0.27
DRI [a]	36	30	0.31
ODI [a]	30	27	0.79
Working	50%	52%	0.90

[a] The best and worse status are scored as 0 and 100, respectively.

Table 5
Self-reported global outcomes for patients treated by PLF with or without instrumentation or with exercise after long-term follow-up (ns).

Global outcome (%)	Instrumented	Non-instrumented	All PLF combined	Exercise
Much better	33	44	39	27
Better	49	22	37	23
Unchanged	9	22	13	27
Worse	9	12	10	23

Fig 4
The mean long-term scores for the surgical and exercise groups on the eight subscales of the SF-36 and the corresponding scores for the general Swedish population. There were no significant differences between the two treatment groups with respect to any of the subscales. In contrast both groups differed significantly from the general population with respect to all 8 subscales (PF= Physical Function, RP=Role Physical, BP= Bodily Pain, GH=General Health, VT=Vitality, SF=Social Functioning, RE=Role Emotional, MH= Mental Health).



Conclusions

PLF results in a limited, but significant positive effect on long-term, self-reported global outcome in comparison to a one-year exercise program and, most likely, in comparison to the natural history of the disorder. The natural long term course of unoperated symptomatic spondylolisthesis in adults involves continued disability affecting many aspects of life. The patients operated on appear to have experienced less severe symptoms for many years. In addition, instrumentation clearly exerted neither a positive nor negative influence on the long-term outcome.

1.7.2 Paper 2

Posterior Lumbar Interbody Fusion versus Posterolateral Fusion in Adult Isthmic Spondylolisthesis

Methodological considerations

This prospective comparative study was designed to compare the 2-year outcome for 86 patients operated on by PLIF (study population 2) to that of historical controls. The criteria for inclusion and exclusion, assessment of outcome and data collection were identical to those employed in Study 1, whose participants served as the control group. These PLIF patients were similar to the PLF patients with respect to socioeconomic variables, age and gender distribution, level of slippage, extent of sciatica, level of pain, and disability (Table 2).

The rates of follow-up for these PLIF and PLF groups were 98% and 97%. In an EBM perspective the non-RCT character of the study, with different populations compared, reduces the level of the evidence obtained to Level 3, as compared to Level 1 in paper 1 and 4.

Findings

The mean pain indexes and DRI values for the patients undergoing PLIF or PLF were virtually identical (Figures 4 and 5 in Paper 2) and moreover, 54% and 52% of these individuals, respectively, were working after 2 years of follow-up (Table 6).

Table 6
The mean pain index and DRI and for patients subjected to PLIF or PLF and the percentage of of these individuals working prior to treatment and after one or two years of follow-up.

Treatment	Parameter	Prior to operation	After one year of follow-up	After two years of follow-up	p [b]
PLIF	**Pain index** [a]	66	35	35	<0.0001
	DRI [a]	47	30	30	<0.0001
	Working % (n)	36% (31)		52% (43)	0.0008
PLF	**Pain index** [a]	64	35	37	<0.0001
	DRI [a]	49	31	29	<0.0001
	Working % (n)	25% (19)		54% (40)	<0.0001

[a] The best and worse values are scored as 0 and 100, respectively.
[b] for the pretreatment score versus the score after two years of follow-up.

The mean ODI for both groups was 25 after two years of follow-up. In addition, the self-reported global outcome was similar for both groups (p=0.3) (Table 7).

Table 7
Self-reported global outcome for the PLIF and PLF groups after 2 years of follow-up.

Self-reported global outcome % (n)	Treatment	
	PLIF (n=83)	PLF (n=73)
Much better	43% (36)	55% (40)
Better	31% (26)	19% (14)
Unchanged	16% (13)	11% (8)
Worse	10% (8)	15% (11)

There were four major complications in the patients subjected to PLF compared to 12 major complications after PLIF (p=0.06), as follows:
The PLF group

- two permanent nerve root injuries at L5, resulting in dermatomal pain and in one case weakness of foot extension
- one patient with permanent unilateral blindness
- one patient who experienced transient dermatomal pain, which resolved after a month.

The PLIF group

- three deep wound infections
- two patients with permanent leg pain
- two who experienced with transient leg pain
- one patient with permanent foot-drop
- one with transient foot-drop
- one patient with a deep vein thrombosis
- one patient with a pulmonary embolus
- one patient with postoperative paraparesis subsequently diagnosed as psychogenic

Conclusions

Despite the several theoretical advantages associated with the PLIF-procedure its outcome here was no better than with the "gold standard" PLF. Furthermore PLIF resulted in a larger number of complications.

1.7.3 Paper 3

Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis

Methodological considerations

The prospective study examined the predictive value of the pre-operative pain drawings and certain other factors with respect to patient outcome two years after lumbar fusion. The 164 patients in studies population 1 and 2 who were treated surgically filled out pre-operative questionnaires that included a pain drawing and items concerning the duration of their symptoms, age, gender, height, weight, work status, marital status, type of domicile, births and number of children, possible immigrant status, level of education, smoking, drinking and exercise habits, dominant hand, etc. The level and degree of slippage were also documented radiologically. The pain drawings were classified on the basis of general impression into organic, possibly organic, possibly non-organic and non-organic [168]. In order to increase the statistical power, the first two categories were combined into a single organic group (O; n=126) and the latter two into a non-organic group (NO; n=38) (Table 8).

The primary parameter of outcome selected for assessment was PI and we also evaluated the DRI and ODI values, as well as self-reported global outcome. The rate of two-year follow-up was 98% (160/164). To control for confounders univariate as well as multivariate analyses were performed.

Employing PI as the primary indicator of outcome univariate analysis was employed to test the possible predictive value of the following factors: gender, age, work status, blue/white-collar work, immigrant status, marital status, smoking habit (none, at all, > 15 cigarettes/day), alcohol consumption, regular exercise, living in a house or apartment, height, weight, BMI, experiencing sciatica, degree of slippage, level of slippage/lysis, treatment by laminectomy, dominant hand, parenthood and children still living at home.

Findings

NO PD, female gender, not working, lack of regular exercise and short stature were all found to be correlated to a less favourable outcome in terms of the absolute PI (Table 9). In addition to experiencing a lower level of pain patients who had been working pre-operatively; also had a *greater reduction* of pain, lower DRI and ODI scores, a more pronounced *improvement* in function (as reflected in the DRI scores) and a better self-reported global outcome following two years of follow-up. This was also the case, although not for all outcome variables, for gender, exercise habits and O/NO PD (Table 10 + 11). At this same time-point, patients who produced O pain drawings demonstrated significantly better function than those with NO pain drawings and significantly more in the O group considered themselves to be “much better” (54% versus 33%) (Table 11).

Table 8
Demographic characteristics, life-style factors, symptoms, weight, BMI, functional disability and pain index prior to treatment for patients who produced "organic" (O) or "non-organic"(NO) pre-operative pain drawing.

Parameter	All patients (n=164)	Patients with "O" pain drawings (n=126)	Patients with "NO" pain drawings (n=38)	p-value (O versus NO)
Mean age (years)	40	40	37	0.062
Gender: Proportion Women	57%	54%	66%	0.1845
On sick-leave or recieving disability pension (%)	69%	63%	89%	0.0021
Blue-collar worker	80%	78%	89%	0.1025
Immigrants	26%	22%	39%	0.0274
Married	70%	69%	74%	0.5653
Smokers	39%	38%	42%	0.7234
Teetotaller (no alcohol)	24%	23%	27%	0.6465
Exercise regularly (≥ once a week)	49%	51%	43%	0.4061
Live in a house (not an apartment)	40%	43%	30%	0.1661
Mean height (cm)	170	170	168	0.1183
Mean weight (kg)	76	77	73	0.0722
Mean BMI	26	26	26	ns
Mean DRI	48	46	53	0.0263
Mean Pain Index	65	63	73	0.0066

ns= not statistically significant

Table 9
The estimated β-slope coefficients, confidence intervals, significances (as determined by univariant linear regression analysis) and associated mean pain indices after two years of follow-up for the pre-operative factors actively working, gender, regular exercise and pain drawing (PD).

Factor/Variable	Estimated β-slope coefficient (95% CI)	p-value	Mean PI after two years of follow-up for the respective variable 0;1
Working (=0) or not (=1)	22.8 (13.3, 22.8)	<0.0001	20; 43
Man (=0) or Woman (=1)	15 (5.7, 24.1)	0.002	27; 42
Exercise (=0) or not (=1)	9.7 (0.31, 19.1)	0.043	30; 40
O (=0) or NO PD (=1)	19 (7.9, 30)	0.001	31; 50
Height pre-op (cm)	-0.64 (-1.13, -0.14)	0.013	35 (mean PI for all)

Table 10
The mean pain indices (PI) and improvement in this index, DRI values and improvements in these values, Oswestry scores and self-reported global outcomes two years post-operatively associated with the factors actively working, gender and regular exercise.

	Working			Gender			Regular exercise		
Parameter	**Yes (n= 51)**	**No (n= 113)**	**p-value**	**♂ (n= 71)**	**♀ (n= 93)**	**p-value**	**Yes (n= 81)**	**No (n= 83)**	**p-value**
Mean PI after 2 years	20	43	<0.0001	27	42	0.0019	30	40	0.046
Mean improvement in PI	39	25	0.0034	33	26	0.1143	31	27	0.266
DRI at 2 years	14	37	<0.0001	25	33	0.032	25	34	0.0371
Mean improvement in DRI	22	15	0.0205	20	15	0.0993	19	16	0.264
Mean ODI after 2 years	13	31	<0.0001	22	28	0.0474	21	29	0.0334
Much better	61%	43%	0.039	61%	40%	0.0082	56%	42%	0.0864

Table 11
The mean pain indices (PI), and improvements in this index, mean DRI value and improvements in these values, Oswestry scores, self-reported global outcome and working status two years post-operatively for the groups of patients who produced Organic (O) or non-organic (NO) pain drawings.

Parameter	O pain drawing (n= 125)	NO pain drawing (n=38)	p-value
PI after two years	32	50	0.0014
Mean improvement in PI	31	23	0.09
DRI after two years	27	41	0.0013
DRI improvement 0-2 years	19	11	0.0495
ODI after two years	22	36	0.0017
% Much better	54%	33%	0.038
On sick-leave or disability pension	41%	59%	0.07

There was no significant difference for any outcome variable or in the nature of pain drawings for patients with or without sciatica (Table 12).

Table 12
The mean pain indices (PI), and improvements in this index, mean DRI value and improvements in these values, Oswestry scores, self-reported global outcomes two years post-operatively, as well as the nature of the pre-operative pain drawing (PD) by LBP patients with and without sciatica.

Parameter	Patients experiencing sciatica (n=118)	Patients without sciatica (n=45)	p-value
Mean PI after 2 years	36	34	0.515
Mean improvement in PI	29	30	0.85
Mean DRI after 2 years	31	27	0.33
Mean improvement in DRI	17	18	0.82
Mean ODI after 2 years	26	23	0.475
Much better (self-reported)	49%	48%	0.87
Organic pre-operative PD	79%	71%	0.30

The forward stepwise multivariate analysis revealed that work status was the major determinant of outcome. Gender, exercise and PD exerted a smaller impact and, indeed, the influence of PD was not statistically significant (p=0.06) (Table 13). With regards to the variability in the PI, 12% could be explained by work alone, 5% by gender alone, 3% by exercise alone and 2% the pain drawing alone. As documented in Table 14, when taken as independent variables in multivariate analyses, the other four parameters of outcome demonstrated a significant correlation to certain of the pre-operative factors after two years of follow-up.

Table 13
The standardized β-coefficient, 95% confidence interval and significance after the last step of multiple linear regression analysis for the pre-operative factors work status, gender, exercise and nature of the pain drawing.

Variable/Factor	β-Standardized coefficient (95% CI)	p-value
(Constant)	7.8 (-1.8, 17.4)	0.110
Work status	19.1 (9.9, 28.3)	0.000
Gender	13.0 (4.3, 21,7)	0.004
Exercise	9.7 (1.2, 18.2)	0.025
O or NO pain drawing	10.6 (-0.5, 21.6)	0.060

Table 14
Factors significantly correlated to the Pain index and four other outcome parameters (as determined by stepwise multiple linear regression analysis).

Outcome / variable after two years of follow-up	**Significantly correlated factors**
Pain index	Work, gender, exercise
Oswestry index	Work, exercise, length
DRI	Work, immigration, exercise
Global outcome: better vesus not better	Work
much better versus all other alternatives	Immigration, exercise, gender

Conclusions

We found here that working status (in particular), male gender and regular exercise were the pre-operative factors that positively influenced the outcome of fusion. These three factors, however, together could only explain 20% of the variability in the outcome. Thus, other predictive factors with a major impact remain to be identified. The nature of the pain drawing was not a statistically significant factor when all of the parameters were considered in the multivariate analysis.

1.7.4 Paper 4

A Prospective randomized study on the long-term effect of lumbar fusion on adjacent disc degeneration.

Methodological considerations and limitations

This long-term prospective randomized controlled investigation was focused on degenerative changes at the level adjacent to a fusion or the slippage in 111 patients assigned randomly be treated by fusion (PLF) or conservatively with exercise (EX) (Study population 1, Table 1). The 9 EX patients who were operated on were excluded, as were two patients in the PLF group in whom fusion was considered to have been unsuccessful. Thus, radiographs were obtained from 63 patients after lumbar fusion and 17 patients after physiotherapy (giving an overall follow-up rate of 72%) after an average of 13 years (range10-17) of follow-up.

Three different radiographic procedures were employed for quantification of ASD: digital radiographic measurements performed with our hospitals' standard software system for viewing and measuring distances in radiographs (IDS5, Sectra PACS™); Quantitative Motion Analysis (QMA™), which is a computerized, FDA-approved radiographic technique [126,189]; and the semiquantitative UCLA grading scale for disc degeneration[55]. A comparison of the intraobserver and interobserver reliabilities of our own digital radiographic procedure and the QMA™ is presented in Table 15.

The limited number of patients treated with exercise introduce the risk for a type-2 error. It could further be argued that exclusion of the nine (26%) patients in the EX group who eventually underwent operation may have unduly influence our results on ASD since these patients may have been generally more prone to disc degeneration. Two observations, however, suggest that the patients remaining in the EX group are representative of the group originally assigned randomly to receive conservative treatment. In the first place the patients excluded were operated on relatively soon (median time = 27 months) after inclusion in our study, most likely because the symptoms from their spondylolisthesis continued, and not because of degenerative changes at the level above, which require a longer period to develop[58]. Secondly, only two of these 9 patients were also fused above the spondylolisthetic level, i.e., essentially the same proportion of the PLF group treated in this manner which contraindicates early development of "ASD" among those excluded.

Findings

Both our own digital radiographic and the QMA™ revealed that PLF patients exhibited shorter disc height in the adjacent segment than those in the EX group although this difference was not statistically significant in the case of the QMA™ (Tables 16 and 17). Both of these approaches demonstrated significantly more height reduction in the posterior than the anterior disc in both group of patients. The alteration of slippage (sagittal translation) into retrolisthesis was somewhat, but non-significantly greater in the fused than in the EX group according to both methods of assessment. Typically, after long term follow-up, both treatment groups contained more patients with retrolisthesis than with anterolisthesis (Table 19). Comparison of radiographs taken prior to treatment and after long-term follow-up with the third approach, the

UCLA grading scale, showed a significantly higher level of DDD in the fused group (p=0.026) (Table 19).

The prevalence of radiographically verifiable ASD varied from 6- 38 % in the fused group and 0-6% in the EX group depending on the definition of ASD applied (Table 20). The only demonstrable significant difference in the outcome for the fused group concerned the self-reported global outcomes associated with ASD defined as a reduction in posterior disc height > the mean +2SD observed for the EX group (Table 21).

Table 15

Intra- and interobserver reliabilities for our digital radiographic measurements (determined here) and the QMA™ (values taken from Pearson et al[126]). The Shrout-Fleiss intraclass correlation coefficients (ICCs) are shown.

ICC	**Intraobserver reliability**	**Interobserver reliability**
Digital X-ray	0.93 - 0.97	0.87 – 0.99
QMA™	0.95 – 0.96	0.70 – 0.79

Table 16

Alterations in anterior, posterior and mean disc heights and in degree of slippage (sagital translation) (regardless of the presence of antero- or retrolisthesis) as determined by to the digital radiographic procedure.

Parameter	**Conservative treatment (EX)**	**Fusion (PLF)**	**p (PLF vs EX)**
Ant Disc height	+3%	-7% ***	0.0098
Post Disc height	-11% *	-30% ***	0.0079
Mean Disc height	-2%	-15% ***	0.0016
Sagital translation	6%	15%	0.6358

* p< 0.05, *** p<0.001 within the group, as calculated by longitudinal analysis

Table 17

Alterations in absolute anterior and posterior disc heights, disc angle and degree of slippage in PLF and EX patients as determined by the QMA™.

Parameter	**Conservative treatment (EX)**	**Fusion (PLF)**	**p (PLF vs EX)**
Anterior disc height (mm)	0.21	-0.33	ns
Posterior disc height (mm)	-0.84*	-1.28***	ns
Disc angle (°)	1.8**	1.7***	ns
Sagital Translation (mm)	0.08	-0.55**	ns

* p< 0.05, *** p<0.001 within the group, as calculated by longitudinal analysis

Table 18
The number of EX and PLF patients identified as demonstrating anterolisthesis or retrolisthesis after long-term follow-up employing our digital radiographic procedure.

Condition	EX patients	PLF patients
Anterolisthesis	1	7
Retrolisthesis	11	55

Table 19
The number of patients classified as exhibiting disc degeneration through comparison of radiographs taken prior to to treatment and after long-term follow-up with the UCLA grading scale. The difference between the EX and PLF groups was statistically significant (p=0.026).

UCLA scale	Exercise	Fusion
Grade 1 (normal)	17	39
Grade 2 (disc heght reduction)	0	11
Grade 3 (and /or osteophytes)	0	12
Grade 4 (and/or endplate sclerosis)	0	1

Table 20
The prevalence of radiographically verifiable ASD in the two treatment groups employing different definitions of ASD

Definition of ASD	Fusion group (n=63)	Exercise group (n= 17)
Reduction in posterior disc height of > 2 SD, as observed in the EX group (i.e., natural process), % (n)	14 % (9)	6 % (1)
A remaining mean disc height < 20% of the anterior vertebral height, % (n)	11 % (7)	6 % (1)
Any deterioration on the UCLA grading scale	38 % (24)	0 %
Totally elimination of posterior disc height	6 % (4)	0 %

Table 21
Global outcome self-reported by fused patients with and without ASD (p=0.036). ASD was defined as reduction posterior disc height of >2SD in comparison to natural processes (i.e., the corresponding reduction in the EX-group).

Self-reported global outcome % (n)	Patients with ASD (n=9)	Patients without ASD (n=45)
Much better	11% (1)	49% (22)
Better, unchanged or worse	89% (8)	51% (23)

The mean Pain Index and DRI and ODI values were all non-significantly worse for the few patients defined as exhibiting ASD, regardless of the criteria used to define this condition, with one exception: using the UCLA grading scale, the outcome for the 24 patients defined as having ASD was virtually identical to that in the absence of this condition (Table 22). Subgroup analysis of our PLF patients revealed that as defined by the UCLA criteria (grade 2-4), ASD was present almost exclusively in those who had undergone laminectomy (p=0.015) (Table 23). By any definition of ASD the use of instrumentation did not influence the prevalence of this condition (Table 24).

Table 22
The differences in the mean Pain index and DRI and ODI values for fused patients with and without ASD (as identified according to four different definitions).

Difference in the mean values for PLF patients with and without ASD	**Definition of ASD**			
	Reduction in posterior disc height of >2SD (n=9)	**A remaining mean disc height of <20% (n=7)**	**UCLA grades 2-4 (n=24)**	**Elimination of posterior disc height (n=4)**
Pain index	6 (p=0.77)	19 (p=0.09)	0 (ns)	24 (p=0.09)
DRI	8 (p=0.45)	20 (p=0.08)	0 (ns)	20 (p=0.11)
ODI	4 (p=0.64)	13 (p=0.09)	2 (ns)	5.5 (p=0.24)

ns = not significant

Table 23
The number of PLF patients subjected or not subjected to laminectomy who exhibited ASD (UCLA grades 2-4) (p=0.015).

Procedure, % (n)	**Patients without ASD**	**Patients with ASD**
Laminectomy	53% (25)	47% (22)
No laminectomy	87.5% (14)	12.5% (2)

Table 24
The numbers of PLF patients fused with (n=32) or without instrumentation (n=31) identified as having ASD according to four different definitions.

Fusion	**Definition of ASD**			
	Reduction in posterior disc height of > 2SD (n)	**A remaining disc height of < 20% (n)**	**UCLA grades 2-4 (n)**	**Elimination of posterior disc height (n)**
Instrumented (32)	4	2	11	1
Non-instrumented (31)	5	5	13	3

No significant differences in any of the radiological parameters examined here (disc height and alterations in disc height and sagittal translation) for PLF patients operated on with or without instrumentation could actually be demonstrated. Both of these groups demonstrated similar tendencies for retrolishesis of the vertebra above the level of slippage, as well as for development of retrolisthesis or anterolisthesis. Furthermore, the mean pain indices, DRI and ODI values and self-reported global outcomes were for patients with ASD who had undergone instrumented or non-instrumented fusion not significantly different.

Conclusions:

Fusion and laminectomy are associated with degenerative changes at the adjacent level that develops more rapid than those that occur naturally, with the height of the posterior disc being affected more frequently than that of the anterior disc. Moreover, the use of instrumentation in connection with fusion does not influence the development of ASD. Except in its rarer, more severe forms, ASD has only a marginal effect on most patients. In view of the methodological considerations discussed, however, further investigations in this controversial area of ASD are certainly warranted.

1.8 GENERAL DISCUSSION

Treatment of chronic low back pain has long been a challenge for spine surgeons. Despite the fact that Level 1 RCT studies in spine surgery have become much more common the basic question whether fusion improves outcome in patients with LBP is still unclear. Only three RCTs have compared fusion to conservative treatment in DDD, and two out of these three are negative, i.e. showing no advantage with fusion[18,41,48]. In IS only the patient population of the present study includes a conservatively treated comparison group.
As a consequence of the reports on the positive short-term effects of lumbar fusion in our previous report on IS, the Swedish lumbar spine study and other studies with no controls, this treatment is used more and more frequently[48,108]. However, few long-term prospective randomized studies concerning DDD or IS have been reported. Furthermore, none of these have included control subjects treated conservatively[2,177].

Medium- and long-term assessment of the outcome of fusion

Both in connection with medium- (two-year) and long-term follow-up, most of the measures of outcome we employed here (i.e, the pain index, DRI, ODI, SF-36 and self-reported global outcome) were more favourable for patients who underwent fusion rather than conservative treatment with exercise[108]. However, as a result of the slight deterioration after fusion and the slight improvement following conservative treatment, the only statistically significant difference was in global outcome. The long-term subjective Quality of life (SF-36) was the same for both groups and considerably lower than that of the general population.

The reasons for this slight deterioration of outcome in the PLF group as compared to that of the EX group are unclear, but there are several theoretical possibilities:

- Regression of the symptoms of the group treated conservatively towards a more favourable mean.
  The symptoms of CLBP, like those of most diseases, vary in intensity and patients tend to consult spine surgeons (and thus get recruited for clinical studies) when their symptoms are more disturbing than usual. With time, the findings of follow-ups tend to be better than the original situation because of this common variability.
- A diminishing placebo effect.
  Surgery has, of course, a pronounced placebo effect that is probably stronger than that associated with conservative treatment. However, nine years after treatment the placebo effect has probably diminished considerably and become more even between the groups.
- Effects of aging on the outcome of fusion.
  Such an effect seems unlikely since the group conservatively treated in this randomized controlled trial did not deteriorate significantly and the age distribution within the two groups was similar (Table 1).
- Insufficient analytical power to detect differences between the two groups.

The power to a detect 15-point difference in the DRI values for the groups treated conservatively (n=29) or surgically (n=71) was calculated to be 95%. Exclusion of the 9 l patients treated conservatively who were operated on reduced this power to 89%. Since the power of this study to detect a 10-point difference on the 5% level was only 70%, minor differences between the treatment groups may not have been detected.

- Accelerated Adjacent Segment Degeneration following fusion. Development of ASD is an obvious possible reason for worsening outcome following fusion. Here ASD exerted a significant, albeit limited effect on outcome.

The difference in the global outcome self-reported by our two treatment groups was quite substantial. 76% of those in the fusion group rated themselves as "much better" or "better", whereas the corresponding value for the conservative group was only 50% indicating a clinically important difference in patient satisfaction. Hägg and colleagues[64] reported that sef-reported global outcome demonstrates significant correlations with pain (VAS), ODI and psychological distress. These investigators also found that the differences in global assessment between groups of patients tend to be greater than differences in other outcome parameters[64].

Interestingly, global outcome was also the only parameter associated with a statistically significant difference between patients with or without ASD in the present investigation, for reasons which remain unclear. Perhaps self-reported global outcome allows more sensitive detection of clinical improvement than can be obtained with other outcome parameters limited to specific aspects of pain or disability. On the other hand, global outcome may be more susceptible to motivational bias, such as having a close relationship with one's surgeon, as well as to the possible influence of the larger placebo effect of surgery, which is a more invasive and cumbersome treatment than exercise. In addition recall bias, i.e., unreliable memory of an earlier state of health, can lead to overestimation of the improvement experienced[5]. Although Hägg has shown that after a two-year period of follow-up, global outcome still provides a valid measure of outcome, no documentation concerning the validity for longer periods of follow-up, such as those used here, is presently available. Thus, a possible influence of in particular recall bias on our findings cannot be completely ruled out.

Radiodiographic signs of ASD and their effect on outcome

It is well-established that radiological degenerative changes in the spine detected radiologically have little, if any, impact on the patient's symptoms and, indeed, many such changes are present in asymptomatic individuals[13,15,172]. Similarly, many investigators have found no correlation between radiological signs of ASD and outcome[51,77,86,88,95,99,106,119,141,161], although there are a few reports of such a correlation[78,79,128].

According to two recent review articles[97,125], the reported prevalences of "radiographic" and "symptomatic" ASD range widely between 5.2% and 100%. These values are, however, based solely on retrospective studies and studies that did not include a valid control group treated conservatively and therefore, representing only level-4 evidence. Here, the scientifically valid randomized controlled trial described in Paper 4 revealed "radiographic" ASD with a prevalence of 6-38%, depending on the definition employed for this condition. Correlation with significant symptoms was,

however, marginal with only the less frequent, more severe forms of ASD appearing to affect outcome. In addition, with all definitions except that for mild ASD provided by the UCLA grading scale, the measurements of outcome used here were somewhat, albeit not statistically significantly worse for our patients defined as having ASD (Table 22).

We also discovered that reduction in the height of the posterior disc is the most common first radiographic sign of this degenerative process. This could reflect compensatory hyperextension, possibly triggered by fusion of the spondylolytic segment in a non-physiological kyphotic position. Favourable outcome has also been shown to be dependent on preservation of an appropriate sagittal alignment[84 169 92,117,124], which may prevent ASD. However, in a randomized controlled trial involving 130 patients Korsgaard and coworkers[87] found no correlation between lumbar lordosis and outcome.

In this context we also observed that patients whose operation involved performance of laminectomy were the ones who most commonly developed ASD. Possible explanations for this finding include the fact that loss of the structural support of the posterior column may result in posterior disc compression or, that loss of the posterior tension band may enhance instability and in this way give rise to ASD with time. The importance of this tension band is indicated by radiological findings that loss of the posterior ligamentous complex leads to instability in the motion segment of the spine[90].

Comparison of the long-term outcome following fusion with and without instrumentation

As has been reported for short-term follow-up we observed no difference in the long-term outcome following fusion with or without instrumentation. Moreover, radiographic measurements showed no significant differences in the risk for degeneration in the adjacent segment following these two procedures, in agreement with the conclusions of Wiltse[180]. Other earlier reports indicating a higher incidence of symptomatic ASD in patients fused with than without pedicle fixation[40,125,146] were not confirmed here. On the contrary, more of the patients in our non-instrumented group tended to develop ASD (Table 24).

Anterior support or additional interbody fusion

On theoretical grounds, it has been proposed that additional fusion of the anterior interbody should result in a more favourable outcome. The support provided by the anterior column (where 80% of the axial compression forces are absorbed)[67], indirect foraminal decompression, restoration of lordosis, reduction of the slippage via ligamentotaxis and removal of the degenerated disc which is supposed to generate pain have all been put forth in favour of anterior fusion, or PLIF [7,94,111,157]. A Danish comparison between 360˚ fusion and PLF revealed only a non-significant tendency towards better outcome with 360˚ fusion after two years[28] but, surprisingly, a significantly more favourable outcome for the 360˚-group in DDD 5-9 years post-operatively. It may be that anterior support in the long run prevents ASD because of better maintenance of lordosis [177].

Otherwise, scientific support for the advantageousness of anterior or 360˚ fusion is

weak and, indeed, the only two randomized controlled trials performed more recently in this area do not support such conclusion[49,80,85,89]. Our own comparison of 86 patients operated on with PLIF and 77 patients who underwent PLF, using exactly the same criteria for inclusion and follow-up procedures shows a strikingly similar outcome for these two groups with respect to all of the parameters assessed including pain, disability and global assessment. On the other hand, complications were more frequent in our PLIF group, in agreement with earlier reports[47,89,100]. Long term results on PLIF are, however, lacking and until then no definite conclusion on its usefulness can be drawn.

The selection of patients

Our present long-term investigation reveals that patients who undergo fusion rather than conservative treatment much more often consider themselves to be "much better". As pointed out by numerous investigators, it is essential for the surgeon to recommend fusion only for those patients who will most likely benefit from this procedure. Unfortunately, no definitive means of making such a selection are presently available, so the surgeon must to a large extent rely primarily on intuition.

We demonstrate here that work status, gender and exercise habits can help the surgeon make this selection, always bearing in mind that these three variables together could account for no more than 20% of the variation in the PI during follow-up. Alternatively, we also found that the nature of a pre-operative pain drawing is strongly associated with work status and these two can be considered to be confounding variables, strongly reflecting one another. This is the first time that, as indicated by univarite analysis, the nature of such a PD can be used to predict the outcome of fusion, i.e., the assessment of pain, function and global outcome was also more favourable for patients who produced such a drawing classified as organic. Previously Möller demonstrated that the nature of the pain drawing can be used to predict the outcome two years after performance of PLF (n=77)[107]. The present study is considerably larger (n=163), involving patients who underwent PLIF (Study population 2), in addition to those treated with PLF, as well as other spine-centers. This means that our present findings can be considered to be of more general validity.

In agreement with the conclusions of the Swedish lumbar spine study [63] our investigation demonstrates that patients with a non-organic pain drawing experience worse pain and more pronounced functional disability pre-operatively. Since we also found this to be the case after two years post-operatively, even if there was a certain degree of improvement, many of the symptoms experienced by this group persists.

In addition, our multivariate regression analysis revealed that work status is the strongest prognostic factor for outcome following fusion. Moreover, *the improvements* in both pain (PI) and function (DRI) were more pronounced for patients who were actively working prior to their operations, indicating that the prognosis provided by these factors does not simply reflect lower pre-operative values of PI and DRI . A similar conclusion was arrived at earlier by Anderson and co-workers with regards to discogenic low-back pain[3]. Also in agreement with our present observations, Gehrchen and colleagues showed that being actively at work pre-operatively and being male were predictive of a more favourable outcome [54].

The natural history of isthmic spondylolisthesis

Only two-thirds of the patients provided with an exercise program completed the one-year regimen of physiotherapy and, thereafter all of these patients were left on their own. The extent to which they continued their exercises is unknown but it seems likely that they gradually discontinued training, so that their outcome 9 and 13 years later can be considered to reflect the natural history of isthmic spondylolisthesis involving CLBP. In this case, this natural history involves continuing pain, along with severe impairment of functional ability and quality of life. We also conclude that the original pathogenesis is almost certainly the reason for the patients' continued problems, since no degenerative changes in the level adjacent to the slippage were observed in the conservative group. Seitsalo[144] has reported earlier that isthmic spondylolisthesis is associated with disc degeneration and spontaneous stabilization of the olisthetic segment and it can be speculated that this spontaneous stabilization is responsible for the (non-significant) minor improvement in the conservative group between short-term and long-term follow-up.

Application of our findings to CLBP in general

For comparison of different treatment procedures the study population of adult isthmic spondylolisthesis employed here has the obvious advantage of being more homogenous than the population of patients with DDD and CLBP, who probably suffer from a number of different causes of pain, including generalised pain syndrome. At the same time, the symptoms expected by these two groups are indistinguishable from one another [109] and, moreover, they respond similarly to fusion. For example Gehrchen[54] and colleagues found that the clinical outcome of spinal fusion in patients with isthmic spondylolisthesis and those with degenerative disc disease of the lumbar spine is identical. Therefore, it is likely that the overall findings described here can, at least to a large extent, be applied to the much larger general population of patients with DDD and CLBP.

Limitations of the present study

The major limitations associated with the long-term studies described in Paper 1 and 4 was the relatively limited number of patients treated with exercise which introduced the risk for a type-2 error. It should be remembered that the study was initiated in 1989, a time point when the positive effect of fusion was expected to be more pronounced than we today know it is. Accordingly, the study was powered to detect a 15 points change in DRI, at that point by us considered the minimally clinically important difference. We now know, however, that the effect of fusion is less pronounced and that studies have to be powered for smaller levels of change, and, hence, larger study populations. The power problem is even more pronounced for the ASD analysis, not a defined outcome measure at the initiation of the study, since it occurs in only a subgroup of patients, the size of which was unknown in 1989.

Despite the limited number of conservatively treated patients we did, however, find significant differences between the groups. This was true for the clinical as well as the radiological (ASD) outcome. Larger number of patients in the conservative group would of course have improved the potential for strong conclusions. Adequately

powered future studies may provide such data. It must, however, be realised that the difficulties with large RCTs with adequate long term follow-up rates (>90% in the present study) are monumental and that we most likely will have to depend on less than ideal data also in the future.

Another source of error is the drop out rate in the conservative group. In order to address the possibility of selection bias introduced by the exclusion of nine patients from the EX group we also analyzed this group with inclusion of the data from these patients in accordance with the intention-to-treat principle (Paper 1). This approach did not alter the clinical outcome for the EX group, i.e., the outcome was not improved by the removal of patients with a less favorable prognosis. Thus, although it cannot be excluded entirely this possible source of error does not seem to have influenced the conclusions arrived at.

The exclusion of the nine (26%) patients in the EX group who eventually underwent operation may further have unduly influenced our results on ASD, since these patients may have been generally more prone to disc degeneration (paper 4). Although this situation certainly renders interpretation more difficult, a drop out analysis suggest that the patients remaining in the EX group most likely are representative of the group originally assigned randomly to recive conservative treatment. Firstly, the patients excluded were operated on relatively soon (median time = 27 months) after inclusion in our study, most likely because the symptoms from their spondylolisthesis continued, and not because of degenerative changes at the level above, which require a longer period to develop[58]. Secondly, only two of these 9 patients were also fused above the spondylolisthetic level, i.e., essentially the same proportion as in the PLF group, not suggesting an early development of "ASD" among those excluded. In light of the considerations discussed above, however, further investigations in this controversial area of ASD are certainly warranted.

In summary

In the long term PLF results in a moderatly improved outcome compared to natural history of IS. The effect of posterolateral fusion is, however, somewhat limited, reflecting the need of improvements in fusion techniques and/or patient selection. In the short term PLIF does not seem to provide the answer. In the long term we still do not know whether PLIF, or other anterior interbody fusion strategies, improves outcome, as suggested by the Danish study on 360°-fusion using ALIF + PLF. Patient selection remains problematic. Although often stated as the key to a successful outcome only a small proportion (20%) of the variability of outcome was explained by working status, exercise habits and male gender. Although several methodological problems, such as study power, disc degeneration measurements and drop out frequency, was associated with the ASD study, the data does demonstrate the existence of ASD. Its clinical importance was more difficult to demonstrate, but seemed limited.

Important areas in future research on surgical treatment in adult IS, as well as DDD, are the long term effect of interbody fusion; whether it improves lordosis and sagittal balance, whether it thereby prevents ASD, and ultimately whether it improves long term outcome.

1.9 CONCLUSIONS

The major conclusions to be drawn from the investigations described in this thesis are the following:

1. The natural course of symptomatic spondylolisthesis in adults does not involve spontaneous improvement, but rather continuing disability that adversely affects many aspects of life over a period of many years.

2. The long-term outcome following fusion is moderately better than that achieved with a one-year exercise program.

3. Instrumentation in connection with fusion has no influence either positive or negative on the long-term outcome.

4. The additional anterior fusion involved in the PLIF procedure does not improve the medium-term outcome in comparison to the "gold standard" PLF.

5. The patients pre-operative work status (actively work or not) is a major predictor of outcome following lumbar fusion. Gender and exercise exert a smaller, but nonetheless significant impact on this outcome.

6. Degenerative changes at the level adjacent to the spondylolisthesis were more rapid in patients subjected to fusion than in patients treated conservatively.

7. The use of instrumentation in connection with fusion does not affect subsequent development of ASD.

8. Except in its rarer, more severe forms, ASD exerted only marginal effects on the outcome for most patients.

9. In connection with ASD the height of the posterior disc is affected more frequently than that of the anterior disc height.

2 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all who have contributed to the research described in this thesis, as well as of you others who have supported me during my years of doctorial studies. I offer special thanks to:

Rune Hedlund, my friend and tutor in connection with both research and spinal surgery, for your constant helpful support. Without you, this book would never have been written (and no skiing enjoyed).

Hans Möller, also my tutor and generous friend, who always encourages me to perform additional and thorough hard work.

My other co-authors-- Adel Shalabi, Yiang XiaoYu, Tycho Tullberg and Pavel Neumann - -for invaluable and enjoyable collaboration.

Helene Fils, Christina Ekman, Ulrica Skoog and Anneli Andersson, for their invaluable secretarial help, often at the most inconvenient of times.

Hans Pettersson and Per Näsman for statistical help and Joe DePierre for linguistic assistance.

Anita Hansson, for all the x-rays taken on sometimes reluctant patients.

All of the patients that I have bothered time and again and again and again and again….

Göran Modin, for continuing to employ me at one of the "leanest" orthopaedic clinics in Scandinavia.

Ricard Miedel and Simon Printz, for so willingly allowing me time to do research and attend congresses (and go on vacation now and then).

Ulric Willers and Peter Elkan, for being such tolerant and good friends and colleagues at the Spine Unit. I will start working again any day now!

Sari Ponzer, for invaluable help, e.g., with the SF-36 calculations and Urban Lindgren for wise advise concerning, among other things, the labyrinths of the doctoral and institutional bureaucracies.

My colleagues and the staff at the Orthopaedic Department of Södersjukhuset and the Spine Unit of Karolinska University Hospital Huddinge.

My father Gunnar and mother Carin, for their lifelong support and for always believing in me. Sorry that I haven't called you so often lately. I promise improvement now!

Eva, for putting up with a hopeless man.

Anton, Lisa and Maja, for putting up with a father who hasn't had much time to do things that are fun.

All the rest of my relatives and friends, who have patiently endured my anti-social behaviour during the past few years.

3 SUMMARY IN SWEDISH

3.1 STELOPERATION MOT KRONISK LÄNDRYGGSVÄRK VID KOTFÖRSKJUTNING AV ISTHMISK TYP

Isthmisk spondylolisthes (IS), dvs. kotförskjutning pga. ett avbrott på kotbågen, finns hos c:a 5% av befolkningen. Symtomen varierar från ett besvärsfritt tillstånd till mycket svåra besvär. Besvärsbilden är densamma som hos den betydligt större gruppen av patienter med ospecifik kronisk ländryggsvärk, ofta då tillskrivet diskdegeneration, d.v.s. förslitning av mellankotskivorna (DDD), som orsak. En positiv effekt av steloperation (fusion) på kort sikt (2år) har visats för IS och DDD. Effekten på lång sikt liksom den möjliga förekomsten av en accelererad diskdegeneration av segmentet närmast invid steloperationen (ASD) är dock okänd.

De huvudsakliga målen med detta avhandlingsarbete var att bestämma långtidsresultatet för ländryggsfusion (med eller utan tillägg av instrumentering med skruvar och stag) jämfört med sjukgymnastisk behandling samt effekten av steloperation på angränsande disk inklusive den kliniska betydelsen av eventuell uppkommen ASD. En ytterligare målsättning var att bestämma om en adderad främre fusion med förbening mellan kotkropparna (PLIF) förbättrar resultatet på kort sikt (upp till 2 år) jämfört med en steloperation mellan de bakre kotbågarna, benämnt posterolateral fusion (PLF). En ytterligare målsättning var att hitta faktorer med vilka man kan förutsäga och i vilket avseende de påverkade slutresultatet efter en steloperation i ländryggen. Detta för att om möjligt kunna välja ut patienter som har störst vinst av en sådan steloperation.

Två patientmaterial innefattas i avhandlingsarbetet. I det första inkluderades 111 patienter med IS, ålder 18-55 år och uttalade besvär sen minst 1 år tillbaka. Patienterna lottades (randomiserades) till sjukgymnastbehandling (EX, n=34 st) eller posterolateral fusion (PLF, n= 77) med (n=37) respektive utan instrumentering (n=40). Det andra patientmaterialet bestod av 86 patienter, med exakt samma intagnings- och uteslutningskriterier, som opererades med PLIF.

Samma utvärderingsmått för att bestämma slutresultatet användes på båda grupperna: Smärta bestämd genom en så kallad VAS-skala (0-100), funktionsmåtten Disability Rating Index (DRI) och Oswestry Disability Index (ODI), arbetsstatus och patientens egen skattning av behandlingsresultatet till mycket förbättrad, förbättrad, oförändrad eller sämre. Livskvalité bestämdes med hjälp av SF-36 frågeformulär. Före operation analyserades även patientens smärtritning (PD) och klassificerades efter det generella intrycket till en organisk (O) eller en icke organisk (NO) grupp. Tre röntgenologiska mätmetoder användes för att kvantifiera ASD: två digitala röntgenologiska mätmetoder och en diskdegenerations-graderingsskala då förutom uppskattad diskhöjd även benpålagringar och ökad bentäthet av övre och angränsande undre kotkroppskanter vägs in (UCLA-skalan).

Uppföljningsfrekvensen var 91% (101/111) efter i medeltal 9 år (5–13år) för långtidsstudien (Delarbete1) och för röntgenologiska långtidsuppföljningen avseende ASD 72 % (80/111) efter i medel 13 år (10-17år) (Delarbete 4). Uppföljningsfrekvensen var 98% (84/86) för PLIF-gruppen och 97 % (75/77) för PLF-gruppen efter 2 år (Delarbete 2 och 3).

Vid långtidsuppföljningen efter 9 år rapporterade PLF-gruppen ett signifikant bättre resultat jämfört med den sjukgymnastiskt behandlade icke opererade EX-gruppen då 76% av de kirurgiskt behandlade ansåg sig vara mycket bättre eller bättre jämfört med 50% av EX-patienterna (p=0.015). Alla andra utfallsmått var bättre för den

kirurgiskt behandlade gruppen men av icke statistiskt signifikant/säkerställd grad. Livskvalitén bestämd genom SF-36 var avsevärt lägre för både de kirurgiskt och konservativt behandlade grupperna jämfört med en svensk normalpopulation.

Inga signifikanta skillnader av långtidsutfallet sågs mellan de instrumenterade och icke instrumenterade patienterna både vad gäller symptom och röntgenologiska ASD-fynd. Hos de stelopererade patienter som man också tagit bort en stor del av bakre kotbågen av för att avlasta nerverna (d.v.s. dekomprimerats med laminektomi) visade sig i högre grad utveckla ASD, definierat enligt UCLA-skalan (p=0.015).

Vid i medeltal 13 år efter operation visade den ena digitala mätmetoden en medeldisksänkning på 2% i EX-gruppen och 15% i PLF-gruppen (p=0.0016), och den andra mätmetoden visade 0.5 mm mer disksänkning i PLF-gruppen jämfört med EX-gruppen (ns). UCLA-skalan visade normala angränsande disker hos alla i EX-gruppen jämfört med ASD hos 38% av de i PLF-gruppen (n= 0.026). Då ASD var definierat som mer disksänkning än 2 standardavvikelser av den som observerats i den konservativt behandlade gruppen (=naturalförloppet), ansåg sig 11% av de fusionerade patienterna med ASD jämfört med 49% av de utan ASD vara mycket bättre (p=0.036). I övrigt var det ingen statistskt signifikant skillnad i smärta eller funktion mellan de med respektive utan röntgenologiskt konstaterad ASD.

Efter 2 år var utfallsmåtten nästan identiska för PLIF- och PLF-gruppen: medelsmärtan var 35 mot 37 (ns) i respektive grupp, medel-DRI 30 mot 29 (ns). Medel-ODI var identiska och patientens egen skattning av behandlingsresultatet (i mycket bättre – sämre) var mycket likartad. Det inträffade 4 (5%) större komplikationer i PLF-gruppen jämfört med 12 (14%) i PLIF-gruppen (p=0.06).

Vid analysen för att finna faktorer som kan förutsäga slutresultatet 2 år efter steloperation fann vi att 54% av patienterna med en O smärtritning var mycket bättre jämfört med bara 33% av de i NO-gruppen (p=0.038). Även medel-smärtan, DRI och ODI var signifikant bättre i O-gruppen. Multivariat regressionsanalys, då flera faktorer vägs in i analysen, visade dock att 3 andra faktorer då var mer betydelsefulla; vara i arbete, manligt kön och regelbundna motionsvanor var faktorer som förutspådde ett bättre slutresultat.

Sammanfattningsvis så kan man dra slutsatserna att posterolateral fusion resulterar i ett måttligt förbättrat långtidsresultat jämfört med konservativ behandling i form av 1 års sjukgymnastbehandling. Eftersom sjukgymnastbehandlingen var begränsad till denna korta tid så reflekterar sjukgymnastgruppen sannolikt det normala åldringsförloppet. Därav kan ingen större förbättring förväntas med tiden vid IS.

Radiologiska långtidsstudien visade att steloperation accelererade degenerativa förändringar på angränsande disknivå jämfört med naturalförloppet. Dessutom visade studien att samtidig dekompression med laminektomi verkar bidra till utvecklingen av ASD. Den kliniska betydelsen av ASD verkade dock begränsad då endast de mer uttalade formerna av ASD påverkade slutresultatet. Resultaten visade också att instrumentering inte påverkade vare sig utveckling av ASD eller slutresultatet.

Trotts fler komplikationer gav tillägg av en främre fusion mellan kotkropparna (PLIF) samma resultat, avseende smärta och funktion, som sedvanlig bakre fusion (PLF) efter 2 år.

En stor del av behandlingsresultatet kunde ej förklaras av de studerade prediktiva faktorerna, emellertid indikerade våra resultat att jämfört med ej arbetande och icke motionerande kvinnor, hade fysiskt aktiva män som inte var sjukskrivna en bättre chans till större förbättring efter steloperation.

4 REFERENCES

1. Aalto TJ, Malmivaara A, Kovacs F, et al. Preoperative predictors for postoperative clinical outcome in lumbar spinal stenosis: a systematic review. *Spine* 2006;31:E648-63.
2. Andersen T, Videbaek TS, Hansen ES, et al. The positive effect of posterolateral lumbar spinal fusion is preserved at long-term follow-up: a RCT with 11-13 year follow-up. *Eur Spine J* 2008;17:272-80.
3. Anderson PA, Schwaegler PE, Cizek D, et al. Work status as a predictor of surgical outcome of discogenic low back pain. *Spine* 2006;31:2510-5.
4. Asche CV, Kirkness CS, McAdam-Marx C, et al. The societal costs of low back pain: data published between 2001 and 2007. *J Pain Palliat Care Pharmacother* 2007;21:25-33.
5. Aseltine RH, Jr., Carlson KJ, Fowler FJ, Jr., et al. Comparing prospective and retrospective measures of treatment outcomes. *Med Care* 1995;33:AS67-76.
6. Axelsson P, Johnsson R, Stromqvist B. The spondylolytic vertebra and its adjacent segment. Mobility measured before and after posterolateral fusion. *Spine* 1997;22:414-7.
7. Barrick WT, Schofferman JA, Reynolds JB, et al. Anterior lumbar fusion improves discogenic pain at levels of prior posterolateral fusion. *Spine* 2000;25:853-7.
8. Bell DF, Ehrlich MG, Zaleske DJ. Brace treatment for symptomatic spondylolisthesis. *Clin Orthop Relat Res* 1988;236:192-8.
9. Beutler WJ, Fredrickson BE, Murtland A, et al. The natural history of spondylolysis and spondylolisthesis: 45-year follow-up evaluation. *Spine* 2003;28:1027-35; discussion on p.35.
10. Blanda J, Bethem D, Moats W, et al. Defects of pars interarticularis in athletes: a protocol for nonoperative treatment. *J Spinal Disord* 1993;6:406-11.
11. Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. *Spine* 1993;18:1186-9.
12. Bolton JE, Christensen MN. Back pain distribution patterns: relationship to subjective measures of pain severity and disability. *J Manipulative Physiol Ther* 1994;17:211-8.
13. Bono CM. Low-back pain in athletes. *J Bone Joint Surg Am* 2004;86A:382-96.
14. Boos N, Weissbach S, Rohrbach H, et al. Classification of age-related changes in lumbar intervertebral discs: 2002 Volvo Award in basic science. *Spine* 2002;27:2631-44.
15. Borenstein D. Does osteoarthritis of the lumbar spine cause chronic low back pain? *Curr Pain Headache Rep* 2004;8:512-7.
16. Boxall D, Bradford DS, Winter RB, et al. Management of severe spondylolisthesis in children and adolescents. *J Bone Joint Surg Am* 1979;61:479-95.
17. Brodsky AE, Kovalsky ES, Khalil MA. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. *Spine* 1991;16:S261-5.
18. Brox JI, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913-21.

19. Brox JI, Storheim K, Grotle M, et al. Systematic review of back schools, brief education, and fear-avoidance training for chronic low back pain. *Spine J* 2007.
20. Buck JE. Direct repair of the defect in spondylolisthesis. Preliminary report. *J Bone Joint Surg Br* 1970;52:432-7.
21. Bushell GR, Ghosh DP, Taylor TK, et al. The effect of spinal fusion on the collagen and proteoglycans of the canine intervertebral disc. *J Surg Res* 1978;25:61-9.
22. Carnes D, Ashby D, Underwood M. A systematic review of pain drawing literature: should pain drawings be used for psychologic screening? *Clin J Pain* 2006;22:449-57.
23. Carragee EJ. Single-level posterolateral arthrodesis, with or without posterior decompression, for the treatment of isthmic spondylolisthesis in adults. A prospective, randomized study. *J Bone Joint Surg Am* 1997;79:1175-80.
24. Carreon LY, Glassman SD, Djurasovic M. Reliability and agreement between fine-cut CT scans and plain radiography in the evaluation of posterolateral fusions. *Spine J* 2007;7:39-43.
25. Cavalier R, Herman MJ, Cheung EV, et al. Spondylolysis and spondylolisthesis in children and adolescents: I. Diagnosis, natural history, and nonsurgical management. *J Am Acad Orthop Surg* 2006;14:417-24.
26. Chan CW, Goldman S, Ilstrup DM, et al. The pain drawing and Waddell's nonorganic physical signs in chronic low-back pain. *Spine* 1993;18:1717-22.
27. Chen CS, Cheng CK, Liu CL, et al. Stress analysis of the disc adjacent to interbody fusion in lumbar spine. *Med Eng Phys* 2001;23:483-91.
28. Christensen FB, Hansen ES, Eiskjaer SP, et al. Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. *Spine* 2002;27:2674-83.
29. Dahl B, Gehrchen PM, Kiaer T, et al. Nonorganic pain drawings are associated with low psychological scores on the preoperative SF-36 questionnaire in patients with chronic low back pain. *Eur Spine J* 2001;10:211-4.
30. Davis H. Increasing rates of cervical and lumbar spine surgery in the United States, 1979-1990. *Spine* 1994;19:1117-23; discussion on pp. 23-4.
31. Dekutoski MB, Schendel MJ, Ogilvie JW, et al. Comparison of in vivo and in vitro adjacent segment motion after lumbar fusion. *Spine* 1994;19:1745-51.
32. den Boer JJ, Oostendorp RA, Beems T, et al. A systematic review of bio-psychosocial risk factors for an unfavourable outcome after lumbar disc surgery. *Eur Spine J* 2006;15:527-36.
33. DeWald CJ, Vartabedian JE, Rodts MF, et al. Evaluation and management of high-grade spondylolisthesis in adults. *Spine* 2005;30:S49-59.
34. DeWald RL, Faut MM, Taddonio RF, et al. Severe lumbosacral spondylolisthesis in adolescents and children. Reduction and staged circumferential fusion. *J Bone Joint Surg Am* 1981;63:619-26.
35. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine* 1998;23:2003-13.
36. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;12:142S-58S.
37. Deyo RA, Mirza SK. Trends and variations in the use of spine surgery. *Clin Orthop Relat Res* 2006;443:139-46.
38. Eck JC, Humphreys SC, Hodges SD. Adjacent-segment degeneration after lumbar fusion: a review of clinical, biomechanical, and radiologic studies. *Am J Orthop* 1999;28:336-40.

39. Ekman P, Moller H, Tullberg T, et al. Posterior lumbar interbody fusion versus posterolateral fusion in adult isthmic spondylolisthesis. *Spine* 2007;32:2178-83.
40. Etebar S, Cahill DW. Risk factors for adjacent-segment failure following lumbar fixation with rigid instrumentation for degenerative instability. *J Neurosurg* 1999;90:163-9.
41. Fairbank J, Frost H, Wilson-MacDonald J, et al. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Br'Med J* 2005;330:1233.
42. Fairbank JCT DJ, Couper J, O´Brien JP. The Oswestry low back pain disability quetionaire. *Physiotherapy* 1980;66:271-3.
43. Fogel GR, Toohey JS, Neidre A, et al. Fusion assessment of posterior lumbar interbody fusion using radiolucent cages: X-ray films and helical computed tomography scans compared with surgical exploration of fusion. *Spine J* 2007.
44. France JC, Yaszemski MJ, Lauerman WC, et al. A randomized prospective study of posterolateral lumbar fusion. Outcomes with and without pedicle screw instrumentation. *Spine* 1999;24:553-60.
45. Fredrickson BE, Baker D, McHolick WJ, et al. The natural history of spondylolysis and spondylolisthesis. *J Bone Joint Surg Am* 1984;66:699-707.
46. Frennered K. Isthmic spondylolisthesis among patients receiving disability pension under the diagnosis of chronic low back pain syndromes. *Spine* 1994;19:2766-9.
47. Fritzell P, Hagg O, Nordwall A. Complications in lumbar fusion surgery for chronic low back pain: comparison of three surgical techniques used in a prospective randomized study. A report from the Swedish Lumbar Spine Study Group. *Eur Spine J* 2003;12:178-89.
48. Fritzell P, Hagg O, Wessberg P, et al. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521-32; discussion on pp.32-4.
49. Fritzell P, Hagg O, Wessberg P, et al. Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish lumbar spine study group. *Spine* 2002;27:1131-41.
50. Frobin W, Brinckmann P, Kramer M, et al. Height of lumbar discs measured from radiographs compared with degeneration and height classified from MR images. *Eur Radiol* 2001;11:263-9.
51. Frymoyer JW, Hanley EN, Jr., Howe J, et al. A comparison of radiographic findings in fusion and nonfusion patients ten or more years following lumbar disc surgery. *Spine* 1979;4:435-40.
52. Gaines RW. L5 vertebrectomy for the surgical treatment of spondyloptosis: thirty cases in 25 years. *Spine* 2005;30:S66-70.
53. Ganju A. Isthmic spondylolisthesis. *Neurosurg Focus* 2002;13:E1.
54. Gehrchen PM, Dahl B, Katonis P, et al. No difference in clinical outcome after posterolateral lumbar fusion between patients with isthmic spondylolisthesis and those with degenerative disc disease using pedicle screw instrumentation: a comparative study of 112 patients with 4 years of follow-up. *Eur Spine J* 2002;11:423-7.
55. Ghiselli G, Wang JC, Bhatia NN, et al. Adjacent segment degeneration in the lumbar spine. *J Bone Joint Surg Am* 2004;86A:1497-503.
56. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine* 1999;24:1820-32.
57. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: updated Cochrane Review. *Spine* 2005;30:2312-20.

58. Gillet P. The fate of the adjacent motion segments after lumbar fusion. *J Spinal Disord Tech* 2003;16:338-45.
59. Goto K, Tajima N, Chosa E, et al. Effects of lumbar spinal fusion on the other lumbar intervertebral levels (three-dimensional finite element analysis). *J Orthop Sci* 2003;8:577-84.
60. Gutke A, Josefsson A, Oberg B. Pelvic girdle pain and lumbar pain in relation to postpartum depressive symptoms. *Spine* 2007;32:1430-6.
61. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *J Am Med Assoc* 1994;271:59-63.
62. Ha KY, Schendel MJ, Lewis JL, et al. Effect of immobilization and configuration on lumbar adjacent-segment biomechanics. *J Spinal Disord* 1993;6:99-105.
63. Hagg O, Fritzell P, Hedlund R, et al. Pain-drawing does not predict the outcome of fusion surgery for chronic low-back pain: a report from the Swedish Lumbar Spine Study. *Eur Spine J* 2003;12:2-11.
64. Hagg O, Fritzell P, Oden A, et al. Simplifying outcome measurement: evaluation of instruments for measuring outcome after fusion surgery for chronic low back pain. *Spine* 2002;27:1213-22.
65. Hamlin C, Hitchcock M, Hofmeister J, et al. Predicting surgical outcome for pain relief and return to work. *Best Pract Benchmarking Healthcare* 1996;1:311-4.
66. Haraldsson S, Willner S. A comparative study of spondylolisthesis in operations on adolescents and adults. *Arch Orthop Trauma Surg* 1983;101:101-5.
67. Harms. Screw-threaded rod system in spinal surgery. *In SPINE: State of the Art Reviews, Vol 6, No. 3* 1992.
68. Harris IE, Weinstein SL. Long-term follow-up of patients with grade-III and IV spondylolisthesis. Treatment with and without posterior fusion. *J Bone Joint Surg Am* 1987;69:960-9.
69. Haukipuro K, Keranen N, Koivisto E, et al. Familial occurrence of lumbar spondylolysis and spondylolisthesis. *Clin Genet* 1978;13:471-6.
70. Helenius I, Lamberg T, Osterman K, et al. Posterolateral, anterior, or circumferential fusion in situ for high-grade spondylolisthesis in young patients: a long-term evaluation using the Scoliosis Research Society questionnaire. *Spine* 2006;31:190-6.
71. Hensinger RN. Spondylolysis and spondylolisthesis in children. *Instr Course Lect* 1983;32:132-51.
72. Hensinger RN. Spondylolysis and spondylolisthesis in children and adolescents. *J Bone Joint Surg Am* 1989;71:1098-107.
73. Herman MJ, Pizzutillo PD. Spondylolysis and spondylolisthesis in the child and adolescent: a new classification. *Clin Orthop Relat Res* 2005:46-54.
74. Herzog RJ, Marcotte PJ. Assessment of spinal fusion. Critical evaluation of imaging techniques. *Spine* 1996;21:1114-8.
75. Howe J, Frymoyer JW. The effects of questionnaire design on the determination of end results in lumbar spinal surgery. *Spine* 1985;10:804-5.
76. Huskisson EC. Measurement of pain. *Lancet* 1974;2:1127-31.
77. Ishihara H, Osada R, Kanamori M, et al. Minimum 10-year follow-up study of anterior lumbar interbody fusion for isthmic spondylolisthesis. *J Spinal Disord* 2001;14:91-9.
78. Iwamoto J, Abe H, Tsukimura Y, et al. Relationship between radiographic abnormalities of lumbar spine and incidence of low back pain in high school and college football players: a prospective study. *Am J Sports Med* 2004;32:781-6.
79. Iwamoto J, Abe H, Tsukimura Y, et al. Relationship between radiographic abnormalities of lumbar spine and incidence of low back

pain in high school rugby players: a prospective study. *Scand J Med Sci Sports* 2005;15:163-8.
80. Jacobs WC, Vreeling A, De Kleuver M. Fusion for low-grade adult isthmic spondylolisthesis: a systematic review of the literature. *Eur Spine J* 2006 Apr;15(4):391-402.
81. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *Bmj* 2001;323:42-6.
82. Kant AP, Daum WJ, Dean SM, et al. Evaluation of lumbar spine fusion. Plain radiographs versus direct surgical exploration and observation. *Spine* 1995;20:2313-7.
83. Katz JN. Lumbar spinal fusion. Surgical rates, costs, and complications. *Spine* 1995;20:78S-83S.
84. Kawakami M, Tamaki T, Ando M, et al. Lumbar sagittal balance influences the clinical outcome after decompression and posterolateral spinal fusion for degenerative lumbar spondylolisthesis. *Spine* 2002;27:59-64.
85. Kim KT, Lee SH, Lee YH, et al. Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. *Spine* 2006;31:1351-7; discussion on p. 8.
86. Korovessis P, Papazisis Z, Koureas G, et al. Rigid, semirigid versus dynamic instrumentation for degenerative lumbar spinal stenosis: a correlative radiological and clinical analysis of short-term results. *Spine* 2004;29:735-42.
87. Korsgaard M, Christensen FB, Thomsen K, et al. The influence of lumbar lordosis on spinal fusion and functional outcome after posterolateral spinal fusion with and without pedicle screw instrumentation. *J Spinal Disord Tech* 2002;15:187-92.
88. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. *Eur Spine J* 2001;10:309-13.
89. Kwon BK, Albert TJ. Adult low-grade acquired spondylolytic spondylolisthesis: evaluation and management. *Spine* 2005;30:S35-41.
90. Lai PL, Chen LH, Niu CC, et al. Relation between laminectomy and development of adjacent segment instability after lumbar fusion with pedicle fixation. *Spine* 2004;29:2527-32; discussion on p. 32.
91. Larsen JM, Rimoldi RL, Capen DA, et al. Assessment of pseudarthrosis in pedicle screw fusion: a prospective study comparing plain radiographs, flexion/extension radiographs, CT scanning, and bone scintigraphy with operative findings. *J Spinal Disord* 1996;9:117-20.
92. Lazennec JY, Ramare S, Arafati N, et al. Sagittal alignment in lumbosacral fusion: relations between radiological parameters and pain. *Eur Spine J* 2000;9:47-55.
93. Lee CK. Accelerated degeneration of the segment adjacent to a lumbar fusion. *Spine* 1988;13:375-7.
94. Lee CK, Vessa P, Lee JK. Chronic disabling low back pain syndrome caused by internal disc derangements. The results of disc excision and posterior lumbar interbody fusion. *Spine* 1995;20:356-61.
95. Lehmann TR, Spratt KF, Tozzi JE, et al. Long-term follow-up of lower lumbar fusion patients. *Spine* 1987;12:97-104.
96. Lenke LG, Bridwell KH, Bullis D, et al. Results of in situ fusion for isthmic spondylolisthesis. *J Spinal Disord* 1992;5:433-42.
97. Levin DA, Hale JJ, Bendo JA. Adjacent segment degeneration following spinal fusion for degenerative disc disease. *Bull NYU Hosp Jt Dis* 2007;65:29-36.
98. Levy HI, Hanscom B, Boden SD. Three-question depression screener used for lumbar disc herniations and spinal stenosis. *Spine* 2002;27:1232-7.
99. Luk KD, Lee FB, Leong JC, et al. The effect on the lumbosacral spine of long spinal fusion for idiopathic scoliosis. A minimum 10-year follow-up. *Spine* 1987;12:996-1000.

100. Madan S, Boeree NR. Outcome of posterior lumbar interbody fusion versus posterolateral fusion for spondylolytic spondylolisthesis. *Spine* 2002;27:1536-42.
101. Main CJ, Wood PL, Hollis S, et al. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine* 1992;17:42-52.
102. Marchetti PG, Bartolozzi P. *Classification of spondylolisthesis as a guideline for treatmen*ted: In Bridwell KH, De Wald RL (eds). The Textbook of Spinal Surgery, 2nd ed. Philadelphia, Lippincott-Raven: pp 1211-54, 1997.
103. Margolis RB, Tait RC, Krause SJ. A rating system for use with patient pain drawings. *Pain* 1986;24:57-65.
104. Meyerding HW. Spondylolisthesis. *Surg Gynecol Obstet ;* 1932;54:371-7.
105. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine* 2007;32:816-23.
106. Miyakoshi N, Abe E, Shimada Y, et al. Outcome of one-level posterior lumbar interbody fusion for spondylolisthesis and postoperative intervertebral disc degeneration adjacent to the fusion. *Spine* 2000;25:1837-42.
107. Moller H. Isthmic Spondylolisthesis in Adults. A Randomized Controlled Trial. PhD. Thesis: Stockholm, Sweden: Karolinska Institutet, 1999.
108. Moller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis--a prospective randomized study: part 1. *Spine* 2000;25:1711-5.
109. Moller H, Sundin A, Hedlund R. Symptoms, signs, and functional disability in adult spondylolisthesis. *Spine* 2000;25:683-9; discussion on p. 90.
110. Morscher E, Gerber B, Fasel J. Surgical treatment of spondylolisthesis by bone grafting and direct stabilization of spondylolysis by means of a hook screw. *Arch Orthop Trauma Surg* 1984;103:175-8.
111. Mulholland RC, Sengupta DK. Rationale, principles and experimental evaluation of the concept of soft stabilization. *Eur Spine J* 2002;11 Suppl 2:S198-S205.
112. Nachemson A, Zdeblick TA, O'Brien JP. Lumbar disc disease with discogenic pain. What surgical treatment is most effective? *Spine* 1996;21:1835-8.
113. Nachemson AL. Newest knowledge of low back pain. A critical look. *Clin Orthop Relat Res* 1992:8-20.
114. Nicol RO, Scott JH. Lytic spondylolysis. Repair by wiring. *Spine* 1986;11:1027-30.
115. Nordlund A, Waddell G. Cost of back pain in some OECD Countries. In Nachemson A, Jonsson E (eds.) *Neck and Back Pain. The Scientific Evidence of Causes, Diagnosis and Treatment.* Philadelphia: Lippincot Williams & Wilkins, 2000:pp.421-5.
116. O'Sullivan PB, Phyty GD, Twomey LT, et al. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine* 1997;22:2959-67.
117. Oda I, Cunningham BW, Buckley RA, et al. Does spinal kyphotic deformity influence the biomechanical characteristics of the adjacent motion segments? An in vivo animal model. *Spine* 1999;24:2139-46.
118. Ohlund C, Eek C, Palmbald S, et al. Quantified pain drawing in subacute low back pain. Validation in a nonselected outpatient industrial sample. *Spine* 1996;21:1021-30; discussion on p.31.
119. Okuda S, Iwasaki M, Miyauchi A, et al. Risk factors for adjacent segment degeneration after PLIF. *Spine* 2004;29:1535-40.

120. Olsewski JM, Schendel MJ, Wallace LJ, et al. Magnetic resonance imaging and biological changes in injured intervertebral discs under normal and increased mechanical demands. *Spine* 1996;21:1945-51.
121. Osterman K, Lindholm TS, Laurent LE. Late results of removal of the loose posterior element (Gill's operation) in the treatment of lytic lumbar spondylolisthesis. *Clin Orthop Relat Res* 1976:121-8.
122. Osterman K, Schlenzka D, Poussa M, et al. Isthmic spondylolisthesis in symptomatic and asymptomatic subjects, epidemiology, and natural history with special reference to disk abnormality and mode of treatment. *Clin Orthop Relat Res* 1993:65-70.
123. Pande KC, Tripathi S, Kanoi R. Limited clinical utility of pain drawing in assessing patients with low back pain. *J Spinal Disord Tech* 2005;18:160-2.
124. Park JY, Cho YE, Kuh SU, et al. New prognostic factors for adjacent-segment degeneration after one-stage 360 degrees fixation for spondylolytic spondylolisthesis: special reference to the usefulness of pelvic incidence angle. *J Neurosurg Spine* 2007;7:139-44.
125. Park P, Garton HJ, Gala VC, et al. Adjacent segment disease after lumbar or lumbosacral fusion: review of the literature. *Spine* 2004;29:1938-44.
126. Pearson A SD, Wharton N , Genuario J, et al. Precision of lumbar intervertebral measurements: Does a computer assisted technique improve reliability? The International Society for the Study of the Lumbar Spine (ISSLS). Hong Kong, China 2007.
127. Phillips FM, Reuben J, Wetzel FT. Intervertebral disc degeneration adjacent to a lumbar fusion. An experimental rabbit model. *J Bone Joint Surg Br* 2002;84:289-94.
128. Pye SR, Reid DM, Smith R, et al. Radiographic features of lumbar disc degeneration and self-reported back pain. *J Rheumatol* 2004;31:753-8.
129. Ransford AO, Cairns D, Mooney V. The pain drawing as an aid to the psychological evaluation of patients with low-back pain. *Spine* 1976;1:127-34.
130. Robaina-Padron FJ. [Controversies about instrumented surgery and pain relief in degenerative lumbar spine pain. Results of scientific evidence]. [Spanish] *Neurocirugia (Astur)* 2007;18:406-13.
131. Roberts N, Smith R, Bennett S, et al. Health beliefs and rehabilitation after lumbar disc surgery. *J Psychosom Res* 1984;28:139-44.
132. Roche MB, Rowe GG. The incidence of separate neural arch and coincident bone variations; a survey of 4,200 skeletons. *Anat Rec* 1951;109:233-52.
133. Rosenberg NJ, Bargar WL, Friedman B. The incidence of spondylolysis and spondylolisthesis in nonambulatory patients. *Spine* 1981;6:35-8.
134. Rowe GG, Roche MB. The etiology of separate neural arch. *J Bone Joint Surg Am* 1953;35A:102-10.
135. Ryden O, Lindal E, Uden A, et al. Differentiation of back pain patients using a pain questionnaire. *Scand J Rehabil Med* 1985;17:155-61.
136. Salen BA, Spangfort EV, Nygren AL, et al. The Disability Rating Index: an instrument for the assessment of disability in clinical settings. *J Clin Epidemiol* 1994;47:1423-35.
137. Santos ER, Goss DG, Morcom RK, et al. Radiologic assessment of interbody fusion using carbon fiber cages. *Spine* 2003;28:997-1001.
138. Saraste H, Brostrom LA, Aparisi T. Radiographic assessment of anatomic deviations in lumbar spondylolysis. *Acta Radiol Diagn (Stockh)* 1984;25:317-23.
139. Schlenzka D, Poussa M, Seitsalo S, et al. Intervertebral disc changes in adolescents with isthmic spondylolisthesis. *J Spinal Disord* 1991;4:344-52.
140. Schlenzka D, Remes V, Helenius I, et al. Direct repair for treatment of symptomatic spondylolysis and low-grade isthmic spondylolisthesis in young patients: no benefit in comparison to segmental fusion after a mean follow-up of 14.8 years. *Eur Spine J* 2006;15:1437-47.

141. Schulte TL, Leistra F, Bullmann V, et al. Disc height reduction in adjacent segments and clinical outcome 10 years after lumbar 360 degrees fusion. *Eur Spine J* 2007 Dec;16(12):2152-8.
142. Schwartz DP, DeGood DE. Global appropriateness of pain drawings: blind ratings predict patterns of psychological distress and litigation status. *Pain* 1984;19:383-8.
143. Seitsalo S, Osterman K, Hyvarinen H, et al. Progression of spondylolisthesis in children and adolescents. A long-term follow-up of 272 patients. *Spine* 1991;16:417-21.
144. Seitsalo S, Schlenzka D, Poussa M, et al. Disc degeneration in young patients with isthmic spondylolisthesis treated operatively or conservatively: a long-term follow-up. *Eur Spine J* 1997;6:393-7.
145. Shah RR, Mohammed S, Saifuddin A, et al. Comparison of plain radiographs with CT scan to evaluate interbody fusion following the use of titanium interbody cages and transpedicular instrumentation. *Eur Spine J* 2003;12:378-85.
146. Shah RR, Mohammed S, Saifuddin A, et al. Radiologic evaluation of adjacent superior segment facet joint violation following transpedicular instrumentation of the lumbar spine. *Spine* 2003;28:272-5.
147. Shen FH, Samartzis D, Andersson GB. Nonsurgical management of acute and chronic low back pain. *J Am Acad Orthop Surg* 2006;14:477-87.
148. Sinikallio S, Aalto T, Airaksinen O, et al. Depression is associated with poorer outcome of lumbar spinal stenosis surgery. *Eur Spine J* 2007;16:905-12.
149. Sivik TM. Personality traits in patients with acute low-back pain. A comparison with chronic low-back pain patients. *Psychother Psychosom* 1991;56:135-40.
150. Slover J, Abdu WA, Hanscom B, et al. The impact of comorbidities on the change in short-form 36 and oswestry scores following lumbar spine surgery. *Spine* 2006;31:1974-80.
151. Songer M. Repair of the pars interarticularis defect with a cable-screw construct. *Spine* 1998 Jan 15;23(2):284.
152. Sorensen LV. Preoperative psychological testing with the MMPI at first operation for prolapsed lumbar disc. Five-year follow up. *Dan Med Bull* 1992;39:186-90.
153. Spangfort E. Disc surgery. *Textbook of Pain* 1984:pp601-7.
154. Spengler DM, Ouellette EA, Battie M, et al. Elective discectomy for herniation of a lumbar disc. Additional experience with an objective method. *J Bone Joint Surg Am* 1990;72:230-7.
155. Steiner ME, Micheli LJ. Treatment of symptomatic spondylolysis and spondylolisthesis with the modified Boston brace. *Spine* 1985;10:937-43.
156. Stromqvist B, Fritzell P, Hagg O, et al. One-year report from the Swedish National Spine Register. Swedish Society of Spinal Surgeons. *Acta Orthop Suppl* 2005;76:1-24.
157. Sudo H, Oda I, Abumi K, et al. Biomechanical study on the effect of five different lumbar reconstruction techniques on adjacent-level intradiscal pressure and lamina strain. *J Neurosurg Spine* 2006;5:150-5.
158. Szypryt EP, Twining P, Mulholland RC, et al. The prevalence of disc degeneration associated with neural arch defects of the lumbar spine assessed by magnetic resonance imaging. *Spine* 1989;14:977-81.
159. Tallarico RA, Madom IA, Palumbo MA. Spondylolysis and spondylolisthesis in the athlete. *Sports Med Arthrosc* 2008;16:32-8.
160. Taylor VM, Deyo RA, Cherkin DC, et al. Low back pain hospitalization. Recent United States trends and regional variations. *Spine* 1994;19:1207-12; discussion on p.13.
161. Throckmorton TW, Hilibrand AS, Mencio GA, et al. The impact of adjacent level disc degeneration on health status outcomes following lumbar fusion. *Spine* 2003;28:2546-50.
162. Toomey TC, Gover VF, Jones BN. Spatial distribution of pain: a descriptive characteristic of chronic pain. *Pain* 1983;17:289-300.

163. Transfeldt EE, Dendrinos GK, Bradford DS. Paresis of proximal lumbar roots after reduction of L5-S1 spondylolisthesis. *Spine* 1989;14:884-7.
164. Transfeldt EE, Mehbod AA. Evidence-based medicine analysis of isthmic spondylolisthesis treatment including reduction versus fusion in situ for high-grade slips. *Spine* 2007;32:S126-9.
165. Trief PM, Grant W, Fredrickson B. A prospective study of psychological predictors of lumbar surgery outcome. *Spine* 2000;25:2616-21.
166. Trief PM, Ploutz-Snyder R, Fredrickson BE. Emotional health predicts pain and function after fusion: a prospective multicenter study. *Spine* 2006;31:823-30.
167. Tubach F, Beaute J, Leclerc A. Natural history and prognostic indicators of sciatica. *J Clin Epidemiol* 2004;57:174-9.
168. Uden A, Astrom M, Bergenudd H. Pain drawings in chronic back pain. *Spine* 1988;13:389-92.
169. Umehara S, Zindrick MR, Patwardhan AG, et al. The biomechanical effect of postoperative hypolordosis in instrumented lumbar fusion on instrumented and adjacent spinal segments. *Spine* 2000;25:1617-24.
170. Waddell G, Richardson J. Observation of overt pain behaviour by physicians during routine clinical examination of patients with low back pain. *J Psychosom Res* 1992;36:77-87.
171. van Tulder M, Furlan A, Bombardier C, et al. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003;28:1290-9.
172. van Tulder MW, Assendelft WJ, Koes BW, et al. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine* 1997;22:427-34.
173. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. *Pain* 1995;62:233-40.
174. Ware JE, Jr. SF-36 health survey update. *Spine* 2000;25:3130-9.
175. Weinhoffer SL, Guyer RD, Herbert M, et al. Intradiscal pressure measurements above an instrumented fusion. A cadaveric study. *Spine* 1995;20:526-31.
176. Weisskopf M, Herlein S, Birnbaum K, et al. [Comparative analysis of lumbar spine degeneration documented by x-rays versus large specimen cryomicrotome sections]. [German] *Z Orthop Ihre Grenzgeb* 2003;141:86-91.
177. Videbaek TS, Christensen FB, Soegaard R, et al. Circumferential fusion improves outcome in comparison with instrumented posterolateral fusion: long-term results of a randomized clinical trial. *Spine* 2006;31:2875-80.
178. Wilke HJ, Rohlmann F, Neidlinger-Wilke C, et al. Validity and interobserver agreement of a new radiographic grading system for intervertebral disc degeneration: Part I. Lumbar spine. *Eur Spine J* 2006;15:720-30.
179. Wiltse LL, Newman PH, Macnab I. Classification of spondylolisis and spondylolisthesis. *Clin Orthop Relat Res* 1976:23-9.
180. Wiltse LL, Radecki SE, Biel HM, et al. Comparative study of the incidence and severity of degenerative change in the transition zones after instrumented versus noninstrumented fusions of the lumbar spine. *J Spinal Disord* 1999 Feb;12(1):27-33.
181. Wiltse LL, Rocchio PD. Preoperative psychological tests as predictors of success of chemonucleolysis in the treatment of the low-back syndrome. *J Bone Joint Surg Am* 1975;57:478-83.
182. Wiltse LL, Rothman LG. Spondylolisthesis: classification,diagnosis, and natural history. *Semin Spine Surg* 1989;1:78-94.
183. Wiltse LL, Widell EH, Jr., Jackson DW. Fatigue fracture: the basic lesion is in isthmic spondylolisthesis. *J Bone Joint Surg Am* 1975;57:17-22.
184. Wiltse LL, Winter RB. Terminology and measurement of spondylolisthesis. *J Bone Joint Surg Am* 1983;65:768-72.

185. Virta L, Ronnemaa T, Osterman K, et al. Prevalence of isthmic lumbar spondylolisthesis in middle-aged subjects from eastern and western Finland. *J Clin Epidemiol* 1992;45:917-22.
186. Voorhies RM, Jiang X, Thomas N. Predicting outcome in the surgical treatment of lumbar radiculopathy using the Pain Drawing Score, McGill Short Form Pain Questionnaire, and risk factors including psychosocial issues and axial joint pain. *Spine J* 2007;7:516-24.
187. Wynne-Davies R, Scott JH. Inheritance and spondylolisthesis: a radiographic family survey. *J Bone Joint Surg Br* 1979;61B:301-5.
188. Zanoli G, Stromqvist B, Padua R, et al. Lessons learned searching for a HRQoL instrument to assess the results of treatment in persons with lumbar disorders. *Spine* 2000;25:3178-85.
189. Zhao K, Yang C, Zhao C, et al. Assessment of non-invasive intervertebral motion measurements in the lumbar spine. *J Biomech* 2005;38:1943-6.

5 PAPERS 1 – 4

5.1 ERRATA CONCERNING PAPERS 1 AND 2

- Table 4 in Paper 1: “2 years” following the symbol †should be changed to “pretreatment” and the symbol * should be placed after the ODI value 0.223, instead of after the At-work p-value for the surgery group.

- Table 3 in Paper 2: “long term” should be changed to “2-year” and p=0.3 to p=0.18

- Page header on uneven pages in Paper 2: Erkman should be changed to Ekman.