

# Sciatica and Disc Herniation: Outcome Measures and Prognostic Factors

Anne Julsrud Haugen



Faculty of Medicine, University of Oslo



Østfold Hospital Trust  
Department of Rheumatology  
January 2012

© Anne Julsrud Haugen, 2012

*Series of dissertations submitted to the  
Faculty of Medicine, University of Oslo  
No. 1334*

ISBN 978-82-8264-450-1

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.  
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.  
The thesis is produced by Unipub merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.



## Table of contents

Preface.....	4
Acknowledgements .....	5
Funding.....	7
List of papers.....	8
Abbreviations.....	9
1 Introduction .....	10
1.1 Historical and current perspective on sciatica.....	10
1.1.1 Biopsychosocial model in back pain and sciatica.....	11
1.1.2 Pain and fear of movement (kinesiophobia) .....	12
1.2 The normal disc and potential pathophysiology mechanisms in sciatica .....	13
1.3 Disease characteristics and diagnosis of sciatica caused by disc herniation .....	14
1.4 Epidemiology .....	15
1.5 Outcome measures .....	16
1.6 Treatment and clinical course of sciatica.....	17
1.7 Prognostic factors for poor outcome .....	27
1.8 Aims.....	29
2 Methods.....	30
2.1 Design.....	30
2.2 Study population .....	30
2.2.1 Sciatica cohort.....	30
2.2.2 Inclusion and exclusion criteria .....	30
2.3 Data collection .....	30
2.3.1 Procedures.....	30
2.3.2 Test–retest sample .....	31
2.4 Treatment.....	31
2.5 Measurements .....	31
2.5.1 Patient-reported outcomes .....	33
2.6 Statistics .....	34
2.6.1 Sample size calculation .....	34
2.6.2 Statistical methods used.....	35
2.6.3 Presentation of data .....	35
2.6.4 Missing items.....	35
2.6.5 Bivariate examinations .....	35
2.6.6 Principal component analysis .....	36
2.6.7 Floor or ceiling effects .....	36
2.6.8 Reliability .....	36
2.6.9 Internal consistency .....	37
2.6.10 Construct validity.....	37
2.6.11 Responsiveness.....	38
2.6.12 Logistic regression .....	38
2.7 Ethical issues.....	39
3 Summary of the results .....	40
3.1 Paper I.....	40
3.2 Paper II.....	41
3.3 Paper III .....	42
4 General discussion.....	44

4.1	Methodological considerations .....	44
4.1.1	Study design.....	44
4.1.2	Patient selection criteria .....	46
4.1.3	Selection of prognostic and outcome variables.....	47
4.1.4	Missing data on items .....	49
4.1.5	Validation of patient-reported outcomes.....	49
4.1.6	Estimating success or not.....	51
4.1.7	Statistical methods for prognostic factors.....	52
4.2	Discussion of the results in comparison with current evidence.....	53
4.2.1	Paper I.....	53
4.2.2	Paper II.....	55
4.2.3	Paper III .....	57
5	Clinical implications and further research .....	60
6	Conclusions.....	61
	References .....	62
	Papers.....	75

## Preface

In Norway, it is uncommon for back clinics to belong to a rheumatologic department; however, Østfold is an exception. The back clinic was included in our department in 1999 when Aage Indahl moved from Østfold after having finishing his doctoral thesis with the “Rygg i Revers” project at Sentralsykehuset i Østfold, Fredrikstad. The back clinic provides investigation and treatment for approximately 1,700 outpatients and 250 inpatients per year, with many of them suffering from sciatica. As clinicians, we realized that the literature referring to sciatica and disc herniation was limited, especially on prognosis and the influence of psychological and social factors. Each week we held meetings with the back surgeons in the hospital regarding patients who may be suitable candidates for surgical treatment of disc herniation. We understood the uncertainty about the indications for surgery, as Weber’s randomized controlled study from 1983 was referred to often.<sup>1</sup>

My colleague Lars Grøvle and I have worked on this project together since its beginning with the goal of writing two doctoral theses on the same material. Due to the interrelationship of our work, we suggest those interested in this topic to read both this and Lars’ dissertation, “Sciatica and Disc Herniation. The complexity of self-reported symptoms, health complaints and return to work”.

## Acknowledgements

First, I would like to thank my colleague Lars Grøvlø. This project would not have started without his enthusiasm, interest in research, and conviction that clinicians have to rely on the scientific literature. I am also incredibly grateful to have developed a very good friendship with Lars and for his close cooperation over many years. I appreciate his wisdom, patience, and encouragement to continue working. Recalling our countless talks and discussions about this project, I know this process has been educational and has given me insight into research.

I would like to express my sincere gratitude to my principal supervisor Margreth Grotle. I came to know her because part of the material in her doctoral thesis was from our back clinic in Østfold. I appreciate her proximity and willingness to answer my questions and for challenging me to improve my efforts.

Margreth and my cosupervisors Jens Ivar Brox and Bård Natvig, have given me supervision of very high quality, with thorough feedback, good comments, and advice that led to inspiring discussions. Thank you for sharing your knowledge with me.

I also want to thank Anne Keller for her contribution to the work and for her quick replies and interesting discussions, and Dag Soldal for contributing to the experience of working in a back clinic in a rheumatologic department.

My greatest gratitude goes to all those who contributed by collecting patient data for the study: all my rheumatologist colleagues and the physiotherapists Knut Morten Huneide and Anett Bjørnødegård at Sykehuset Østfold; Dag Soldal and Bjarte Justnæs at Sørlandet Sykehus; Anne Keller at Oslo Universitetssykehus Ullevål; and Eli Molde Hagen and Arne Skoglund at Sykehuset Innlandet. Without your efforts, it would not have been possible to include all patients. Additionally, our study nurse Eli Minge did a great job distributing the questionnaires to the right patient at the right time and following up the patients who did not respond.

My thanks also go to Bjørn Finnanger who was head of the Rheumatology Department at Sykehuset Østfold when we started the project and who always had trust in us.

Many thanks go to Holger Ursin and Hege R. Eriksen at Nasjonalt Ryggnettverk, Forskningsenheten, which was the network for back pain research at the University of Bergen when we started planning the project. Lars and I were invited to the Geilo

meetings and were introduced to the research environment on spinal diseases in Norway, which was very inspiring. There I also met Camilla Ihlebæk, who contributed to the discussion on comorbid subjective health complaints in Paper III.

My thanks go to Leiv Sandvik for his statistical support; often I looked forward to the first Friday of every month, the day Prof Sandvik came to give statistical advice to the researchers at Sykehuset Østfold.

I would also like to thank the Department of Research, Sykehuset Østfold. Special thanks to Famara Sanyang, Marianne Eckhoff, and Morten Jacobsen for continuous support and to the staff at the medical library for very good service.

During the time I worked on this project, I also worked in the clinic as a rheumatologist. At times it was frustrating when I had to work full-time in the clinic for several months and only had the opportunity to work on this project in my leisure time. However, for a great part of the last 1.5 year, 60% of my employment was dedicated to finishing the project. I owe my gratitude to all my colleagues at the rheumatology department for their understanding that it is not possible to be 100% involved in the clinic and do research at the same time. Special thanks go to Jonas Berglund and Grete Jespersen, leaders of the Rheumatology Department, Sykehuset Østfold, who understood the need for time to finish this research project.

Finally, I extend my warmest thanks to my family: my children Hanna, Hallvard, and Astrid for patience with their busy mother, and particularly my husband Eirik, who has supported me all these years. His knowledge about computers, the Internet, formatting, and cooking has been a great help to me.



## **Funding**

The South Eastern Regional Health Authority of Norway financed this PhD work.

## List of papers

- I. Haugen A.J., Grøvle L., Keller A., Grotle M.,  
Cross-Cultural Adaptation and Validation of the Norwegian Version of the Tampa  
Scale for Kinesiophobia  
*Spine* 2008;33:(17):E595–E601
- II. Haugen A.J., Grøvle L., Brox J.I., Natvig B., Keller A., Soldal D., Grotle M.  
Estimates of success in patients with sciatica due to lumbar disc herniation  
depend upon outcome measure  
*Eur Spine J.* (2011) 20:1669-1675
- III. Haugen A.J., Brox J.I., Grøvle L., Natvig B., Keller A., Soldal D., Grotle M.  
Prognostic factors for non-success in patients with sciatica and disc herniation.  
Submitted

## Abbreviations

ADL	Activity of Daily Living
AUC	Area Under the ROC Curve
BMI	Body Mass Index
CES-D	Center for Epidemiologic Studies Depression scale
CI	Confidence Interval
COS	Clinical Overall Score
CT	Computed Tomography
ECLT	Euglobulin Clot Lysis Time
EMG	Electromyography
EQ-5D	EuroQol-5D
HAD-A	Hospital Anxiety and Depression scale
HQM	Hannover Mobility Questionnaire
LBOS	Low Back Outcome Score
MCIC	Minimal Clinically Important Change
MCID	Minimum Clinically Important Difference
MIC	Minimal Important Change
MMPI	Minnesota Multiphasic Personality Inventory
MRI	Magnetic Resonance Imaging
MSPQ	Modified Somatic Perception Questionnaire
ODI	Oswestry Disability Index
OR	Odds Ratio
PAI-1	Plasminogen Activator Inhibitor 1
PWC	Pain and Working Capacity
RCT	Randomized Controlled Study
RMDQ	Roland Morris Disability Questionnaire
ROC	Receiver Operating Characteristic
SBI	Sciatica Bothersomeness Index
SC	Stauffer-Coventry scale
SCL	The Symptom Checklist-90
SD	Standard Deviation
SF-36	Medical Outcomes Study Short-Form 36
SFI	Sciatica Frequency Index
SLR	Straight-leg-raising test
SPORT	Spine Patient Outcomes Research Trial
VAS	Visual Analogue Scale

# 1 Introduction

## 1.1 Historical and current perspective on sciatica

Sciatica was described by early physicians and scientists in Greece<sup>2</sup> through the Greek word “iskhion” and later the Latin word “ischiadicus”. These words meant hip pain, but it was not until the early 1900s that the symptom was thought to emanate from spinal structures through the description of the pathology of the lumbar disc.<sup>3</sup>

In 1929, Dandy<sup>279</sup> operated on two cases with cauda equina syndrome removing “a completely detached fragment of cartilage from an intervertebral (lumbar) disc bulging dorsally into the spinal canal”. Both patients recovered, and Dandy concluded, “these cases offer pathologic evidence of a definite lesion including symptoms of so-called sciatica, with its all too meager pathology”. Five years later, Mixter and Barr<sup>5</sup> demonstrated that surgical removal of herniated disc material encroaching on a nerve root led to relief of radicular pain. Over the next few decades, the discectomy increased in popularity<sup>6,7</sup> and is still the commonly performed procedure.

Since its discovery in the early 1930s, the classic meaning of sciatica has been pain in the distribution of the sciatic nerve, and a herniated intervertebral disc is believed to cause most cases, around 90%. Other lesions affecting the integrity of the lumbosacral nerve roots (L4–S3) or the sciatic nerve, including lumbar canal or foraminal stenosis, tumors, cysts, hemorrhages, abscesses, fractures, and some more uncommon conditions, may produce the same clinical picture. In addition to back and leg pain, muscular weakness and sensory disturbances may occur. Affection of autonomic nerve fibers may cause bladder, bowel, and genital dysfunction. The condition varies from short-lasting, single episodes to a remitting or permanent course over months or years.

In the medical and scientific literature, the term sciatica is defined in many ways and is interchangeable with other terms, such as lumbosacral radicular syndrome, sciatic neuralgia, radiculopathy, and nerve root entrapment. In this thesis, I have chosen to use the term sciatica and the definition of sciatica as nerve root pain or radicular pain in the leg below the knee, due to its widespread use in the literature.<sup>8</sup>

Allan and Waddell<sup>3</sup> reviewed the history of low back pain and disability over the last 3,500 years. Clearly, the current epidemic of chronic disability due to low back pain, both nonspecific low back pain and more specified conditions, such as sciatica, which is

seen in all Western countries, is worrying from both an individual and a social perspective. Allan and Waddell suggested that low back pain has not changed throughout recorded human history, but how low back pain is understood and managed has changed. The fact that there is no evidence for a change in physical or neurophysiological pathology in patients with low back pain suggests that cultural and social factors strongly influence the consequences of experiencing a back pain episode. A broader understanding of pain in general, and low back pain in particular, therefore is suggested.<sup>9-12</sup> This is in line with a modern definition of pain, which states that “pain is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”.<sup>13, 14</sup> Therefore, in this thesis, a broader perspective on sciatica, including both psychological and social factors, is emphasized in addition to the physical and clinical factors.

### **1.1.1 Biopsychosocial model in back pain and sciatica**

The biomedical theory of health has been the dominant paradigm of the 19<sup>th</sup> and 20<sup>th</sup> century medicine. In this biomedical model, the human being is mainly considered as an organism or a sophisticated machine, and to a lesser degree, as a whole person.<sup>15</sup> The model pays little attention to the environment in which an individual person lives his or her life. As a reaction to the limitations of the biomedical model, Engel,<sup>16</sup> a professor in psychiatry and medicine, introduced the biopsychosocial model in 1977. The model proposes that the interrelationships of biological, psychological, and social factors play a significant role in the explanation of a disease. The biopsychosocial model became an important model in the back pain research field, in particular through Gordon Waddell’s<sup>17, 18</sup> and Alf Nachemson’s<sup>19</sup> significant works. They both emphasized the influence of cultural, social, and psychological environment on the pain and function of an individual person. They suggested that sole emphasis on the presumed structural features of the symptoms of spinal pain created a model that missed the important psychosocial part of the personal experience of pain.

A range of studies within low back pain research supports the notion that various psychosocial factors are associated with negative outcome.<sup>20-25</sup> Cognitive and behavioral factors contribute to disease continuation and disability in low back pain. Affective factors, particularly fear, have also been demonstrated to be important in persisting or worsening symptoms. According to recent reviews of studies on sciatica, psychosocial

factors, in addition to certain physical factors, have been shown to play an important role in the clinical course of sciatica.<sup>26, 27</sup> Hence, it is important to use the biopsychosocial model in the study of sciatica, similar to other back-pain conditions.

### **1.1.2 Pain and fear of movement (kinesiophobia)**

Fear is a natural response to a pain experience.<sup>28</sup> Based on scientific findings of the relationship between psychological factors and disability due to low back pain, Waddell stated that “fear of pain and what we do about pain may be more disabling than pain itself”.<sup>29</sup> He further pointed out there is little evidence that activity is harmful in low back pain, and that pain and disability are not directly related to each other.<sup>18</sup> Although Waddell introduced the fear-avoidance model into the back pain research field, the psychologist Lethem first described it in 1983.<sup>30</sup> The model proposes that confrontation leads to a reduction of fear over time, in contrast to avoidance that leads to more suffering and the perception of higher pain. Several studies have supported this model, both with respect to the fear of work-related activities<sup>29</sup> and the fear of movement following an injury.<sup>31</sup>

The term kinesiophobia was introduced by Kori<sup>31</sup> in 1990 and refers to “an irrational and debilitating fear of physical movement resulting from a feeling of vulnerability to painful injury or reinjury”. One of the authors, Miller, developed the Tampa Scale for Kinesiophobia.<sup>32</sup> The questionnaire was translated to Dutch and elaborated on by Vlaeyen, who named the phenomenon fear of movement/(re)injury and suggested a cognitive-behavioral treatment model.<sup>33</sup> According to the model, the patient reacts either with catastrophizing or with non-catastrophizing after an injury or painful experience. Patients with a non-catastrophizing attitude confront the pain symptoms, and this is associated with recovery. Contrarily, the catastrophizing patients develop fear of movement and avoidance, which might enhance the painful experience of the injury and disability. The model is used to discuss the role of pain-related fear in the explanation of chronic musculoskeletal pain.<sup>11, 34</sup>

In patients undergoing lumbar disc surgery, fear of movement/(re)injury predicted more disability and severe pain at 6 weeks, and more severe pain at 6 months.<sup>35</sup> In another study and contrary to what was anticipated, fear of movement and pain-catastrophizing were not associated with residual complaints after lumbar disc surgery at 3 and 12 months follow-up.<sup>36</sup> In two Swedish cross-sectional studies,

kinesiophobia was associated with disc diseases: one study was on patients with specific back pain (defined as disc herniation, isthmic spondylolisthesis, or spinal stenosis),<sup>37</sup> the other on patients treated with surgery for disc herniation.<sup>38</sup> According to the model, kinesiophobia may play an apparent role in the transition from acute to chronic pain. Therefore, in sciatica, a possible association might be between kinesiophobia and the development of chronic pain, but the results for patients operated on for sciatica are conflicting.

## **1.2 The normal disc and potential pathophysiology mechanisms in sciatica**

The intervertebral discs are pads of fibrocartilage that link the vertebral bodies together. Their main function is mechanical, transmitting loads arising from body weight and muscle activity through the spinal column and allowing bending, flexion, and torsion. The disc is composed of a central core, the nucleus pulposus, which is surrounded by the annulus fibrosus, a thick outer ring of fibrous cartilage. The nucleus pulposus consists of a matrix of proteoglycan and water gel held together by a network of collagen type II and elastin fibers, maintained by fibroblast-like and chondrocyte-like cells. The proteoglycan macromolecules are hydrophilic; the nucleus contains 70% water, generating large hydrostatic pressures within the nucleus and the inner annulus fibrosus. The annulus is a series of 15 to 25 concentric rings, or lamellae, cross-linked by radiating elastin fibers. Inferiorly and superiorly, cartilaginous endplates attach the disc to the adjacent vertebral bodies. A young healthy disc behaves like a waterbed, with the high water content of the nucleus and inner annulus enabling the tissue to act like a fluid.<sup>39</sup>

Adult discs are mostly avascular, thus the disc obtains nutrition mostly by diffusion, although blood vessels supply the surface of the annulus and the cartilaginous endplates.<sup>40, 41</sup> The discs are innervated by nerves located in the outer lamellae of the annulus, whereas the endplates are aneural. The posterior longitudinal ligament, which passes directly posterior to the disc, is richly innervated by nociceptive fibers.

From the early teenage years, a process of proteoglycan loss and consequently lower disc water content is seen, especially in the nucleus, continuing with increasing age.<sup>42</sup> The annular lamellae become irregular and bifurcating, and the collagen and elastin networks become more disorganized.<sup>43</sup> Gradually, degenerated discs behave less

hydrostatically under load, and the annulus becomes stiffer and weaker.<sup>44</sup> These changes are followed by the appearance of nuclear clefts and annular tears<sup>45</sup> which may permit nucleus tissue to be displaced into the annulus, forming herniations. Disc herniations may range from protrusions (outer annular lamellae remain intact) to extrusions (annular lamellae are ruptured) to sequestrations (herniation is completely detached from the body of the disc).<sup>46</sup> Studies in twins have shown a substantial genetic predisposition to disc degeneration.<sup>47</sup>

Within the cauda equina the nerves run downward and laterally before exiting their respective foramina. At their takeoff from the dural sac, the sciatica nerve roots are fastened by ligamentous attachments to the vertebral body and the subjacent pedicle within the foramen. Thus, a disc herniation may cause stretching and compression of the nerve root and dorsal root ganglion. A posterior disc herniation usually affects the root of the nerve exiting at the level below the herniation; a herniation between the vertebrae of L5 and S1 will usually affect the S1 nerve root. Herniations extending far laterally may affect the root at the same level. Very large herniations may compress more than one nerve.

It has been shown that stimulation of compressed roots causes pain whereas manipulation of normal roots does not.<sup>48</sup> Rydevik<sup>49</sup> and Olmarker<sup>50</sup> reported that compression was associated with edema formation and decreased electrical impulse propagation in the nerve root. They also showed that application of the tissue from the nucleus pulposus on the root induced inflammatory reactions.<sup>51</sup> Histological evaluation of herniated disc tissue has revealed prominent infiltration of inflammatory cells, most markedly macrophages and cytokines that promote lymphocyte activation that further recruit and activate macrophages toward phagocytosis and proteolytic enzyme secretion.<sup>52</sup> The combination of compression and inflammation is now widely accepted as an important pathophysiologic factor.<sup>53,54</sup> Longstanding root compression may result in axon loss and intra- and extraneural fibrosis. All types of fibers in the nerve roots may be affected.

### **1.3 Disease characteristics and diagnosis of sciatica caused by disc herniation**

Although patients with sciatica present symptoms that can be recognized and diagnosed as sciatica relatively easily, it is crucial to distinguish this condition from two



of the other major back problems: simple backache (or nonspecific low back pain) and back pain due to possible serious spinal pathology. This diagnostic triage, recommended by most clinical guidelines for treatment of low back pain,<sup>55</sup> is based mainly on clinical history, a physical examination, and if necessary, imaging.

The cardinal symptoms of sciatica include radiating pain with or without sensory disturbances or weakness. Pain is described typically as sharp, lancinating, or burning and is often exacerbated by coughing and sneezing. Clinical signs of nerve dysfunction support the diagnosis, including an abnormal straight-leg raising, or Lasègue's test,<sup>56</sup> reduced dermatomal sensibility, muscular strength, or tendon reflexes. Examination of a patient suspicious of cauda equina syndrome includes testing bladder and anal function. A recent Cochrane review indicated poor diagnostic performance to identify lumbar disc herniation when physical tests were used in isolation.<sup>57</sup> A diagnosis of sciatica due to disc herniation therefore requires identification by magnetic resonance imaging (MRI) or computed tomography (CT) with a site and level corresponding to symptoms and clinical findings.

CT and MRI show equal ability to identify disc herniations<sup>58,59</sup> and may be used to classify herniations according to morphology, volume, or location in the sagittal or horizontal plane.<sup>46</sup> There are weak associations between self-reported symptoms and the size of the herniation or whether it is a protrusion, extrusion, or sequestration.<sup>60,61</sup> Clinically silent disc herniations are common. MRI findings have shown prevalences between 20 and 36% among asymptomatic individuals.<sup>62,63</sup> Electrophysiological tests<sup>64</sup> do not provide diagnostic information beyond history, imaging, and clinical examination. Hence, there is no strong association between findings on imaging and the clinical presentation of sciatica. The diagnostic performance of MRI and CT on disc herniation and disc pathology are presented in two systematic reviews that propose that MRI or CT may incorrectly classify a considerable proportion of patients for disc herniation.<sup>65,66</sup> From a clinical perspective, it is important to be aware of the poor precision of imaging as a diagnostic screening tool.

## **1.4 Epidemiology**

As epidemiological studies of sciatica from the general population are based mostly on physicians' diagnoses or self-reported questionnaires without any radiological findings, there is limited knowledge on the incidence or prevalence of

sciatica caused by disc herniation. In a Finnish general population study, the point prevalence was estimated to be 4.8% based on symptoms and clinical examination.<sup>67</sup> In another Finnish epidemiological study, the lifetime cumulative incidence was estimated to be 12.5% for males and 12.0% for females when applying the clinical diagnosis of sciatica made by a physician.<sup>68</sup> According to a recent review on the prevalence of sciatica, the wide spectrum of definitions of sciatica has resulted in a large variation in prevalence estimates, with values ranging from 1.2% to 43%.<sup>69</sup>

Bladder, bowel, and sexual dysfunction, constituting a cauda equina syndrome, is rarely reported.<sup>70</sup> The exact occurrence of this potentially devastating complication is not known but the one year incidence is reported to be from 1–3/100,000.<sup>71</sup>

## 1.5 Outcome measures

Since radiological or clinical findings are weakly associated with symptoms, disability, and quality of life in sciatica,<sup>72</sup> there is no recommended outcome measures that can be considered to represent objective measures of sciatica. Therefore, clinicians and researchers have to rely on patients' subjective indicators to evaluate the severity and prognosis of the disease. Validated outcome measures have been developed to assess the impact of symptoms and disability of back pain patients.<sup>73,74</sup> Compared with condition-specific questionnaires for evaluating low back pain,<sup>75,76</sup> few self-report symptom questionnaires exist for sciatica. Patient-reported outcomes used in clinical studies on sciatica are patient-rated global change;<sup>77-80</sup> leg pain;<sup>77,78,81-83</sup> back pain;<sup>83</sup> Medical Outcomes Study Short-Form 36 (SF-36) Bodily Pain and Physical Functioning;<sup>79</sup> North American Spine Society neurogenic symptoms score;<sup>84</sup> Oswestry Disability Index;<sup>79,81,82</sup> and the Roland Morris Disability Questionnaire.<sup>78,80,83,85</sup> Few studies have used sciatica-specific outcomes, such as the Sciatica Bothersomeness Index and the Sciatica Frequency Index, and then only as secondary outcomes.<sup>78-80</sup>

At the time of planning the current study there was no consensus as to which health-related quality of life measure should be recommended for patients with sciatica in clinical studies. Patients with sciatica differ from patients with low back pain not only because of the radicular symptoms, but also because they generally report more severe pain, have longer absences from work, and return to work less.<sup>86-89</sup> Additionally, the sensory and motor dysfunctions and complications after surgery should be addressed.

Another aspect is that no measures had been advocated for use among single patients in daily clinical use. Therapeutic decisions were taken and still are taken on a global subjective judgment by the clinician, with no support of standardized and validated patient-report measures.

Clinical relevance and feasibility are key elements for questionnaires intended to be used in daily clinical care. Such questionnaires must be short, easily scored, easily interpretable, and require little staff time.<sup>90</sup>

## **1.6 Treatment and clinical course of sciatica**

Sciatica is usually treated by conservative or surgical therapy. Due to early implementation of various conservative treatment modalities and in particular, surgical therapy, the natural course of sciatica and disc herniation is not fully known. There are large variations in both the referral practice to surgery and the various types of treatment provided across different countries.<sup>91</sup> Over the last decades, a discectomy has been an increasingly used surgical treatment in industrial countries. The mean annual rate of discectomy per 100,000 inhabitants is reported to about 24 in Sweden,<sup>92</sup> 60 in Norway,<sup>4</sup> 80 in Denmark,<sup>93</sup> and 100 in the United States.<sup>94</sup> This variation in the rates of surgery exceeds the expected variations of sciatica and indicates that nonmedical factors influence the referral to discectomy treatment.

In the clinical guidelines for diagnosis and treatment of sciatica<sup>8</sup> it is recognized that patients with persistent pain over 6–8 weeks due to disc herniation with nerve root affection should be assessed for surgery. However, there are no strict criteria or evidence-based appropriateness criteria for surgery in patients with sciatica except for cauda equina syndrome, and there are conflicting results with respect to the beneficial long-term effect of surgery compared with conservative treatment. Several studies have compared surgical and conservative treatment for patient with lumbar disc herniation. Few studies have used a randomized, controlled design<sup>1, 78, 82, 83, 95, 96</sup> but there have been many observational studies.

In 1983, Weber et al. from Norway carried out the first randomized, controlled trial in which surgical therapy was compared with conservative care.<sup>1</sup> Patients with definite indications for surgery and patients with moderate symptoms who showed continuous improvement were excluded from the study. There was a significant difference in favor of the surgical group at the one-year follow-up, with a good or fair

result in 92% compared with 82% in the conservative group. After four and ten years, no significant differences were found between the groups and about 90% of all patients had a good or fair result. More recent randomized trials support evidence that patients treated with surgery recover more rapidly than patients treated conservatively, but the effect of surgery seems to disappear after one year.<sup>78, 96</sup> Similar results have also been found in many of the observational studies with long-term follow-up. For example, the Maine study,<sup>97, 98</sup> which followed patients with sciatica prescribed by the physician to surgical or non-surgical treatment, found that surgically treated patients had a higher severity of symptoms at baseline but were better at the follow-up. The relative advantage of surgery was greatest early in the observation period and narrowed over five and ten years.<sup>80, 97, 98</sup> The authors concluded that discectomy could be avoided, particularly in patients with milder symptoms.<sup>80</sup>

The effect of various other surgical treatment methods for lumbar disc prolapse was reviewed in a Cochrane review in 1999, which identified 26 randomized controlled studies: 15 about chemonucleolysis using intradiscal injections of chymopapain, a proteolytic enzyme; ten about surgical techniques; and one unclassified.<sup>99</sup> Meta-analyses from this Cochrane review concluded that chemonucleolysis was better than placebo and less effective than discectomy. Later, concerns about chemonucleolysis' safety and effectiveness led to limited use of chemonucleolysis using chemopapain.<sup>100</sup> The results comparing the different surgical techniques were divergent, but there is evidence for surgical treatment of the disc herniation for faster relief for carefully selected patients.<sup>100, 101</sup> Some trials have shown techniques reducing scar formation, but there is limited evidence on the effect on clinical outcomes.<sup>100, 102, 103</sup>

No study so far has compared surgical treatment of disc herniation with placebo. In a double-blind design comparing surgical and non-surgical treatment, sham surgery should be included; however, ethical issues make this difficult to implement. In a review from 1972 on 2,504 operations for lumbar disc diseases, Sprangfort<sup>104</sup> showed that patients reported complete relief of sciatic pain in 37% of patients and complete relief of back pain in 43% where no disc herniation was discovered. Hence, the placebo effects influence patients' outcomes after any treatment, including surgery.<sup>105</sup>

We cannot tell patients the exact effect of surgical treatment for disc herniation without knowing the effect of a double-blind randomized controlled study with sham surgery. The placebo effect was illustrated in a randomized, controlled study on

osteoporotic fractures, where no beneficial effect was found for vertebroplasty compared with a sham procedure.<sup>106</sup>

Treatments directed to the specific inflammatory pathogenesis have also been studied. A Finnish randomized trial with an anti-tumor necrosis factor alpha did not show a favorable effect of infliximab at the one-year of follow-up.<sup>107</sup> In a recent placebo-controlled randomized trial with adalimumab on patients with sciatica and disc herniation, the three-year results showed a reduction in the need for back surgery in the adalimumab group, but no significantly different results in back or leg pain.<sup>108</sup>

Commonly used conservative treatment modalities for sciatica, such as traction, physical therapy, bed rest, manipulation, medication, and acupuncture, were included in a systematic review including 30 trials.<sup>109</sup> This review reported that none of the treatment modalities was shown to be more effective in the long-term. In addition, according to randomized, controlled trials, epidural injections of corticosteroids, which is another popular conservative treatment for sciatica, have shown conflicting results.<sup>110-112</sup> A recent placebo-controlled Norwegian study concluded that the efficacy of epidural corticosteroids was no better than epidural injections of saline or subcutaneous injections of local anesthetics.<sup>111</sup>

In summary, the results regarding the effect of conservative versus surgical treatment are conflicting, and there are still no evidence-based criteria for the selection of patients to surgical treatment, nor to various conservative treatment modalities for sciatica. Methodological variation and limitations across the studies make it difficult to interpret and compare the results. For example, different definitions of sciatica due to disc herniation are used, resulting in variations between patient samples at baseline. Different outcome measures and definitions of recovery and/or successful outcomes also vary between the studies, which limits a comparison of the results across studies. Another limitation is small sample sizes, making the results vulnerable for type II errors.

Despite these limitations studies suggest that a relatively large proportion of patients seem to have a poor prognosis of sciatica from a long-term perspective. Table 1 provides an overview of prospective studies with either no surgical treatment or mixed cohorts, and table 2 provides an overview of prospective studies with surgical treatment for sciatica and lumbar disc herniation. The proportion of patients with poor outcome after ten years varies from 10% in the first study from Weber<sup>1</sup> up to between 40 and 60% in the later Maine study.<sup>80, 97, 98</sup> In addition, the prospective studies, including only

surgically treated patients show that the proportion of patients with a poor long-term outcome is approximately 20-40%. These results indicate that the clinical course of sciatica in a long-term perspective is poor for a relatively large proportion of the patients, regardless of which treatment they receive. This also suggests that there is a strong need for more studies on this topic.

Table 1. Studies on clinical course and prognostic factors of sciatica for patients who received conservative treatment, or mixed surgical/conservative treatment, presented in chronological order.

Author / Year	Patients and settings	Design and treatment	Outcome measures	Follow-up	Main results	Significant baseline prognostic factors
Weber <sup>1</sup> 1983	126 patients with radicular pain, positive SLR test, and uncertain indication for surgery. Imaging: Radiculography Age 25-55 years (Norway)	RCT, conservative care or surgery  About ¼ of the patients in the conservative group were operated during the first year.	Good, fair, poor or bad	1, 4 and 10 years	Good or fair: 1 year: Surgical group 92%, conservative group 82%  4 and 10 years: 90% in both treatment arms	In univariate analyses younger age, shorter sick leave and higher physical activity predicted better outcome at 4 years. Younger age predicted better outcome at 10 years.
Weber <sup>63</sup> 1993	214 patients with less than 2 weeks of radiating pain and positive SLR test consulting primary care physicians. Imaging: No Age 18-75 years (Norway)	The first 2 weeks of the study: RCT, Piroxicam versus placebo. Thereafter an inception cohort. 2% of patients had surgery.	Back pain VAS, leg pain VAS, modified RMDQ (17 items), duration of sick leave, use of analgesics, restrictions at work and leisure.	4 weeks, 3 and 12 months	4 weeks: Improvement in pain and RMDQ scores, but Piroxicam was not more effective than placebo for pain relief. Mean duration sick leave: 28 days.  12 months : 30% reported restrictions at work and leisure, and used analgesics often. 8% were sick listed.	Univariate analyses: previous sciatica episodes predicted worse outcome.
Hasenbri ng <sup>113</sup> 1994	111 hospitalized patients with acute radicular pain and a disc prolaps or protrusion Imaging: Myelography and CT Age 17-72 years (Germany)	Prospective cohort. 1/3 of patients operated.	Pain intensity, surgeon's evaluation of pain, application for early retirement.	6 months	Results according to treatment not reported. 6 months :20% severe pain, 40% no or minimal pain. 20% had applied for early retirement.	Predictors worse outcome: more disc displacement, more scoliosis, lower pain coping ability and social status. Multivariate regression analysis. The analysis was adjusted for treatment which was not found to be a significant predictor.  Somatic parameters accounted for 12%, social variables for 6%, and psychological variables for 37% of the variance.

Author / Year	Patients and settings	Design and treatment	Outcome measures	Follow-up	Main results	Significant baseline prognostic factors
Nykvist <sup>77, 114, 115</sup> 1995	342 hospitalized patients, Imaging: Myelography Age <55 years (Finland)	Observational cohort. 2/3 of patients operated.	Poor outcome 5 years: moderate-extremely occupation handicap  Pain in leg or back, modified ADL score (0-20), Global change scale, Global satisfaction, Employment and pension status.	5 and 13 years	5 years: poor outcome surgical: 31% men and 54% women; non-surgical 59% men and 65% women. Of operated patients, 68% still had sciatica and 21% were retired. Of non-operated patients, 82% still had sciatica and 26% were retired.  13 years: surgical: patient's report of sciatica 67%, no change/worse 19%, dissatisfied 26%, retired 21%, ADL score 5.1. Non-surgical: patient's report of sciatica 68%, no change/worse 41%, dissatisfied 24%, retired 21%, ADL score 5.8.	5 years: clinical findings predicting poor outcome: for the operated patients: sensory deficits of legs, tenderness in lumbar extension, decreased repetitive trunk flexion capacity, decreased lumbar lordosis, tightness of hamstrings. For the non-operated patients, due to heterogeneity, no statistically acceptable regression model could be established.  Clinical, social, psychological factors predicting poor outcome for operated: subjective working incapacity, sensory deficit, tightness of hamstrings, age, pain in lumbar extension. For non-operated: increased occurrence of occupational hazards, co-morbidity.
Carragee <sup>116</sup> 1997	87 patients with sciatica referred to hospital for MRI. Radicular pain, positive SLR test or motor weakness, and abnormal MRI Imaging: MRI Age 20-70 years (California, USA)	Combined retrospective and prospective observational study.  Non-surgical cohort (surgical presented Table 2)	Composite measure (sciatic pain, medication, restriction of activity, satisfaction) <sup>117</sup> on a 0-10 scale >6= excellent/good ≤6 = fair/ poor	>2 years	56% excellent/good 44% fair/poor	Predictors for a good outcome: a shorter duration of current sciatica episode, no involvement with litigation, younger age, and a small ratio of the disc area to the remaining spinal canal assessed by MRI Multivariate regression analysis
Balague <sup>8</sup> 1999	82 hospitalized patients with sciatica lasting up to 2 months Imaging: CT or MRI Age <65 years (Switzerland)	Prospective observational cohort  1/3 of patients had surgery	Recovered if pain VAS≤15, ODI score ≤20 and normal muscle function. Non-recovered if otherwise.	1 year	37% of patients were recovered at 1 year, 12 % reported radiating pain.	Predictors non-recovery: abnormal neurological test OR 4.3 (95% CI: 1.37, 13.28). No statistically significant association between most variables for patients who opted for surgery and those who did not. Multivariate analyses
Vroomen <sup>118</sup> 2002	183 patients to general practitioner for the first time, sciatica with at least two defined typically signs and symptoms Imaging: no (Netherlands)	Observational cohort	Worsened, unchanged, improved or greatly improved	3 months	At 3 months 73% had improved (without surgery)	Predictors worse outcome: duration of pain >30 days and positive SLR Multivariate regression analysis
Beauvais <sup>119</sup> 2003	75 patients with sciatica or femoral neuralgia of < 1 month duration attended to rheumatology department Imaging: CT Age 18-70 years (France)		SC's treatment response (complete, partial, failure)  Definition failure: persistent pain+continuous analgesic use+ no return to work.	3 months	30% partially, 45% completely and 25% no recovery	Proportion of patients admitted to hospital for their pain was higher in the failure group ( $P = 0.01$ ). The prognostic value of CT could not be determined because the number of patients in these subgroups was small.

Author / Year	Patients and settings	Design and treatment	Outcome measures	Follow-up	Main results	Significant baseline prognostic factors
Atlas <sup>80, 97, 98, 120</sup> 2005 Edwards <sup>121</sup> 2007	507 patients recruited from orthopedic surgeons, neurosurgeons, and occupational medicine. Imaging: Not required (CT, MRI or myelography available in 55% of patients) Age >18 years (Maine, USA)	Observational cohort. 1/2 of patients operated.	Change in predominant symptom (leg or back pain). SBI, SFI, Modified RMDQ, SF-36, Work status, Patient satisfaction.	1, 5 and 10 years 3 years <sup>121</sup>	1 year: predominant symptom completely gone, much better or better: 80% (surgical) and 56% (non-surgical) Patients satisfied with current state: 60% (surgical) and 40% (non-surgical). 5 years: predominant pain symptom completely gone, much better or better: 70% (surgical) and 56% (non-surgical). Patients satisfied with current state: 63% (surgical) and 46% (non-surgical). 10 years: predominant pain symptom completely gone, much better or better: 69% (surgical) and 61% (non-surgical). Patients satisfied with current state: 71% (surgical) and 56% (non-surgical). Among those employed at baseline more than 80% in both treatment groups were working at 5 and 10 years.	In multivariate analyses surgical therapy predicted better outcome of predominant symptom at 1 year (OR 4.3 (95% CI, 2.5–7.3); 5 years (OR, 2.1; 95% CI, 1.3–3.6); but not at 10 years (OR, 1.4; 95% CI, 0.9–2.3) If better was restricted to completely gone or much better at 10 years: (OR, 2.1; 95% CI, 1.2–3.7). Patients treated surgically were more likely to report satisfaction with their current state at 5 years (OR 3.0; 95% CI, 1.8–5.2) and 10 years (OR, 2.2; 95% CI, 1.4–3.6). At 10 years, in different models, receiving Workers' Compensation at baseline was associated with worse outcomes (OR, 0.4–0.6). Psychological distress predicted greater pain and self perceived disability at 3 years. <sup>121</sup>
Osterma <sup>n92</sup> 2006	56 patients with radicular pain below knee 6-12 weeks and at least one specific physical finding Imaging: CT Age 20-50 years (Finland)	RCT comparing microdiscectomy and conservative management 10 patients in the control group crossed over to surgery	Pain in leg (VAS) Back pain (VAS) Work ability ODI The 15D instrument of health-related quality of life <sup>122</sup>	6 weeks, 3, 6, 12, 24 months	6 weeks: surgical treated patients less leg pain, more full recovery. Other results did not find clinically significant differences between the groups. 1 year: full recovery 33% (surgical group) and 30% (non-surgical group) 2 years: full recovery not assessed	Operative treatment resulted in a superior outcome at L4–L5, no difference was noted at L5–S1. Patients older than median (37 years) seemed to benefit more from surgery.
Jensen <sup>85</sup> 2007	154 non-surgical patients who had participated in a RCT of active conservative treatment. Imaging: MRI (no disc herniation in 10%). Age 18-65 years (Denmark)	Prospective observational cohort with focus on MRI findings	Recovery defined as absence of sciatic leg pain and RMDQ ≤3	1 year	53% recovered; 52% of the actively treated patients and 53% of patients in the control group.	Significant predictors of good outcome: Broad based disc protrusion OR 13.6 (95% CI; 1.9, 95.4), disc extrusion OR 10.6 (95% CI; 1.9, 58.7), male gender OR 2.6 (95% CI; 1.3, 5.0), absence of canal stenosis (males only) OR 4.2 (95% CI; 1.2, 14.7) (Paper difficult to interpret because of discrepancy between text and tables).
Peul <sup>78, 95, 123</sup> 2008	283 patients with symptoms of sciatica of 6-12 weeks duration. Indication for surgery confirmed by neurosurgeons. Imaging: MRI Age 18-65 years (Netherlands)	RCT, prolonged conservative care versus early surgery. 16/141 recovered before early surgery 55/142 assigned to prolonged conservative care, underwent surgery	RMDQ for sciatica, Leg pain (VAS), Global recovery 7 point scale, SF-36, SBI, SFI. Recovery: ≤2 on a Likert scale (1-7)	1 and 2 years	1 year recovery 95%, both groups 2 years recovery 80% both groups Early surgery achieved more rapid relief of sciatica than conservative care, but outcomes were similar by 1 and 2 years.	Predictors worse outcome: female sex, smoking and positive Bragards test (dorsiflexion of the foot aggravates radicular pain during SLR) Multivariate analyses.



Author / Year	Patients and settings	Design and treatment	Outcome measures	Follow-up	Main results	Significant baseline prognostic factors
Weinstei n <sup>96, 124</sup> 2006, 2008 Atlas <sup>125</sup> 2009 Rihn <sup>126</sup> 2011	1244 surgical candidates with symptoms of radiculopathy for at least 6 weeks enrolled; 501 in a RCT and 743 in an observational study. Imaging: MRI or CT Age >18 years (13 medical centers in 11 states in USA)	Prospective observational cohort: ¾ of patients operated. RCT (Surgery versus conservative care): of those randomized to surgery, 50% had surgery by 3 months, 57% by 1 year and 59% by 4 years Of those randomized to non-operative care, 30% of the patients received surgery by 3 months, 41% by 1 year, and 45% by 4 years	SF-36 bodily pain, SF-36 physical functioning, Modified ODI, Self-reported improvement, SBI, Satisfaction with current symptoms, Work status, Satisfaction with care.	6 weeks, 3 months, 6 months, 1, 2 and 4 years	Major improvement: RCT: 1 year 76% surgery, 67% non-surgery; 2 years 76% surgery, 69% non-surgery (intent to treat) Observational 1 year 80% surgery, 60% non-surgery, 2 years 76% surgery, 58% non-surgery  RCT: because large numbers of patients did not adhere to assigned treatment groups, conclusions based on intent-to-treat analysis were not warranted. In as-treated analyses, approximately 75% of operated and 50% of non-operated patients were somewhat or very satisfied with symptoms at 2 and 4 years.	In combined assessment of patients included in the randomized and the observational cohorts: Worse outcome in patients with symptom duration >6 months irrespective of treatment group, (univariate analyses). <sup>126</sup>  Baseline worker's compensation status not significantly associated with outcome in multivariate analyses. <sup>125</sup>
ADL	Activity Of Daily Living <sup>127</sup>					
COS	Clinical Overall Score <sup>128</sup>					
CI	Confidence Interval					
CT	Computed Tomography					
MRI	Magnetic Resonance Imaging					
ODI	Oswestry Disability Index <sup>129</sup>					
OR	Odds Ratio					
RCT	Randomized Controlled Trial					
RMDQ	Roland Morris Disability Questionnaire <sup>130</sup>					
SC	Stauffer-Coventry Scale <sup>131</sup>					
SBI	Sciatica Bothersomeness Index <sup>72</sup>					
SF-36	Medical Outcomes Study Short Form 36 <sup>132</sup>					
SFI	Sciatica Frequency Index <sup>72</sup>					
SLR	Straight-Leg Raising					
VAS	Visual Analogue Scale <sup>133</sup>					

Table 2. Prospective studies on clinical course and prognostic factors of sciatica for patients who received surgically treatment, presented in chronological order.

Author / Year	Patients and settings	Design and treatment	Outcomes measures	Follow-up	Main results	Significant baseline prognostic factors
Hurme <sup>134</sup> 1986	220 patients discectomy and partial laminectomy Imaging: EMG, rhizography Age <55 years (Finland)	Prospective observational study	Pain Index (0-30) ADL index (0-20) PWC	6 months	71% much better	Predictors -reduced PWC: short education, divorced or widowed, older age, physical hard work, somatization (SCL), long duration of sciatica, desire to retire -decrease of ADL: low preoperative ADL, desire to retire, short education, somatization (SCL), long duration of sciatica, older age, higher BMI, not self employed. Multivariate regression analysis
Sørensen <sup>135</sup> 1987	57 patients discectomy 16-62 years (Denmark)	Prospective observational study	Poor outcome defined as VAS >50, pain chart score ≥ 2 and poor health state	6 months	17.5% had poor outcome	Predictors for poor outcome: Female sex, action for damages, prolonged disease of the back, prolonged current attack, report of long-term illness, severe pain reported immediately post-operatively, pathological pain chart, increases on the MMPI clinical scales hypochondriasis, depression, hysteria, psychastenia, and the special scales panic fear, admission-of-symptom. <sup>136</sup>

Author / Year	Patients and settings	Design and treatment	Outcomes measures	Follow-up	Main results	Significant baseline prognostic factors
Haaland <sup>137</sup> 1992	122 patients discectomy incl laminectomy Imaging: CT or myelography in Age <70 years (Norway)	Prospective observational study	COS (0-1000): Pain in back and leg (VAS) Clinical and neurological examination Functional status Analgesics	1 year	Median score preoperative 524, and 1 year after surgery 104 (0-1000)	High levels of preoperative serum euglobulin clot lysis time (ECLT) and plasminogen activator inhibitor 1 (PAI-1) predicted the clinical outcome of surgery. This indicates poorer outcome in patients with impaired fibrinolytic activity. Multivariate regression analysis
Graver <sup>138-140</sup> 1995, 1998, 1999	134 patients discectomy incl laminectomy Imaging: CT or myelography in Age <70 years (Norway)	Prospective observational study	COS (0-1000): Pain in back and leg (VAS) Clinical and neurological examination Functional status Analgesics	1 and 7 years	1 year: mean score preoperative 520 and postoperative 157 (0-1000)  7 years: mean pain (VAS) low back 29 and leg 22. 88% fully satisfied with having undergone an operation	1 year. Predictors for poor outcome: Study A: high levels of anxiety and somatic distress (HAD-A scale, MSPQ) The psychological tests maintained when the effects of the fibrinolytic variables (ECLT and PAI-1) were controlled for. Explained variance physical and psychological variables 24% Study B: predictors poor outcome: low body height, high body weight, high BMI Multivariate regression analysis  7 years. Predictors for poor outcome: Female gender, high preoperative psychological distress (MSPQ) and impaired fibrinolytic activity. Multivariate regression analysis
Kjellby-Wendt <sup>141</sup> 1999	50 patients partial discectomy Imaging: CT or MRI Age: 21-68 years (Sweden)	Prospective observational study	Patients' satisfaction with the treatment; discontented: ≤ 3 (0-5)	2 years	37% were discontented	Predictors for poor outcome: high scores of preoperative pain, anxiety and depression. Bivariate analysis
Carragee <sup>116</sup> 1997	48 patients Imaging: MRI Age 20-70 years (California, USA)	Prospective observational study (non-surgically treated patients same cohort presented in Table 1)	Composite measure (sciatic pain, medication, restriction of activity, satisfaction) <sup>117</sup> on a 0-10 scale >6 = excellent/good ≤6 = fair/poor	>2 years	75% excellent/good 25% fair/poor	Predictors for good outcome: a larger anteroposterior disc length, larger ratios of disc area to canal area, Large right-left canal widths and small disc widths. Predictors for poor outcome: concurrent medical illness, workers' compensation involvement, and female gender, smaller (<6 mm) disc herniation. Multivariate regression analysis

Author / Year	Patients and settings	Design and treatment	Outcomes measures	Follow-up	Main results	Significant baseline prognostic factors
Woertge n <sup>142,143</sup> 1997 and 1999	121 patients discectomy Imaging: CT (87%), MRI (22%), or myelography (8%) Age 15-76 years (Germany)	Prospective observational study	LBOS Good 50-75, poor <50 Prolo scale <sup>144</sup> good 8-10, poor <8 Quality of life a little/no Pain grading scale Good 0-50, poor 51-100	1 year (univariate analysis) <sup>142</sup> 3,12, 28 months (multivariate analysis) <sup>143</sup>	Poor outcome (1 year) <sup>142</sup> : LBOS 24% Prolo scale 30% Pain grading scale 14% Quality of life 12 % Favorable outcome (LBOS) <sup>143</sup> : 3 months: 64% 12 months: 74% 28 months: 66%	Predictive factors for poor outcome, bivariate analysis <sup>142</sup> : -quality of life: medial prolapse, duration of pain and paresis -pain: smoking, duration of pain and paresis -Prolo: SLR <30°, walking <500 meters, smoking, duration of pain and paresis -LBOS: smoking, duration of pain and paresis Predictive factors for good outcome bivariate analysis <sup>142</sup> : - pain: sensory loss, walking <500 meters - quality of life: walking <500 meters LBOS: walking <500 meters  Predictive factors for poor outcome, multivariate analysis <sup>143</sup> : -3 months: education level, crossed SLR, score on the LBOS on admission -12 months: education level, SLR <30°, duration of paresis -28 months: radicular sensory deficit, duration of paresis, score on the LBOS on admission
Schade <sup>145</sup> 1999	46 patients discectomy Imaging: MRI Age 20-50 years (Switzerland)	Prospective observational study	Pain (VAS) RMDQ Return to work SC's treatment response with an additional excellent category	2 years	Leg pain: 83% complete relief SC: 74% excellent/good	Predictors of -pain relief: MRI-identified neural compromise and social support by the spouse -subjective disability: MRI-identified neural compromise and work-related resignation -return to work: depression and occupational mental stress -good result of surgical outcome: extent of herniation and depression Multiple regression, accounted for 30% (pain) and 58% (SC) of the variance.
Nygaard <sup>146</sup> 2000	132 patients microdiscectomy Imaging: CT or MRI Age <60 years (Norway)	Prospective observational study on surgically treated patients	COS (0-1000): Pain in back and leg (VAS) Clinical and neurological examination Functional status Analgesics	1 year	108 patients returned to work within the first year after surgery Value of COS not reported	Predictors poor outcome: long duration of sick leave and long duration of symptoms Accounted for 21% of the variance in multiple linear regression analysis.
Rothoerl <sup>147</sup> 2002	219 patients discectomy Imaging: Not specified Age 15-76 years (Germany)	Prospective observational study	Prolo scale favorable 8-10, poor 2-7 points	9.9 months	Prolo scale: 40% poor outcome	Predictive factors for poor outcome: longer duration of preoperative pain, smoking, paresis, and sensory deficit. Bivariate analysis
Asch <sup>148</sup> 2002	220 patients Microsurgery and i.v. Decadron or epidural steroids Imaging: CT/MRI Age: All ages (New York, USA)		VAS (0-10) in leg and back; success: ≤4 ADL (0-4) Time return to ADL Time return to work Satisfaction with surgery results ODI, success; <40%	1 and 10 days, 6 weeks, 6 and at least 12 months	Median follow up 2.07 years: Success leg pain: 80%, back pain 77%, ODI 78%, satisfaction and at least surgery 76%, return normal ADL 65%, return to work 61%.	Predictor for poor outcome: receiving workers' compensation status, higher age. Bivariate. Multivariate (ODI and leg pain VAS)

Author / Year	Patients and settings	Design and treatment	Outcomes measures	Follow-up	Main results	Significant baseline prognostic factors
Carragee <sup>149</sup> 2003	187 patients discectomy, some with a small laminotomy Imaging: MRI Age 18-65 (California, USA)	Combined retrospective and prospective observational study comparing intraoperative findings of disc herniations classified into four categories	ODI Stanford Score <sup>116</sup> (0-10) Persistent of sciatica Reoperation	>2 years	Mean 73% improvement	Poor clinical outcome in -ODI: no Fragment-Contained group (p<0.001), Fragment Defect group (p< 0.05) - Stanford Score: no Fragment-Contained group (p<0.001), Fragment fissure group (p< 0.05) - Persistent sciatica: no Fragment-Contained group (p<0.001) -reoperations: fragment-Defect group (21%) Multivariate regression analysis
Ng <sup>150</sup> 2004	113 patients discectomy Imaging: Radiological signs, not specified Age 12-66 years (United Kingdom)	Prospective observational study	ODI LBOS VAS The patients' subjective evaluation of the surgery (excellent, good, fair or poor).	1 year	77% had a significant reduction (20%) of the ODI 65% rated the operation as excellent or good.	Longer duration of sciatica correlated with less favorable outcome on ODI and LBOS and less degree of satisfaction with surgical outcome Patients with an uncontained herniated disc had a shorter duration of symptoms and a better functional outcome than those with a contained herniation. Multiple regression analysis
Kohlboeck <sup>151</sup> 2004	58 patients microdiscectomy Imaging: Not specified Age <70 years (Austria)	Prospective observational study	Pain (yes/no) Number of pain locations Pain intensity (0-10) HQM Return to work SF-36	6 months	56% benefited from lumbar discectomy At 6 months 25% reported no pain and had an average functional status of 85% in daily activities. 79% of this group had returned to work. Mean values SF-36 were above average except physical functioning	Predictors worse outcome: greater degree of preoperative depression (CES-D) postoperatively. The less the degree of the SLR, the greater the probability to belong to the success group. Multivariate analysis
Lønne <sup>152</sup> 2011	91 patients with limb paresis operated with microdiscectomy for lumbar disc herniation Mean age 45 years (Norway)	Prospective cohort study	Muscle strength of the affected limb (0-5) Recovery: 5 EQ-5D ODI Pain in back and leg (VAS) EQ-5D VAS general health state Employment state	1 year	Paresis: Full recovery: total 75%, with severe paresis 55%, with mild paresis 84% Improved 10% Unchanged 15% Patients who recovered had a significant better outcome in leg pain, ODI, EQ-5D and general health than patients with persistent paresis. Amongst those who did not recover, three times as many received workers compensation	Predictor not having full recovery of muscle function: a severe paresis at baseline, OR: 4.2 (95%CI = 1.6-11.4). Duration of paresis was not associated with recovery. Univariate analysis

Author / Year	Patients and settings	Design and treatment	Outcomes measures	Follow-up	Main results	Significant baseline prognostic factors
ADL	Activity Of Daily Living <sup>127</sup>					
BMI	Body Mass Index					
CES-D	The Center For Epidemiologic Studies Depression Scale <sup>153</sup>					
CI	Confidence Interval					
COS	Clinical Overall Score <sup>128</sup>					
CT	Computed Tomography					
ECLT	Euglobulin Clot Lysis Time					
EQ-5D	Euroqol-5D <sup>154</sup>					
HAD-A	Hospital Anxiety And Depression Scale <sup>155</sup>					
HQM	Hannover Mobility Questionnaire <sup>156</sup>					
LBOS	Low Back Outcome Score <sup>157</sup>					
MMPI	Minnesota Multiphasic Personality Inventory <sup>158</sup>					
MRI	Magnetic Resonance Imaging					
MSPQ	Modified Somatic Perception Questionnaire <sup>159</sup>					
ODI	Oswestry Disability Index <sup>129</sup>					
OR	Odds Ratio					
PAI-1	Plasminogen Activator Inhibitor 1					
PWC	Pain And Working Capacity <sup>160</sup>					
RMDQ	Roland Morris Disability Questionnaire <sup>130</sup>					
SC	Stauffer-Coventry Scale <sup>131</sup>					
SCL	The Symptom Checklist-90 <sup>161</sup>					
SLR	Straight-Leg Raising Test					
SBI	Sciatica Bothersomeness Index <sup>72</sup>					
SF-36	Medical Outcomes Study Short Form 36 <sup>132</sup>					
SFI	Sciatica Frequency Index <sup>72</sup>					
VAS	Visual Analogue Scale <sup>133</sup>					

## 1.7 Prognostic factors for poor outcome

Most studies on the prognostic factors for sciatica have been performed on surgically treated patients. Table 2 presents prospective cohort studies on prognostic factors in patients surgically treated for sciatica and lumbar disc herniation. The surgical treatments were discectomy or microdiscectomy. Excluded studies were those comparing different surgical techniques; studies on other lumbar diseases; for example spinal stenosis or degenerative disc; populations with mixed spinal diseases and where the radicular pain was not specified; and studies on other treatments, for example, chemonucleolysis or corticosteroid injections. Many of the studies in Table 2 were included in two existing reviews of lumbar spine surgery: one of predictors for spinal surgery outcome,<sup>26</sup> and the other of biopsychosocial risk factors after lumbar disc surgery.<sup>27</sup>

According to several of the studies in the surgically treated cohorts, factors associated with negative outcome were long duration of symptoms<sup>134, 135, 142, 146, 147, 150</sup> and MRI findings.<sup>116, 145, 149</sup> Clinical neurological findings were associated with outcome in a few studies.<sup>142, 143, 147, 151</sup> One Norwegian study indicated poor outcome in patients with impaired fibrinolytic activity, which has been suggested as being due to fibrin deposition, connective tissue invasion, and failure of surgery.<sup>137, 140</sup>

Psychosocial factors have also been associated with a poor outcome among surgically treated patients. As shown in Table 2 important predictors for poor outcome were limited education,<sup>134, 143</sup> aspects of work,<sup>116, 134, 145, 146</sup> somatization,<sup>134</sup> depression,<sup>135, 141, 145, 151</sup> anxiety<sup>138, 141</sup> and somatic distress.<sup>138, 140</sup> Smoking was a predictor for a negative outcome in two studies,<sup>142, 147</sup> and being female was associated with a poor outcome in three studies.<sup>116, 135, 140</sup>

There are conflicting results regarding prognostic factors for non-surgically treated patients with sciatica.<sup>162</sup> Table 1 shows that a longer duration of symptoms was associated with poor outcome in three studies.<sup>116, 118, 126</sup> In addition, clinical neurological findings were associated with a poor outcome in three studies,<sup>81, 118, 123</sup> whereas larger disc herniations on MRI were more likely to be associated with good outcome.<sup>85, 116</sup>

Psychosocial factors also seem to influence poor outcome in non-surgically treated patients, for example, poor strategies for coping with pain and social status,<sup>113</sup> longer duration of sick leave,<sup>1</sup> smoking,<sup>123</sup> and psychological distress.<sup>121</sup> Being female was associated with poor outcome in two studies.<sup>85, 123</sup>

In summary, there are many limitations in the existing knowledge on the prognosis and prognostic factors in patients with sciatica. For example, the inclusion criteria of the different studies vary due to the definition of sciatica, duration of symptoms, age limits, and recommendations on imaging. The large number of different outcome variables, ranging from self-reported improvement to complex composite measures of recovery, may also contribute to the conflicting results. The definitions of success on different scales are also not validated. Most outcome variables used in the studies are specific for low back pain or variables on quality of health; few measures are sciatica specific or validated for patients with sciatica. Patrick<sup>72</sup> validated the outcome measures Sciatica Bothersomeness Index, Sciatica Frequency Index, a modified 23-item version of the Roland Morris Disability Questionnaire, and the SF-36 subscales on patients with sciatica in the Maine study. The Maine–Seattle Back Questionnaire was validated in a sciatica cohort,<sup>163</sup> but not used as a predictor in the Maine study.<sup>80</sup> Finally, different types of prognostic factors are selected in the studies, for example, demographic, psychosocial, clinical and blood tests, imaging, and work-related factors. Statistical methodology also differs between the studies, where some use univariate and some others use multivariate analyses. The majority of the studies do not report the proportion of explained variance.

With this knowledge regarding the treatment and prognosis of patients with sciatica, it is evident that there is a need for a prospective observational study on patients treated both with and without surgical treatment. Results will be generalized more easily if as many patients from the clinics as possible are included; hence, wide inclusion criteria and a high rate of follow-up are required. It is also crucial to select prognostic factors of interest, which makes sense in a clinical perspective, and according to previous studies, have been shown to be important for the prognosis of patients with sciatica.

## **1.8 Aims**

The overall purpose of this study was to investigate outcome measures and prognostic factors for sciatica from a biopsychosocial perspective. The specific aims were:

- First, to carry out a translation and cross-cultural adaptation of the Tampa scale for Kinesiophobia from English to Norwegian, and to investigate the reliability, construct validity, and responsiveness of the Norwegian version within a Norwegian-speaking sample of patients with sciatica and disc herniation.
- Second, to investigate the clinical course for leg pain (Visual Analogue Scale [VAS]), back pain (VAS), Sciatica Bothersomeness Index, Maine–Seattle Back Questionnaire, SF-36 Bodily Pain and SF-36 Physical Functioning in patients treated with and without surgery for sciatica and disc herniation. After estimating the most optimal cut-off points for success (or not) on the patient-reported outcome measures (the cut-off point with highest sensitivity and specificity), the success rate after an episode of sciatica using the outcome cut-off values will be determined.
- Third, to investigate prognostic factors associated with non-success after one and two years of follow-up for sciatica in patients referred for secondary care.

## **2 Methods**

### **2.1 Design**

The study was a 2-year prospective observational multicenter cohort study. Patients were treated as usual, and no experiments were conducted. Both cross-sectional (Paper I) and longitudinal observations (Papers II and III) were used.

### **2.2 Study population**

#### **2.2.1 Sciatica cohort**

Patients were recruited from specialty back clinics at four hospitals in southeast Norway: Sykehuset Østfold, Sørlandet Sykehus, Oslo Universitetssykehus Ullevål, and Sykehuset Innlandet. Patients were referred to the specialty back clinics from the primary health care service. The inclusion period was from January 2005 to December 2006. Consecutive eligible patients were invited to participate in the study by the clinic staff.

#### **2.2.2 Inclusion and exclusion criteria**

Inclusion criteria were 18 years of age or older, radiating pain or paresis below knee level, and a lumbar disc herniation at the corresponding level and side verified by MRI or CT. Exclusion criteria were inability to communicate in written Norwegian, pregnancy, prior surgery at the same disc level, and any condition that might cause sciatic symptoms (e.g., fracture, infection, malignancy).

### **2.3 Data collection**

#### **2.3.1 Procedures**

On the day of inclusion sociodemographic variables, back pain history, sciatica history, and patient-reported outcome measures were recorded by questionnaires. A clinical examination was performed by trained physiotherapists or physicians.<sup>164</sup> The testing procedures were standardized in meetings with the participating centers before inclusion of the patients.

Patients were followed up by questionnaires after 3, 6, 12, and 24 months. The questionnaires were sent to the patient by mail and returned in prepaid envelopes. A study nurse called the patient or sent a reminder by text message after 2 weeks if no



reply was obtained. Follow-up assessments included the outcome measures used at baseline and questions about treatments received since the previous follow-up. A reminder letter was sent to non-responders if no reply was obtained after 3 weeks.

### **2.3.2 Test–retest sample**

To investigate the reliability of the Norwegian version of the Tampa scale, a subgroup of 52 patients from Sykehuset Østfold completed the questionnaire twice, at baseline and then 2 days later. At baseline, a prepaid, self-addressed envelope containing the Tampa scale was given to patients to return by mail after 2 days. (Paper I)

## **2.4 Treatment**

Study participation did not involve any specific type of intervention and did not alter treatment considerations for the patients in the clinics. Patients received general information about sciatica and disc herniation, instructions about back exercises, advice to stay active, and pain medication if needed. Surgery was performed in patients with severe symptoms after discussion with an orthopedic surgeon at each center.

## **2.5 Measurements**

The selection of dependent and independent variables was based on a thorough literature review and emphasized measurements and variables that had good methodological properties, made clinical sense, and were easy to use. Table 3 summarizes all variables used in the three papers. The sociodemographic variables were included in a comprehensive questionnaire at baseline. The clinical examination comprised the following tests: Motor function was recorded using tiptoe or heel walking, knee extension, ankle dorsal motion and plantar flexion, big toe extension, and the Trendelenburg test. Sensory loss was tested by a light touch in the dermatomes L4, L5, and S1. Tendon reflexes were assessed for the Achilles tendon and patella. The straight-leg raising test was performed by elevating the leg with passive flexion of the hip and extension of the knee until radiating pain occurred, and the result was deemed abnormal if radiating pain was elicited at an angle of  $< 60^\circ$ .<sup>166</sup> Muscular performance and sensory and tendon reflexes deemed normal or abnormal. The number of tenderpoints (0–18) was recorded according to The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia.<sup>167</sup>

To assess surgery on the herniated disc as a prognostic factor, patients who underwent surgery during the observational period recorded the date of surgery in the next follow-up questionnaire.

Table 3. Summary of variables used in all three papers

Variable	Measure used	Paper I	Paper II	Paper III
<b>Demographics</b>				
Age	Years	X	X	X
Sex		X	X	X
Education	Years*	X	X	X
Current smoker	Yes/no	X	X	X
Work status	Working, on partial sick leave, on sick leave, undergoing rehabilitation, on disability pension, other	X	X	X
Duration current sciatica episode	Weeks	X†	X†	X‡
Duration back problems	<1, 1-5, >5 years§			X
Number of previous sciatica episodes	0, 1, 2, 3-4, 5-10, >10¶		X	X
<b>Clinical examination findings</b>				
Straight-leg raising test (<60°)	Normal, abnormal,	X	X	X
Sensory (dermatomal light touch)	Normal, abnormal	X	X	X
Muscular performance	Normal, abnormal	X	X	X
Reflexes (patellar or Achilles)	Normal, abnormal	X	X	X
Number of tenderpoints	0-18		X	
<b>Patient-reported outcomes</b>				
Kinesiophobia	Tampa Scale for Kinesiophobia	X		X
Pain intensity in the back	VAS back pain	X	X	X
Pain intensity in the leg	VAS leg pain	X	X	X
Disability due to sciatica	Maine-Seattle Back Questionnaire#	X	X	X
Sciatica-specific symptoms	Sciatica Bothersomeness Index#	X	X	X
Sciatica-specific symptoms	Sciatica Frequency Index#	X		
Fear avoidance beliefs	Fear Avoidance Beliefs Questionnaire	X		
General health	SF-36	X	X**	
Emotional distress	Hopkins Symptom Checklist-25	X		X
Comorbidity	Subjective health complaints			X
Change of symptoms	Global change scale (1-7)	X	X	
Satisfaction	Satisfied to spend the rest of life in the current state, scale (1-5)		X	
<b>Surgery</b>				
Previous back surgery	Yes/no	X	X	
Surgery for disc herniation during the observational period	Yes/no		X	X

\* Collapsed to ≤ 12 years or >12 years. † Collapsed to <3, 3-6, or >6 months. ‡ Dichotomized to <3 months or ≥3 months. § Dichotomized to 0 or ≥1 years. ¶ Dichotomized to 0 or ≥1 episodes. # The Maine-Seattle Back Questionnaire, Sciatica Bothersomeness and Frequency Indices were translated to Norwegian according to the same translation procedures as the Tampa Scale for Kinesiophobia<sup>165</sup>\*\* The subscales bodily pain and physical functioning were used

### 2.5.1 Patient-reported outcomes

The **Tampa Scale for Kinesiophobia**<sup>168</sup> comprises 13 items rated on a 4-point Likert scale, and includes the following response choices: strongly disagree, disagree, agree, and strongly agree. The total score ranges from 13 to 52, and higher scores indicate greater kinesiophobia.<sup>33</sup> This version does not include any of the original four negatively phrased items that appeared in the first version.<sup>32</sup>

A horizontal **visual analogue scale (VAS)** ranging from 0 (no pain) to 100 (worst pain ever) was used to measure pain intensity in the back and in the leg (sciatica) during the previous week. This method is used frequently to assess pain.<sup>133</sup>

The **Maine–Seattle Back Questionnaire**<sup>163</sup> was used to assess disability and functional limits caused by sciatic and back pain. This questionnaire was developed from the Roland Morris Disability Questionnaire<sup>130</sup> and was modified for sciatica and spinal stenosis patients in the Maine study.<sup>169</sup> The scale comprises 12 items, each with the answer yes (1) or no (0). The score ranges from 0 to 12, and a higher score indicates worse symptoms.

The **Sciatica Bothersomeness Index** is a composite score of four questions about symptoms during the previous week: (1) leg pain; (2) numbness or tingling in the leg, foot, or groin; (3) weakness in the leg or foot; and (4) back or leg pain while sitting.<sup>72</sup> Scores are in the range of 0 to 6 for each item; the total score ranges from 0 to 24, and a higher score indicates worse symptoms.

The **Sciatica Frequency Index** is a composite score of the frequency during the previous week of the same four questions from the Sciatica Bothersomeness Index.<sup>72</sup> The answer categories are: not at all, very rarely, a few times, about half the time, usually, almost always, and always. Scores are in the range of 0 to 6 for each item. The total score ranges from 0 to 24, and a higher score indicates worse symptoms.

The **Fear Avoidance Beliefs Questionnaire**<sup>29</sup> was used to measure pain-related fear beliefs and behavioral avoidance. The Norwegian-translated and cross-culturally adapted questionnaire was used.<sup>170</sup> It contains two subscales: the Fear-Avoidance Beliefs Questionnaire/work has seven items and the Fear-Avoidance Beliefs Questionnaire/physical activity has four items. Scores are in the range of 0 to 6 for each item. The total scores range from 0 to 42 for the Fear-Avoidance Beliefs Questionnaire/work and from 0 to 24 for the Fear-Avoidance Beliefs Questionnaire/physical activity.

**Medical Outcomes Study Short Form 36 (SF-36)** was used to evaluate general health-related quality of life and includes eight subscales: physical functioning, role-emotional, role-physical, mental health, general health, vitality, social functioning, and bodily pain.<sup>132, 171</sup> The full-length questionnaire and abbreviated versions are used widely, including in studies on sciatica. For example, recent studies on sciatica used the subscales bodily pain and physical functioning.<sup>79, 80, 96</sup> The SF-36 scores range from 0 to 100, and higher values indicate better health.

The **Hopkins Symptom Checklist-25** was designed to measure emotional distress.<sup>172</sup> The questionnaire includes 25 items on depression, anxiety, and somatization during the previous week, and the scores range from 1 (not at all) to 4 (extremely). Adding all item scores and dividing the sum by the number of completed items yields a mean score. A Norwegian epidemiological study found that patients with scores  $\geq 1.75$  are in need of care.<sup>173</sup>

The **Subjective Health Complaints** questionnaire lists 29 common somatic and psychological complaints.<sup>174</sup> Patients grade the intensity of each complaint during the previous month as not at all (0), a little (1), some (2), or severe (3). Each item is dichotomized to absent (0) or present (1–3). Two of the items, low back pain and leg pain during exercise, were excluded, giving a dichotomized score range from 0 to 27.

A **seven-item global change scale** was included for leg and back pain. The responses were completely recovered, much better, better, a little better, no change, a little worse, and much worse.

**Satisfaction to spend the rest of life in the current state** was registered using a 5-item Likert scale with the wording of responses as very satisfied, a little satisfied, neither satisfied nor not, a little dissatisfied, and very dissatisfied.

## 2.6 Statistics

### 2.6.1 Sample size calculation

It has been suggested that for prognostic studies, at least 10 outcome events are required for each factor studied.<sup>175</sup> In the present study, a total of 20 prognostic factors were included. Because there was no consensus about the optimal definition of “outcome events” for sciatica when planning this study, we could not precisely estimate the sample size a priori. However, based on the previous Maine study, we expected that surgical treatment would be necessary for 30% of the patients and that 30% of those

who were treated surgically and 50% of those who were not treated surgically would not experience a successful outcome after 1 year.<sup>80</sup> With a sample of 400 patients and poor events for 50% of the sample, it would be power for about 20 prognostic factors.

## 2.6.2 Statistical methods used

Table 4 provides an overview of all statistical methods used in the three papers. The Statistical Package for Social Sciences versions 14.0-18.0 (SPSS Inc., Chicago, IL), and Medcalc3000 were used for data analyses.

Table 4. Statistical methods used

Statistical method	Paper I	Paper II	Paper III
Student's <i>t</i> -test <sup>176</sup>	X	X	X
Chi-square test <sup>176</sup>	X	X	X
Principal component analysis <sup>177</sup>	X		
Floor or ceiling effects <sup>178</sup>	X	X	
Repeatability <sup>179</sup>	X		
Coefficient of variance	X		
Weighted kappa coefficient <sup>180</sup>	X		
Internal consistency <sup>181</sup>	X		
Pearson's correlation coefficient <sup>176</sup>	X		
Standardized response mean <sup>182</sup>	X		
Standard error of measurement <sup>183</sup>	X		
Receiver operating characteristic (ROC) curve <sup>176</sup>		X	
Area under the ROC curve (AUC) <sup>184</sup>		X	
McNemar's test <sup>176</sup>		X	
Spearman's rho <sup>176</sup>			X
Univariate logistic regression <sup>177</sup>			X
Multivariate logistic regression <sup>177</sup>			X
Variance inflation factor <sup>185</sup>			X
Nagelkerke R squared <sup>177</sup>			X

## 2.6.3 Presentation of data

Means and standard deviations (SDs) were calculated for numerical variables, and frequencies were calculated for categorical variables. The normal distribution of each continuous variable was judged by visual inspection.

## 2.6.4 Missing items

Missing items in questionnaires were substituted with mean values (Paper III).

## 2.6.5 Bivariate examinations

All three papers include statistical analyses to test differences between means in different groups, using independent *t*-tests.

### **2.6.6 Principal component analysis**

Principal component analysis is used to identify similar or identical items to develop a more comprehended and understandable questionnaire, and to establish subscales. Principal component analysis is in many ways similar to factor analysis. The two terms are sometimes used interchangeably, and many authors use the term “factor” to refer to the output of both types of analyses.<sup>177</sup> Factor analysis is a more theoretical solution, whereas principal component analysis gives an empirical summary of the data set, or a data reduction of the correlated observed variables to a smaller set of important independent composite variables.<sup>186</sup>

In the present study, principal component analysis was used to evaluate the factor structure of the Tampa Scale for Kinesiophobia. Factors with eigenvalues >1 were extracted, and the factor structure was explored by Varimax rotation (orthogonal). Factors with a loading  $\geq 0.30$  were included in the model.<sup>177</sup>

### **2.6.7 Floor or ceiling effects**

Positive floor or ceiling effects were classified as >15% of scores in the lowest or highest parts of the scales, respectively.<sup>178</sup> The possible consequence of floor or ceiling effects is that extreme values are not distinguished from high or low values (content validity, reliability) and that changes cannot be registered as well as possible (responsiveness). For example, in a patient with a floor effect, the score can only be increased, whereas in a patient with a ceiling effect, the score can only be reduced.

### **2.6.8 Reliability**

The repeatability test described by Bland and Altman was used to calculate measurement error.<sup>179</sup> The test is based on the standard error of measurement between test and retest scores. The standard error of measurement is calculated from the mean square error term in a repeated-measures analysis of variance model.<sup>183</sup> The error of measurement is the difference between an obtained score and its theoretical true score.<sup>187</sup> The standard error of measurement, which is the measurement error of test-retest reliability, is multiplied by 2.77 ( $1.96 \times \sqrt{2}$ ) to obtain the 95% confidence interval (CI) of repeatability. This is the smallest amount of within-person change between the two different measurements at  $p < 0.05$  that must be observed before the change can be considered to exceed the measurement error; this is called the minimum detectable

change or smallest detectable change.<sup>178</sup> As recommended by Bland and Altman,<sup>179</sup> plots were created to show the difference between the two measures against the mean of the sum scores.

To provide a unitless percentage of the measurement error, the average coefficient of variance for paired measurements was calculated as the ratio of SD divided by the mean and multiplied by 100.<sup>179</sup>

To assess the repeatability of each item separately, a quadratic weighted kappa coefficient was calculated. Interpretations of kappa were as follows:  $K > 0.81$ , almost perfect;  $K > 0.61$ , substantial;  $K > 0.41$ , moderate;  $K > 0.21$ , fair; and  $K > 0.01$ , slight.<sup>180</sup>

### **2.6.9 Internal consistency**

Internal consistency was measured with Cronbach's alpha and describes the pairwise correlations between the items and how well each individual item in a scale correlates with the sum of the remaining items. A Cronbach's alpha between 0.7 and 0.8 is regarded as satisfactory.<sup>181</sup> A very high Cronbach's alpha ( $>0.95$ ) suggests that some items may be too similar and are thus redundant. Cronbach's alpha depends on the number of items; the more items in the questionnaire, the higher the internal consistency.<sup>188</sup>

### **2.6.10 Construct validity**

The Pearson's correlation coefficients were calculated to investigate the relationship between variables in the assessment of construct validity.<sup>189</sup> Construct validity was based on a priori defined hypotheses about the relationship between the Tampa scale and other measures (Paper I). We expected a high correlation between the Tampa scale and the Fear-Avoidance Belief Questionnaire and weaker to moderate correlations between the Tampa scale and SF-36, and between the Tampa scale and VAS pain. A correlation  $r \geq 0.60$  was considered high, a correlation between 0.30 and 0.60 moderate, and a correlation  $r \leq 0.30$  weak.<sup>190</sup>

Spearman's rho, a nonparametric measure of the correlation between variables, was used to quantify the associations between independent variables in the multivariate logistic regression analysis (Paper III). A value  $\geq 0.70$  was considered high, indicating a high correlation between potential prognostic variables.

### 2.6.11 Responsiveness

Responsiveness is the ability of an outcome measure to detect clinically important changes over time and reflects the validity of a change in score. Responsiveness of the Tampa scale was estimated by standardized response means, which is a distribution-based approach.<sup>191</sup> Standardized response means were calculated by dividing the Tampa Scale mean score change between inclusion and 3 months by the SD of the mean change score. Responsiveness was analyzed in two subgroups: *improved* and *unchanged*. Using the global change scale at 3 months, completely gone, much better, and better were categorized as *improved*, a little better as *unchanged*, and a little worse as *unchanged*. Standardized response mean values  $>0.80$ ,  $>0.5$ , or  $>0.2$  were considered large, moderate, and small, respectively.<sup>192</sup>

The receiver operating characteristic (ROC) curve of a continuous variable is an anchor-based approach.<sup>191</sup> In Paper II, we used the global change scale as the anchor or external criterion, and patients were categorized as success if they responded that they had recovered completely or were much better. The ROC curve provides a graph of sensitivity and specificity of all possible cut-offs of a continuous variable. The value on the curve closest to the upper left corner shows the highest sensitivity and specificity, and is generally considered the best cut-off score to distinguish patients according to the external criterion. The area under the ROC curve (AUC) reflects the variable's ability to differentiate between the defined external criterion. An AUC value  $>0.70$  was considered satisfactory.<sup>178</sup>

McNemar's test was used in Paper II to estimate the outcome measure with highest sensitivity where the cut-off was set to a specificity of 80% for each measure using ROC curve analysis.

### 2.6.12 Logistic regression

Paper III includes a logistic regression analysis in which the association between the independent baseline variables and the dependent variable were analyzed. The dependent variable was a dichotomized variable reflecting successful or unsuccessful outcomes. First, univariate logistic regression analysis was used to examine each of the independent baseline variables as one single explanatory variable. Potential prognostic variables were age, sex, history of sciatic and back symptoms, work status, clinical examination findings, comorbidity, and psychosocial factors as smoking, emotional



distress and kinesiophobia. These variables were chosen to cover the most important aspects of the biopsychosocial model.

$p$  is the probability of a subject having the defined outcome, and  $1 - p$  is the probability of not having the outcome. The ratio  $p/(1 - p)$  is called the odds. When comparing two groups the odds ratio (OR) is the proportion between the odds in the two groups.<sup>176</sup> The OR was presented with a 95% CI. An OR value of 1 means that the event is equally likely in both groups. An OR value  $>1$  indicates there is an association between the independent and dependent variable. Larger OR values indicate a stronger association. Nagelkerke R squared was used to measure how well the independent variable(s) in the logistic regression analyses explained the outcome.

In the multivariate logistic regressions analysis, the independent variables that were significantly associated with outcome with a  $p$ -value  $<0.10$  in the univariate regression analyses were included. We decided a priori to adjust for age, sex, and baseline scores of the Maine–Seattle Back Questionnaire. Multicollinearity was checked to judge to what extent one independent variable was associated with the others. If this association was high (i.e., a variance inflation factor  $>5$ ), the variable should be excluded from the model. In the last step, to assess the effect of surgery during the observation period on the outcome, the variable surgery (yes/no) was added to the model.

## **2.7 Ethical issues**

The study protocol was approved by the Regional Committee for Medical Research Ethics, Oslo (October 11, 2004) and The Ombudsman for Privacy in Research at the Norwegian Social Science Data Services (March 2, 2005). All patients received both written and oral information about the study before participating in the study and gave an informed, signed consent. The participants were also informed about their right to withdraw from the study at any time.

### 3 Summary of the results

The study included 466 patients. The mean age (SD) was 43.6 (11.5) years; 268 (57.5%) were men. One hundred ninety-two (41.4%) patients had a duration <3 months counted. Two hundred ten patients (45.3%) experienced their first sciatica episode, and 145 (31.1%) had back problems for <1 year. There were 200 (43.3%) smokers, and 92 (19.8%) were working full time.

#### 3.1 Paper I

The Tampa scale was translated and back-translated according to recent guidelines for cross-cultural adaptation of self-report measures.<sup>193</sup> The principal component analysis of the Tampa scale produced three components. The first component was strongly related to physical activities, the second component was related to fear beliefs, and the third component was associated with catastrophizing. In the total cohort, Cronbach's alpha was 0.81 for the total Tampa Scale for Kinesiophobia, and 0.75, 0.70, and 0.46 for factors 1, 2, and 3, respectively. Component 1 showed a floor effect in which 152 (33.3%) patients achieved the lowest possible score.

Test-retest reliability was calculated for 47 patients (32 men and 15 women) with a 2-day interval. The mean score (SD) on the Tampa scale was 24.5 (6.7) for the test and 24.2 (5.5) for the retest; the test and retest scores did not differ significantly. Repeatability according to Bland & Altman<sup>179</sup> was 8, and the coefficient of variance was 11%, indicating a high level of test-retest reliability for the Tampa Scale for Kinesiophobia. Further, four items (6, 7, 9, and 10) had a moderate quadratic-weighted kappa (0.41–0.60), and the remaining nine items showed substantial concordance (0.61–0.80).

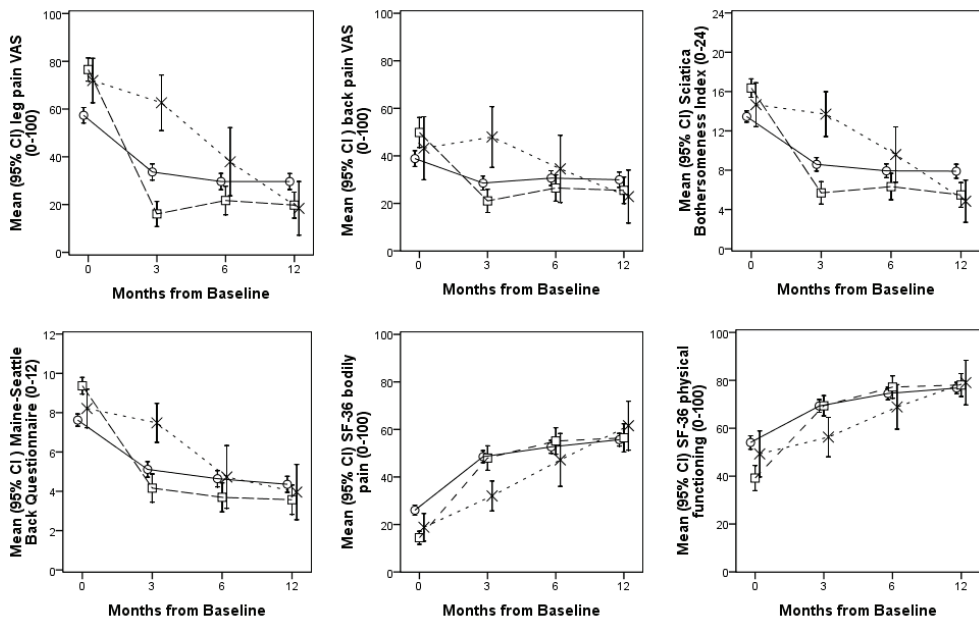
The highest correlations ( $r = 0.51$  for both) were found between the total score on the Tampa scale and the Fear-Avoidance Beliefs Questionnaire/physical activity, and between Tampa scale component 2 and the Fear-Avoidance Beliefs Questionnaire. The Tampa scale correlated poorly with pain, as measured by VAS in the back and leg ( $r = 0.18$  and  $0.15$ , respectively). The correlation coefficient between the Tampa scale and SF-36 bodily pain was also low ( $r = -0.24$ ).

The standardized response means for the Tampa scale and the three factors were low to moderate for the improved group and low for the unchanged group. For the

Tampa scale total score, the standardized response means were  $-0.42$  for the improved group and  $0.14$  for the unchanged group. This indicates that the Tampa scale showed a moderate ability to detect changes in symptoms over a 3-month period in the sciatica patient group.

### 3.2 Paper II

During the first year of follow-up, symptoms decreased in all outcome measures: leg pain; (VAS); back pain (VAS); Sciatica Bothersomeness Index; Maine–Seattle Back Questionnaire; SF-36 bodily pain; and SF-36 physical functioning (Figure 1). The two subgroups of surgically treated patients had worse baseline scores, and the patients operated on before the 3-month follow-up had a faster reduction in symptoms. Figure 1 also shows that the patients treated with surgery between the 3- and 12-month follow-up had a slower reduction in symptoms. At the 1-year follow-up, the mean scores were about the same in the three subgroups.



**Figure 1.** Time course for leg pain (VAS), back pain (VAS), Sciatica Bothersomeness Index, Maine-Seattle Back Questionnaire, SF-36 bodily pain, and SF-36 physical functioning (means with 95% CI) in patients treated without surgery (n=289) —○—, surgery before 3 months registration (n=98) —□—, and surgery between 3 and 12 months registration (n=28) ----x----.

On the global change scale, 54% of the patients reported success at the 1-year follow-up: 69% of the surgically treated group and 48% of the non-surgically treated group. Overall, 46% of the patients indicated they were satisfied to spend the rest of their lives in their current state; 60% were in the surgically treated group and 40% were in the non-surgically treated group.

The cut-off values with the highest sensitivity and specificity that distinguished between patients with and without success of sciatica at the 1-year follow-up were: 17.5 on the VAS for leg pain (range 0–100); 22.5 on the VAS for back pain (range 0–100); 6.5 on the Sciatica Bothersomeness Index (range 0–24); 4.5 on the Maine–Seattle Back Questionnaire (range 0–12); 51.5 for bodily pain on the SF-36; and 81.7 for physical functioning on the SF-36 (ranges for SF-36, 0–100, with higher values indicating better health). The ROC analyses showed high AUCs (0.76–0.89) for all measures at all follow-up times.

The percentage of patients achieving success according to the cutoff values on each measure at the 1-year follow-up were 52% for leg pain VAS, 53% for back pain VAS, 53% for the Sciatica Bothersomeness Index, 58% for the Maine–Seattle Back Questionnaire, 49% for the SF-36 bodily pain, and 50% for the SF-36 physical functioning.

Among the measures tested at the 1-year follow-up, the cut-off on the Maine–Seattle Back Questionnaire showed the highest sensitivity and specificity for distinguishing between patients with and without success. When the specificities of the measures were fixed at 80%, the sensitivity of the Maine–Seattle Back Questionnaire was 86%, which was significantly higher compared with the other outcome measures.

### **3.3 Paper III**

The response rates were 88% (n = 409) at the 1-year follow-up and 82% (n=380) at the 2-year follow-up. Among the responders, 120 patients (29%) received surgical treatment during the first year and nine patients (2%) during the second year of follow-up.

Of the sociodemographic variables, only male sex and smoking remained significantly associated with non-success at 1 year in the multivariate models. None of these were associated with poor outcome at 2 years. Education and work status, were

associated with poor outcome in the univariate analyses but not in any of the multivariate analyses.

Of the back pain/sciatica history variables, duration of back problems >1 year and duration of the index episode >3 months were both significantly associated with poor long-term outcome at 2 years. Previous sciatica episodes did not influence outcome at the 1- or 2-year follow-up.

In the univariate regression analyses, all the reported health status questionnaires were significantly associated with outcome at both 1- and 2-year follow-up. In the multivariate models, only Subjective Health Complaints were consistently associated with non-success at both follow-up times. The adjustment variable, the baseline scores of the Maine–Seattle Back Questionnaire, remained significant all the models both at 1 and 2 year follow-up. Of the pain-intensity scales (leg and back pain, VAS) only back pain remained significant in the final multivariate model at 1 year but not at 2 years. Leg pain was not associated with non-success in the multivariate models.

Univariate analysis of the psychological variables emotional distress and kinesiophobia showed both at 1 and 2 years associations with non-success; emotional distress showed an OR 3.3–3.4, whereas the OR for kinesiophobia was lower (ORs 1.1), although significant. In the multivariate analyses, emotional distress was not retained in the backward stepwise regression analysis, whereas kinesiophobia remained in the multivariate model at 2 years. The correlations assessed at baseline by Spearman's rho were between emotional distress and Subjective Health Complaints (0.59) and between emotional distress and kinesiophobia (0.41).

Of the clinical examination variables, only a reduced or absent tendon reflex at the 1-year follow-up showed a significant association with non-success in both univariate and multivariate analyses. Muscle weakness, sensory loss, and the straight-leg raising test were not significantly associated with non-success.

Finally, in the univariate analysis, patients not treated with surgery were more likely to have non-success at 1 year (OR = 1.66 [95% CI 1.07–2.58]), but no significant association was seen at 2 years (OR = 0.98 [95% CI 0.63–1.52]). The same associations were seen when surgical treatment was added in the multivariate analysis; no surgery was associated with non-success at 1 year (OR = 2.97 [95% CI 1.75–5.04]) but not at 2 years (OR 1.32 [95% CI 0.78–2.23]).

## **4 General discussion**

The main findings can be summarized as follows: Paper I demonstrated good psychometric properties of the Norwegian version of the Tampa scale in terms of high test-retest reliability and internal consistency, moderate construct validity, and low-to-moderate responsiveness.

Paper II reported variation in the proportion of patients who achieved success according to the calculated cut-off values in six outcome measures; they varied from 49% using SF-36 Bodily Pain to 58% using the Maine-Seattle Back Questionnaire. This indicates that success criterion has to be defined before comparing different studies.

Paper III demonstrated that the most important prognostic factors for non-success after one year were male sex, smoking, more subjective health complaints, and not having received surgical treatment; and after two years, a long duration of back pain and sciatica, more subjective health complaints, and a high level of kinesiophobia.

These primary results will be discussed with respect to methodological considerations, including design, patient selection and representativity, selection of prognostic and outcome variables, and statistical methods. Finally, the findings will be compared with other current relevant evidence.

### **4.1 Methodological considerations**

#### **4.1.1 Study design**

In the present thesis, a multicenter observational cohort study was used since the main goal was to investigate prognosis and prognostic factors among patients who experienced an episode of sciatica and who received surgical or non-surgical treatment for their condition. Observational studies are used primarily to identify prognostic factors and predictors for a defined outcome and in studies where a randomized controlled study is considered impossible or unethical.<sup>194</sup> Observational studies may also provide an indication of how treatments are utilized and to evaluate daily medical practice,<sup>195</sup> both providing doctor and patient a more realistic prognosis. The main weakness of using this design is that it limits evaluation of the effectiveness of the treatments provided since observational studies are more prone to bias and confounding than randomized trials are.<sup>196</sup> Hence, our results regarding the influence of

surgical versus non-surgical treatment on outcome after one and two years should be interpreted as exploratory due to the longitudinal observational design.

Longitudinal cohort studies may systematically overestimate the magnitude of the associations between exposure and outcome compared with randomized, controlled trials, but this was not found in well-designed observational studies.<sup>197</sup> Observational studies may be biased by loss to follow-up, but a loss to follow-up of 22% did not find to bias the results in a Norwegian study of 663 patients operated on for degenerative disorders of the lumbar spine.<sup>198</sup> The loss to follow-up in the present study was only 12% at one year and 18% at two years and is not regarded to be a major source of bias to our results, although non-responders were slightly different at baseline.

Selection of patients in a cohort study are usually pragmatic and apply criteria used in daily practice.<sup>199</sup> This design is usually considered to provide results with better external validity, as opposed to a randomized, controlled design, in which the selection criteria are usually stricter. For example a Dutch randomized study on surgical treatment versus prolonged conservative treatment of sciatica included patients aged between 18 and 65 years with symptom duration less than 12 weeks. Exclusion criteria were symptoms in the same dermatome within the past 12 months, severe comorbidity, or psychiatric illness.<sup>200</sup> Such strict selection criteria limits the generalizability of the results.

Moreover, by using a multicenter study, it was possible to recruit a large number of participants over the same time period, and the patients represented clinical practice from several of the hospitals in the South-East region of Norway. This might have strengthened the basis for generalization and the external validity of the study. On the other hand, carrying out a multicenter study creates challenges with respect to internal validity; for example, good interrater reliability is of importance, in particular, for the clinical examinations findings provided in the present study. The fact that we did not include an interrater reliability study of the clinical examinations findings in the present study must be regarded as a potential weakness. However, the doctors who included patients reported their clinical examination findings in a standardized sheet, which was designed to be as close to clinical practice as possible. Furthermore, the clinical examination tests were only distinguished with respect to normal or abnormal findings, and not to several levels of abnormality. Additionally, potential placebo and nocebo

effects through the clinical assessment and treatment can not be ruled out in this study, similar to all study designs that do not provide a placebo-control.<sup>105</sup>

Finally, it is important to mention that the outcome measures at the follow-up time points were not known by the clinicians involved in this study. A project coordinator, blinded for the type of treatment the patients had received, administered the follow-up questionnaires.

#### **4.1.2 Patient selection criteria**

In the present study there was no upper limit of age, which makes the cohort generalizable for all age groups. Further, there was no limit to the duration of sciatica symptoms or treatment given before inclusion, except surgery of the same disc as the current disc herniation. The inclusion period was over two years; thus, patients were recruited in all seasons. MRI or CT scanning verified the disc herniation. The response rate was high, over 80%, which is one of the criteria of a high quality study.<sup>176</sup>

A recent structured literature review reported a large variation in diagnostic eligibility criteria for randomized, controlled trials for lumbar disc herniation and radiculopathy.<sup>201</sup> Compared with for example, rheumatoid arthritis,<sup>202</sup> there is no established classification criterion for sciatica and lumbar disc herniation. This contributes to variation of inclusion criteria for different studies and is a challenge for comparison and generalization of studies. By example, in the Maine study, a disc herniation verified by imaging was not required; patients who according to clinicians had sciatica were accepted to be included in the study.<sup>169</sup> In contrast, in RCT studies disc herniation had to be confirmed either with radiculography,<sup>1</sup> CT,<sup>82</sup> MRI,<sup>200</sup> or either CT or MRI.<sup>96</sup> This is similar to the present study in which we did not include patients with sciatica without a disc herniation.

In the present study, radiating pain or paresis below knee level was included as an inclusion criteria. This allowed patients with paresis but without radicular pain to be included. This group constituted of only 1.5% of the total cohort which documents that radiculopathy with motor signs only in patients with disc herniation is rare. Contrary to other studies we did not require any clinical neurological findings for inclusion in the study. Weber<sup>1,83</sup> did not include patients without a positive straight-leg raising test, and in the Finnish randomized study, at least one specific physical finding was required.<sup>82</sup> In two trials reported after the start of the present study, Peul<sup>200</sup> included patients who had



radicular pain with and without mild neurological deficit, whereas the Spine Patient Outcomes Research Trial (SPORT)<sup>203</sup> included patients with either a positive straight-leg raising test or a positive femoral tension sign. With respect to demographic and clinical variables, our cohort closely resembled the samples of the Maine and the SPORT trials,<sup>79, 97</sup> which both included patients aged 18 years and older.

A potential limitation with respect to patient recruiting was incomplete recording of patients who, according to the inclusion and exclusion criteria, were eligible but for some reason were either not invited or declined to participate. Another minor limitation was that only patients recruited from Sykehuset Østfold participated in the test-retest procedure in the validation study (paper I). This was due to practical reasons of the administration of the retest questionnaire.

Despite a high response rate in the study, the few non-responders may bias the results. Generally, younger males are reported to drop out more often in follow-up studies. This was found in the Maine study at one year.<sup>80</sup> Missing data were also a limitation of the interpretation of results in the SPORT study.<sup>79</sup> Despite these limitations, we consider our patient sample to be representative for patients referred to secondary care due to sciatica.

#### **4.1.3 Selection of prognostic and outcome variables**

According to the biopsychosocial model and previous findings described in the introduction of this thesis, it was expected that both biological and psychosocial factors were associated with negative outcome. Since few studies had investigated the impact of psychosocial factors on sciatica treated conservatively, it was important to include these in this thesis.<sup>113, 121, 162</sup> den Boer<sup>27</sup> reported that a number of biopsychosocial variables was associated with unfavorable outcome after surgery for sciatica, in particular, passive avoidance coping strategies, anxiety, and somatization. The use of Subjective health complaints scale and Hopkins Symptom Checklist-25 in the current study are measures used in previous Norwegian populations<sup>22, 204-211</sup> and were therefore chosen as measures of psychosocial aspects of health in the present study. The Tampa scale has also been used in several studies.<sup>35, 212-214</sup>

Possible predictors that are not evaluated in the current study are: imaging;<sup>81, 85, 116, 119</sup> height and weight;<sup>116, 139, 215</sup> cardiovascular risk factors;<sup>216</sup> occupational

activities;<sup>215</sup> work-related factors;<sup>217</sup> time to surgery;<sup>147</sup> time on sick leave;<sup>218</sup> and identification of the centralization phenomenon.<sup>219, 220</sup>

When selecting outcome measures in the present study, it was important to apply validated and commonly used outcome measures. By choosing the same disease-specific and generic outcome measures as in the Maine study, we would be able to make comparisons with this study although differences in inclusion criteria and social insurance systems suggest a careful interpretation of the findings. At the time of planning the study, only one sciatica-specific outcome measure was validated in patients with sciatica, namely the Maine–Seattle Back Questionnaire.<sup>163</sup> Four questionnaires (the Tampa scale, Sciatica Bothersomeness Index, Sciatica Frequency Index, Maine–Seattle Back Questionnaire) were translated to Norwegian and validated in the present sciatica population. The validation of the Sciatica Bothersomeness Index, Sciatica Frequency Index and Maine–Seattle Back Questionnaire is presented in the other thesis on the same population.<sup>165</sup> The Sciatica Frequency Index was found to be very similar to the Sciatica Bothersomeness Index, except for the middle response category, which many patients avoided on the Sciatica Frequency Index. Hence, we chose the Sciatica Bothersomeness Index as one of the outcome measures in the estimates of success criteria (Paper II) and as an independent variable in the study of predictors for non-success (Paper III). In other studies on sciatica, the use of both the Sciatica Bothersomeness Index and the Sciatica Frequency Index have not provided more information than just using one.<sup>72, 78, 97</sup> In the SPORT study, both the Sciatica Frequency Index and the Sciatica Bothersomeness Index were presented at baseline.<sup>79, 96</sup> The values were very similar at baseline, and in the follow-up at one and two years, only Sciatica Bothersomeness Index was presented. In the SPORT study, the Oswestry Disability Index,<sup>129</sup> SF-36 Bodily Pain and SF-36 Physical Functioning were the primary outcome measures, with the Sciatica Bothersomeness Index as one of the secondary outcomes. Several studies on sciatica have used the Oswestry Disability Index.<sup>79, 81, 82, 111, 148-150</sup> Additionally, modifications of the Roland Morris Disability Questionnaire was used as a primary outcome in the Dutch study on early versus late surgical treatment for sciatica<sup>72, 78</sup> and as a secondary outcome in the Maine study.<sup>80</sup> The Oswestry Disability Index and the Roland Morris Disability Questionnaire were not evaluated as outcome measures in the present study because they do not contain any questions about leg pain,

which is important among patients with sciatica. This is a limitation of the study because it would have been easier to compare results with previously published studies.

Since previous studies have shown that definitions of good and poor outcome, or successful and non-successful outcome, varied largely across sciatica studies, it was important to investigate the optimal outcome measure and cut-off point that discriminated best between success and non-success (Paper II). This turned out to be 4.5 on the Maine–Seattle Back Questionnaire in the present study, and was therefore used as the definition for non-success in Paper III. The thorough validation of this outcome enhances the internal validity of the results from this study.

#### **4.1.4 Missing data on items**

The highest proportion of missing data in the present study was found in the Maine–Seattle Back Questionnaire, which was 5.0% at baseline. The item regarding sexual activity had the most missing data (2.8%).<sup>165</sup> This is in line with many quality-of-life studies, where the item with the highest percentage of missing data queried sexual activity.<sup>221</sup> However, a question about sexuality should be included in a questionnaire about sciatica because sexuality might be affected due to the nature of the disease. To justify for the missing items in questionnaires, the missing items were substituted with the arithmetic mean of values from the available items. This is recommended when items have mean response rates higher than 95% and responses are valid for more than half of the items.<sup>222</sup>

In the current study, we did imputation of data in Paper III, but not in Papers I and II. Additionally to the imputation of the Maine–Seattle Back Questionnaire, we also did imputation of single missing items of the Sciatica Bothersomeness Index, the Tampa scale, Hopkins Symptom Checklist-25, and Subjective health complaints scale in Paper III. This explains the varying results regarding rates of non-success by using the Maine–Seattle Back Questionnaire: 42% in Paper II and 44% in Paper III.

#### **4.1.5 Validation of patient-reported outcomes**

In this study, the translation and cross-cultural adaptation of the Tampa scale was carried out according to recent guidelines,<sup>223</sup> and the psychometric properties were tested according to recommendations from Terwee et al.<sup>178</sup> A cross-cultural adaptation is oriented toward measuring a similar phenomenon in different cultures.<sup>224</sup>

Comparisons between different populations are strictly speaking only possible after the questionnaire has been adapted and checked whether it is suitable in the relevant culture. Properly adapted questionnaires will make it easier to pool data in meta-analyses or systematic reviews when identical outcome variables are applied.

Validation of a self-report measure includes the process of determining whether it measures the underlying construct it is intended to measure and if it is useful for its purpose.<sup>221</sup> The underlying construct in the Tampa scale is kinesiophobia, which is an abstract phenomenon and hence a challenging construct to measure. Since there is no gold standard available for a construct like kinesiophobia, testing construct validity in terms of prespecified hypotheses is the recommended method.<sup>178</sup> We therefore decided to compare the kinesiophobia scores with the fear-avoidance beliefs scores,<sup>29</sup> pain scores (VAS), and the SF-36 scores,<sup>225</sup> which had been frequently used previously among patients with back pain.

We investigated test-retest reliability or measurement error by limits of agreement as described by Bland and Altman.<sup>226</sup> Since this measure is calculated from the variance within patients, it is considered a better method to estimate test-retest reliability than the intraclass correlation coefficients. No strict recommendation for sample sizes for test-retest studies exists, but a number more than 50 has been suggested.<sup>176, 178</sup> We chose to use a test-retest interval of two days, assuming that this would be long enough for patients not to recall their earlier responses. Test-retest intervals in comparable studies varied from 24 hours in one acute cohort<sup>227</sup> and one mixed acute/chronic cohort<sup>228</sup> to two weeks in a Swedish chronic low back pain cohort.<sup>229</sup> Although that the test-retest reliability of the Tampa scale has been validated in a large number of studies, both in different versions and languages and in a variety of musculoskeletal pain conditions,<sup>227-234</sup> none of the studies assessed measurement error by the limits of agreement as recommended by Bland and Altman.<sup>226</sup> Limits of agreement are important when interpreting change scores in research studies. If the change score is less than the measurement error, it can not be interpreted as a clinically meaningful difference.<sup>176, 178</sup>

Although the total score of the Tampa scale is used as a continuous variable, we also used weighted kappa statistics to evaluate each item as a categorical variable. The four response choices are strongly disagree, disagree, agree, and strongly agree, and each item might be validated as an ordinal measure. Terwee<sup>178</sup> recommended the

weighted kappa statistics to be used for ordinal measures to evaluate each item as a categorical variable. The argument for using the weighted kappa coefficient on each item in the Tampa scale is that this may allow for a better understanding of each statement.<sup>235</sup> The response choices are not continuous, and there is probably a larger difference between disagree and agree than between agree and strongly agree. Using a weighted kappa coefficient on each item in the Tampa scale has to our knowledge only been presented in one validation study of the Tampa scale on acute low back pain patients and showed fewer items with substantial and more items with moderate values compared to the current study.<sup>236</sup>

We used the standardized response mean<sup>191, 237</sup> to test the responsiveness of the Tampa scale. This method has recently been criticized by an expert Delphi panel,<sup>238, 239</sup> which argued that effect sizes represent measures of the magnitude of change due to an intervention or other event, rather than measures of the quality of the measurement instrument itself. The panel recommends testing prespecified hypotheses about the relationship of changes in the questionnaires with changes in other measures, which is similar to methods used for assessing construct validity.

#### **4.1.6 Estimating success or not**

In order to estimate the most optimal cut-off values for success or not in the patient-reported outcomes, an anchor-based approach and ROC analysis was used. The anchor or external criterion was a Global change scale assessed at the 12-month follow-up. Since there is no gold standard for an external criterion in situations like this, a global change scale is often used in methodological studies.<sup>240-242</sup> The lack of a gold standard for evaluating outcome is a limitation for the validation of outcomes measures.

Norman<sup>243</sup> argued that using a global change scale means accepting that a single-item global rating is superior to the multi-item measure under study. In this perspective, one might ask why not use the global change scores as the main outcome instead of longer outcome questionnaires, but there are both strengths and weaknesses with the global change scale.<sup>244</sup> Although the scale is quick and easy to administer, the scorings might be affected by potential sources of bias, for example the patient's current health state and other diseases or symptoms affecting health-related quality of life,<sup>245</sup> eager-to-please the clinician, and problems remembering the original status from when the actual

change started. This recall bias is regarded as one of the main limitations with the global change score.<sup>182, 243, 246</sup>

A global change scale has also been used as an external criteria when evaluating the minimum clinically important difference (MCID) of a continuous variable in spine surgery.<sup>247, 248</sup> Minimal Clinically Important Change (MCIC) and Minimal Important Change (MIC) are often used interchangeably with MCID.<sup>249-251</sup> For low back pain outcome measures, a 30% change from the baseline is proposed as a MIC and may be considered the clinically meaningful improvement.<sup>242, 252, 253</sup> However, the level of scorings of the patients might influence a percentage change score.<sup>254</sup> For example, in the current sciatica cohort, the baseline scores indicated a high degree of pain and disability. The mean leg pain VAS (0–100) score at baseline was  $63.2 \pm 28.2$ , and 100 patients (21.5%) had a score  $>90$ . A 30% change score for patients scoring  $>90$  would suggest a reduction to approximately 60 on the 0–100 scale. This indicates considerable leg pain, and hence, a recommendation of 30% improvement is probable not acceptable in this patient. Even a 50% improvement indicates much leg pain. To avoid this type of bias when estimating cut-off values based on change scores, we decided to use a cut-off value of the absolute follow-up scores. Using an absolute score may also be easier in a clinical perspective.

#### **4.1.7 Statistical methods for prognostic factors**

We used a logistic regression model with a stepwise backward approach to evaluate factors associated with non-success. Potential methodological problems in the current analyses are confounding factors with the likelihood of type I and type II errors. Confounders are defined as extraneous variables that correlate both with the dependent and independent variables. This may lead to an overestimate (type I error) or underestimate (type II error) of an effect. A multiple regression model can control for potential confounders. In the present study, none of the variables showed high correlations with the dependent and independent variables. We decided a priori, to adjust for age, gender, baseline variables of the dependent variable Maine-Seattle Back Questionnaire and whether the patients received surgical or non-surgical treatment. Type of treatment was the only potential prognostic factor not registered at the inclusion of the patients into the study because the decision about surgical treatment was taken after baseline factors were registered. Therefore, potential baseline

predictors were first taken into the model and analyzed, and as the last step, the variable surgical treatment (yes/no) was added to the model.

## **4.2 Discussion of the results in comparison with current evidence**

### **4.2.1 Paper I**

This study represents the first report using the Tampa scale among patients with sciatica. Our version of the cross-culturally adapted Tampa scale was used in another validation on two patient groups with widespread pain and low back pain referred to the University Hospital of Northern Norway.<sup>255</sup> Damsgård found that the Tampa scale fitted the Rasch model and seemed to reflect a unidimensional construct of kinesiophobia. This result deviates from the current sciatica cohort, where the principal component analysis showed three components related to physical activities, fear beliefs, and catastrophizing. These differences may be due to the different cohorts, but may also be because of the different statistical methods used. The mean scores also support that the cohorts were different: while the current sciatica population had a mean baseline score of 27, the low back pain and widespread pain cohorts scored 31. In the sciatica cohort, we also found a floor effect of Component 1 physical activities, which indicated that many patients with sciatica did not report to have kinesiophobia. In a Swedish study of patients treated with surgery for disc herniation, half of the patients suffered from kinesiophobia 10-34 months after surgery.<sup>38</sup> One possible explanation for this phenomenon could be that fear of pain remains, even though leg pain is reduced. Additionally, pain-related fear may initiate and maintain chronic disability and pain.<sup>34</sup>

The original version of Tampa scale was in English and developed primarily for chronic low back pain patients.<sup>31</sup> The Tampa scale has been translated into several languages and applied in several musculoskeletal pain conditions. First, the Tampa scale was translated to Dutch, and until the late 1990s, most psychometric research was carried out with the Dutch version for various patient groups, such as chronic low back pain,<sup>33, 34, 256, 257</sup> acute low back pain,<sup>227</sup> fibromyalgia,<sup>256, 257</sup> osteoarthritis,<sup>258</sup> and the general population.<sup>259</sup> During the later years, the Tampa scale has been translated to Swedish,<sup>229</sup> Portuguese,<sup>228</sup> Spanish,<sup>260</sup> Chinese,<sup>261</sup> Italian,<sup>233</sup> and French (article in French).<sup>262</sup> These versions have been applied in patient groups with nonspecific low back pain,<sup>228, 229, 233, 234, 263, 264</sup> but also in neck pain,<sup>231</sup> whiplash,<sup>230</sup> shoulder pain,<sup>232</sup> patients with anterior cruciate ligament reconstruction,<sup>265</sup> and other acute or chronic

musculoskeletal pain conditions,<sup>260, 261</sup> as well as for injured workers.<sup>262</sup> Later, shorter forms were developed, for example, there are 11-item<sup>232, 234, 265</sup> and 12-item versions,<sup>259</sup> with even shorter versions containing only three items<sup>265</sup> and four items.<sup>261</sup> This indicates high interest of the concept of kinesiophobia.

Our principal component analysis, which showed that the Tampa scale contained three factors, (related to physical activities, fear beliefs, and catastrophizing) are inconsistent with earlier comparable studies that have used the same 13-item version as was used in this thesis. These studies have mainly showed a two-factor model (related to activity avoidance and somatic focus or “harm”).<sup>168, 233, 256, 258, 263</sup> The differences in results might be due to differences between the study populations. There might also be cultural differences, but only a multicenter study across country borders would answer that question.

As expected, of all the measures compared in the current construct validity analyses, the highest correlation was found between the Tampa scale and the Fear Avoidance Beliefs Questionnaire. These two instruments measure similar constructs, although the moderate correlation ( $r = 0.51$ ) suggests that these constructs do not overlap completely. None of the instruments provides any direct information regarding the individual’s exact fear. Fear is a negative sensation, usually a response to a specific stimulus, such as pain, or a specific activity. Fear may be an appropriate emotional response to a perceived threat and motivates to a reaction, either a defense behavior or a confrontation.<sup>33</sup> Phobia is an intense and irrational fear of something that in reality is not dangerous. Kori, who introduced the term kinesiophobia in 1990, suggested that chronic pain behavior has more to do with phobic than physical processes.<sup>31</sup> Questions have been raised as to whether kinesiophobia is really a phobia as comprehended in a strict psychological perspective.<sup>11</sup> Vlaeyen found many similarities when Kori’s original theories<sup>31</sup> of pain-related fear in chronic pain were compared with the specific phobia according to the Diagnostic and Statistical Manual of Mental Disorders-IV.<sup>266</sup> One difference was that in specific phobia, the person recognizes that the fear is excessive or unreasonable, while in pain-related fear, the person often does not.<sup>267</sup> In the validation of the Dutch version of the Tampa scale, Vlaeyen identified correlations between the Tampa scale and other psychological measures.<sup>33</sup> Correlations with the Tampa scale and Fear Survey Schedule which contains clusters of phobic complaints, were weak to moderate ( $r = 0.27-0.33$ ), whereas the highest correlations were with catastrophizing ( $r$



= 0.41–0.58) and depression ( $r = 0.50$ ). Finally, phobia is a concept most commonly used within the psychology and psychiatry field, and it can be discussed whether this is a precise term to use in a different clinical field, such as the management of sciatica.

Our study indicated that the Tampa scale had limited ability to detect symptom changes in the sciatica patient group over the first three-month follow-up period and, therefore, is probably more suitable as a predictor than an outcome measure. The low to moderate responsiveness in the present study differs from Woby's<sup>234</sup> report of chronic low back pain patients, in which there was a high responsiveness according to both standardised response means and ROC curves. Also, de Soursa<sup>228</sup> found that the Tampa scale had higher responsiveness compared with the Fear Avoidance Beliefs Questionnaire. One potential explanation for the different findings is that the interventions used in the present study were not directed towards a change in kinesiophobia, thus, there was no reason to expect a large change in the Tampa scale. Future studies should explore whether interventions directed towards kinesiophobia can improve the clinical course of sciatica.

#### **4.2.2 Paper II**

As Kamper et al.<sup>76</sup> presented in the systematic review of 82 studies on low back pain, including 14 studies on sciatica, there is an enormous variation in how recovery has been measured. In two of the studies, a composite measure, including many different scales of pain and function, was used.<sup>81, 85</sup> Eight of the reports did not describe the details of the definition for a good outcome for sciatica. This has contributed to the fact that practically no one has used the same definition and measurement of recovery or a successful outcome. This hampers a comparison across the studies and shows the need for a common definition of recovery or success.

The wording of success as opposed to recovery might also be discussed. In the literature, both terms are used.<sup>26, 76</sup> Success may be defined as achievement of an objective or a goal.<sup>268</sup> Mannion uses the wording "success" as an outcome considered in relation to the predominant aim of the disc surgery.<sup>26</sup> Recovery is to restore, or return to a defined condition, for example, good health.<sup>269</sup> When using the wording "recovery", one might assume that the condition was good before the disease started, for example no pain before the sciatica episode. This is not always the case in sciatica, as for example,

in the present study approximately half of the patients had experienced sciatica previously.

Our results in Paper II also showed that the most optimal outcome measure discriminating between success and non-success in the present patient sample was the sciatica-specific disability measure, the Maine–Seattle Back Questionnaire. To our knowledge, no studies have reported on different cut-off values on the outcome measures for a successful outcome in sciatica. Some studies have reported cut-off values for MCIC or MCID on various outcome measures when used in sciatica populations. For example, in the Maine study, estimates for the MIC (by using a little better as the threshold for clinical importance) after three months for the Sciatica Bothersomeness Index was about 3 points, and about 2-3 points for the Roland Morris Disability Questionnaire.<sup>72</sup> When performing similar analyses on the current population (as part of the other thesis on the same material), similar estimates for the MIC have been found; when using “better” on the global change scale as the cut-off point for clinical importance, a value of about 4 points was found for the Sciatica Bothersomeness Index, and about 2 points for the Maine–Seattle Back Questionnaire.<sup>165</sup>

Estimates of MCIDs have also been carried out on other outcome measures in patients undergoing spine surgery. In a previous study from a large database of lumbar surgeries in the United States, estimates for the VAS on back and leg pain, the Oswestry Disability Index, and the physical component summary of the SF-36 were provided.<sup>247</sup> Similar to this study, they found a wide range in the numbers of patients reaching the thresholds of MCID, dependent upon which outcome and which calculation methods were used. Moreover, MCIDs were also estimated for improvement and deterioration for patients undergoing spine surgery in a spine center in Switzerland by using the Core Outcome Measure Index, which is also a composite outcome measure.<sup>248, 270</sup> When comparing these studies, one should keep in mind that defining a minimum value, as done in the previous studies, is different from defining a desirable value for recovery or success, which our study has done. In addition, as discussed in section 4.1.6, there are certain disadvantages of estimating cut-off values based on a percentage change score because the baseline scorings of the patients may influence the change score. By using an absolute score, as was done in this study, we avoided this potential bias.

Guidelines for the clinical management of sciatica have been developed,<sup>8</sup> but apart from cauda equina syndrome, recommendations for referral to surgical treatment

depends on the patient's report of intractable radicular pain. This is opposite to, for example, inflammatory rheumatic diseases, where recent treat-to-target recommendations have been developed for rheumatoid arthritis and due to values of composite measures, treatment is adjusted until predefined recovery or remission is reached.<sup>271</sup> For chronic low back pain, the Outcome Measures in Rheumatology Special Interest Group has discussed a responder index, but concludes that further studies are needed.<sup>272</sup>

Further studies are needed to explore whether the measures used in the present sciatica cohort respond similarly in other sciatica samples. In this future work our results in terms of the cut-off values for the different instruments, will enable comparison of results between sciatica studies. Additionally, the results of the present study might ease calculations in meta-analyses.

#### **4.2.3 Paper III**

Our results showed that the comorbid subjective health complaints was the only prognostic factor significantly associated with non-success at both one and two years. This finding is in line with previous studies on low back pain patients.<sup>210, 273, 274</sup> Lower thresholds for reporting bodily discomfort when suffering from a painful condition like sciatica might be one potential explanation to this finding. Reports on pain and function are by definition subjective when the patient's perceptions and expectations play a vigorous part. Another study of the current cohort found that patients reported more subjective health complaints than the general population at baseline, and among those with persisting or worsening sciatica at the one year follow-up, nearly twice the number of subjective health complaints was seen compared with the general population.<sup>275</sup> To our knowledge, comorbidity has only been tested as a predictor in one study of sciatica: Carragee<sup>116</sup> identified concurrent medical illness as a significant predictor for poor outcome for surgically treated patients with a mean follow-up of 2.8 years. One Finnish health survey reported comorbidity in patients with sciatica, but the association was weak.<sup>276</sup>

Since Waddell introduced the biopsychosocial model for low back pain in 1987,<sup>18</sup> the importance of psychological and social factors in the transition from acute to chronic low back pain has been supported in several studies.<sup>277</sup> In the present study, kinesiophobia remained a significant prognostic factor in the multivariate analyses at

two years. Thus, kinesiophobia may play a role in the persistence of sciatica in the longer term. To our knowledge, there is no study on sciatica with a follow-up of more than one year where kinesiophobia was a potential predictor. Two studies on patients treated with lumbar disc surgery found an association between poor outcome and kinesiophobia at six weeks, but not at three or 12 month follow-up.<sup>35, 36</sup> Furthermore, a cross-sectional Swedish study of patients treated with surgery for disc herniation<sup>38</sup> reported that half the patients suffered from kinesiophobia 10-34 months after surgery. These patients were more disabled, had more pain, more catastrophizing thoughts, more symptoms of depression, lower self-efficacy, and poorer health-related quality of life than patients without kinesiophobia. This is in line with the current study.

Emotional distress was also associated with non-success with relatively high odds ratio (>3) at both one- and two-year follow-ups in the univariate analysis, but the effect disappeared when adjusting for other significant variables, as subjective health complaint and kinesiophobia. This is contrary to the previous Norwegian study of Graver<sup>138, 140</sup> where somatic distress predicted poor outcome both at one- and seven-year follow-ups, and the Maine study where distress (or mental health according to the subscale of the SF-36) predicted greater pain and self-perceived disability at the three year follow-up.<sup>121</sup> The study results can not be compared directly, however, due to varying measures of emotional distress, different sets of predictors, and outcome variables, however, in summary, the results of the current study indicate the importance of comorbidity, kinesiophobia and distress to the outcome of sciatica.

Long duration of back symptoms (>1 year) and current sciatica episode (>3 months) were also significantly associated with non-success at the two year follow-up. Several studies on patients receiving surgical treatment support that a long duration of back symptoms predict poor outcome.<sup>134, 135, 142, 146, 147, 150</sup> However, most of these studies had only one year of follow-up and were only on surgically treated patients.

Of the clinical examination variables included in this study, only reduced or absent tendon reflex at one year predicted non-success. Other clinical examination findings such as paresis and sensory loss predicted poor outcome in two previous studies on surgically treated patients.<sup>142, 143, 147</sup> The straight-leg raising test has shown divergent results in previous studies.<sup>118, 151</sup> Hence, there are no consistent findings with respect to the impact of clinical examination findings on the long-term prognosis of sciatica.

Moreover, none of the demographic and social factors in the present study (age, sex, work status, education, and smoking) were significant prognostic factors in the multivariate models. Exceptions were being male and smoking at one year. Our finding of associating male sex with non-success is opposite to other findings of mixed surgically and non-surgically treated populations.<sup>85, 123</sup> but these studies did not adjust for psychological predictors. In two studies on surgical cohorts, both being female and psychological variables predicted poor outcome.<sup>135, 140</sup> Hence, there are conflicting results with respect to the impact of sex on the outcome of sciatica. The current evidence also conflicts for smoking. While smoking was found to be a predictor for negative outcome in two studies of surgically treated patients,<sup>142, 147</sup> other studies have reported the opposite finding.<sup>116, 143, 150</sup> Smoking was not associated with outcome in any of the studies in the review of non-surgically treated patients with sciatica.<sup>162</sup>

Furthermore, surgical treatment was significantly associated with non-success at one year but not at the two-year follow-up. This finding confirms the results from other studies, where the benefits of surgical treatment decreased with time.<sup>1, 100, 278</sup>

In summary, although many of the current findings on prognostic factors for non-success are in line with previous research on sciatica, the heterogeneity of the available studies makes it difficult to draw firm conclusions regarding most of the prognostic factors. This underlines the need for future high-quality studies on sciatica.

## 5 Clinical implications and further research

An important finding of importance for the clinical management of sciatica is that the long-term prognosis is not very good for about half the patients, regardless of surgical or conservative therapy. This might be difficult to communicate to the patients suffering from sciatica. On the one hand, the clinician needs to balance the importance of providing an optimistic approach with good opportunities for a placebo effect, and on the other, it is important to provide realistic information regarding the prognosis when recommending surgery or conservative therapy (when it is a choice for the patient).

Similarly, both clinicians and researchers should be aware that the prognosis of sciatica is only weakly associated with clinical examination findings and more strongly associated with pain, disability, comorbid subjective health complaints, and kinesiophobia. This calls for a broader assessment of patients with sciatica than the traditional clinical assessment, in which the physical symptoms and signs are mainly investigated. The biopsychosocial model, therefore, is an important model in both clinic and research.

Moreover, the use of validated measures might be helpful in a broader assessment and may improve the process of clinical decision-making. The use of predefined cut-off values for success might ease the interpretation of the scorings from individual patients, as well as enhance the comparisons of results obtained in different studies/settings. It is important, however, to be careful with transferring results directly from prediction models in research to individual patients in the clinic. As stated by Moons,<sup>175</sup> prediction models are intended to help doctors and other clinicians make decisions by providing more objective estimates of probability as a supplement to other relevant clinical information. Furthermore, they can improve understanding of the determinants of the course and the outcome of patients with a particular disease. With this caution, it is evident that further research regarding interventions that are more effective in patients with sciatica is strongly needed. In particular, it remains to be seen if specific therapies directed toward significant prognostic factors will result in better outcomes.

## 6 Conclusions

- The Norwegian version of the Tampa scale was easily comprehended and demonstrated satisfactory validity and reliability for the assessment of fear of movement and/or (re)injury in patients with sciatica due to disc herniation.
- In a cohort of 466 patients, the success rate, defined as the number of patients who reported to be much better or completely recovered, was 52 % at the one-year follow-up. Using this criterion to calculate cut-off values for success on two sciatica-specific measures; the Sciatica Bothersomeness Index and the Maine–Seattle Back Questionnaire, two generic measures; the SF-36 Bodily Pain and the SF-36 Physical Functioning subscales, and back and leg pain VAS, the success rates varied from 49 to 58%. Thus, the success rates in sciatica depend on the outcome used.
- Our proposed cut-offs for success or non-success on different outcome measures may facilitate the comparison of success rates across studies in patients with sciatica.
- In multivariate analysis more baseline subjective health complaints predicted non-success at both one and two years. Male sex, smoking, more back pain, and a reduced tendon reflex were factors associated with non-success at one year, but not at two years. A longer history of back pain, longer duration of the current sciatica episode, and more kinesiophobia were predictors for non-success at two years. Non-surgical treatment was associated with non-success at the one-year, but not at the two-year follow-up.

## References

1. Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine*. 1983;8:131-140.
2. Robinson Joe S. Sciatica and the lumbar disk syndrome: A historic perspective. *South Med J*. 1983;76:232-238.
3. Allan DB, Waddell G. An historical perspective on low back pain and disability. *Acta Orthop Scand Suppl*. 1989;234:1-23.
4. Magnaes B. Quality assurance of back surgery. *Tidsskr Nor Laegeforen*. 1998;118:2135.
5. Mixter WJ, Barr JS. Rupture of the intervertebral disc with involvement of the spinal canal. *N Engl J Med*. 1934;211:210-215.
6. Mixter WJ. Rupture of the lumbar intervertebral disk: An etiologic factor for so-called "sciatic" pain. *Ann Surg*. 1937;106:777-787.
7. Parisien RC, Ball PA. William Jason Mixter (1880-1958). Ushering in the "Dynasty of the Disc". *Spine*. 1998;23:2363-2366.
8. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ*. 2007;334:1313-1317.
9. Waddell G. Low back pain: A twentieth century health care enigma. *Spine*. 1996;21:2820-2825.
10. Pincus T, Burton AK, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine*. 2002;27:E109-20.
11. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain*. 2000;85:317-332.
12. Keefe FJ, Rumble ME, Scipio CD, et al. Psychological aspects of persistent pain: Current state of the science. *J Pain*. 2004;5:195-211.
13. [http://www.iasp-pain.org/AM/Template.cfm?Section=General\\_Resource\\_Links&Template=/CM/HTMLDisplay.cfm&ContentID=3058#Pain](http://www.iasp-pain.org/AM/Template.cfm?Section=General_Resource_Links&Template=/CM/HTMLDisplay.cfm&ContentID=3058#Pain).
14. Bonica J. The need of a taxonomy. *Pain*. 1979;6:247-252.
15. Weiner BK. Spine update: The biopsychosocial model and spine care. *Spine*. 2008;33:219-223.
16. Engel GL. The need for a new medical model: A challenge for biomedicine. *Science*. 1977;196:129-136.
17. Waddell G. *The Back Pain Revolution*. Elsevier Health Sciences, 2004.
18. Waddell G. A new clinical model for the treatment of low back pain. *Spine*. 1987;12:632-644.
19. Nachemson AL, Jonsson E, Jonsson. *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis, and Treatment*. Lippincott Williams & Wilkins Philadelphia, 2000.
20. Burton AK, Tillotson KM, Main CJ, et al. Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine*. 1995;20:722-728.
21. Linton SJ. Do psychological factors increase the risk for back pain in the general population in both a cross-sectional and prospective analysis? *European Journal of Pain*. 2005;9:355-361.
22. Grotle M, Brox JI, Veierod MB, et al. Clinical course and prognostic factors in acute low back pain: Patients consulting primary care for the first time. *Spine*. 2005;30:976-982.
23. Pincus T, Vogel S, Burton AK, et al. Fear avoidance and prognosis in back pain: A systematic review and synthesis of current evidence. *Arthritis Rheum*. 2006;54:3999-4010.
24. Coste J, Lefrancois G, Guillemin F, et al. Prognosis and quality of life in patients with acute low back pain: Insights from a comprehensive inception cohort study. *Arthritis Rheum*. 2004;51:168-176.
25. Truchon M, Fillion L. Biopsychosocial determinants of chronic disability and low-back pain: A review. *J Occup Rehabil*. 2000;10:117-142.
26. Mannion AF, Elfering A. Predictors of surgical outcome and their assessment. *Eur Spine J*. 2006;15 Suppl 1:S93-108.



27. 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;15:527-536.
28. <http://www.iasp-pain.org/AM/AMTemplate.cfm?Section=Home&CONTENTID=7619&TEMPLATE=/CM/ContentDisplay.cfm&SECTION=Home>.
29. Waddell G, Newton M, Henderson I, et al. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain.* 1993;52:157-168.
30. Lethem J, Slade PD, Troup JD, et al. Outline of a fear-avoidance model of exaggerated pain perception--I. *Behav Res Ther.* 1983;21:401-408.
31. Kori SH, Miller RP, Todd DD. Kinisophobia: A new view of chronic pain behavior. *Pain Management.* 1990;3:35-43.
32. Miller RP, Kori SH, Todd DD. The Tampa scale. Unpublished report. Tampa, FL, 1991.
33. Vlaeyen JW, Kole-Snijders AM, Boeren RG, et al. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain.* 1995;62:363-372.
34. Crombez G, Vlaeyen JW, Heuts PH, et al. Pain-related fear is more disabling than pain itself: Evidence on the role of pain-related fear in chronic back pain disability. *Pain.* 1999;80:329-339.
35. den Boer JJ, Oostendorp RA, Beems T, et al. Continued disability and pain after lumbar disc surgery: The role of cognitive-behavioral factors. *Pain.* 2006;123:45-52.
36. Ostelo RW, Vlaeyen JW, van den Brandt PA, et al. Residual complaints following lumbar disc surgery: Prognostic indicators of outcome. *Pain.* 2005;114:177-185.
37. Lundberg M, Frennered K, Hagg O, et al. The impact of fear-avoidance model variables on disability in patients with specific or nonspecific chronic low back pain. *Spine.* 2011;36:1547-1553.
38. Svensson GL, Lundberg M, Ostgaard HC, et al. High degree of kinesiophobia after lumbar disc herniation surgery. *Acta Orthop.* 2011;82:732-736.
39. Adams MA, Roughley PJ. What is intervertebral disc degeneration, and what causes it? *Spine.* 2006;31:2151-2161.
40. Urban JP, Smith S, Fairbank JC. Nutrition of the intervertebral disc. *Spine.* 2004;29:2700-2709.
41. Grunhagen T, Shirazi-Adl A, Fairbank JC, et al. Intervertebral disk nutrition: A review of factors influencing concentrations of nutrients and metabolites. *Orthop Clin North Am.* 2011;42:465-77.
42. Antoniou J, Steffen T, Nelson F, et al. The human lumbar intervertebral disc: Evidence for changes in the biosynthesis and denaturation of the extracellular matrix with growth, maturation, ageing, and degeneration. *J Clin Invest.* 1996;98:996-1003.
43. Raj PP. Intervertebral disc: Anatomy-physiology-pathophysiology-treatment. *Pain Pract.* 2008;8:18-44.
44. Adams MA, McNally DS, Dolan P. 'Stress' distributions inside intervertebral discs. the effects of age and degeneration. *J Bone Joint Surg Br.* 1996;78:965-972.
45. Haefeli M, Kalberer F, Saegesser D, et al. The course of macroscopic degeneration in the human lumbar intervertebral disc. *Spine.* 2006;31:1522-1531.
46. Fardon DF, Milette PC, Nomenclature and classification of lumbar disc pathology. recommendations of the combined task forces of the north American spine society, American society of spine radiology, and American society of neuroradiology. *Spine.* 2001;26:E93-E113.
47. Sambrook PN, MacGregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: A magnetic resonance imaging study in twins. *Arthritis Rheum.* 1999;42:366-372.
48. Kuslich SD, Ulstrom CL, Michael CJ. The tissue origin of low back pain and sciatica: A report of pain response to tissue stimulation during operations on the lumbar spine using local anesthesia. *Orthop Clin North Am.* 1991;22:181-187.
49. Rydevik B, Brown MD, Lundborg G. Pathoanatomy and pathophysiology of nerve root compression. *Spine.* 1984;9:7-15.
50. Olmarker K, Rydevik B, Hansson T, et al. Compression-induced changes of the nutritional supply to the porcine cauda equina. *J Spinal Disord.* 1990;3:25-29.

51. Olmarker K, Blomquist J, Stromberg J, et al. Inflammatogenic properties of nucleus pulposus. *Spine*. 1995;20:665-669.
52. Shamji MF, Setton LA, Jarvis W, et al. Proinflammatory cytokine expression profile in degenerated and herniated human intervertebral disc tissues. *Arthritis Rheum*. 2010;62:1974-1982.
53. Mulleman D, Mammou S, Griffoul I, et al. Pathophysiology of disk-related sciatica. I.--evidence supporting a chemical component. *Joint Bone Spine*. 2006;73:151-158.
54. Takahashi N, Yabuki S, Aoki Y, et al. Pathomechanisms of nerve root injury caused by disc herniation: An experimental study of mechanical compression and chemical irritation. *Spine*. 2003;28:435-441.
55. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ*. 2006;332:1430-1434.
56. Breig A, Troup JD. Biomechanical considerations in the straight-leg-raising test. cadaveric and clinical studies of the effects of medial hip rotation. *Spine*. 1979;4:242-250.
57. van der Windt DA, Simons E, Riphagen II, et al. Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. *Cochrane Database Syst Rev*. 2010;(2):CD007431.
58. Thornbury JR, Fryback DG, Turski PA, et al. Disk-caused nerve compression in patients with acute low-back pain: Diagnosis with MR, CT myelography, and plain CT. *Radiology*. 1993;186:731-738.
59. Albeck MJ, Hilden J, Kjaer L, et al. A controlled comparison of myelography, computed tomography, and magnetic resonance imaging in clinically suspected lumbar disc herniation. *Spine*. 1995;20:443-448.
60. Karppinen J, Malmivaara A, Tervonen O, et al. Severity of symptoms and signs in relation to magnetic resonance imaging findings among sciatic patients. *Spine*. 2001;26:E149-54.
61. Modic MT, Obuchowski NA, Ross JS, et al. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology*. 2005;237:597-604.
62. Boden SD, Davis DO, Dina TS, et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*. 1990;72:403-408.
63. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*. 1994;331:69-73.
64. Dvorak J. Neurophysiologic tests in diagnosis of nerve root compression caused by disc herniation. *Spine*. 1996;21:39S-44S.
65. Wassenaar M, van Rijn RM, van Tulder MW, et al. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: A diagnostic systematic review. *Eur Spine J*. 2011. epub ahead of print DOI 10.1007/s00586-011-2019-8.
66. van Rijn RM, Wassenaar M, Verhagen AP, et al. Computed tomography for the diagnosis of lumbar spinal pathology in adult patients with low back pain or sciatica: A diagnostic systematic review. *Eur Spine J*. 2011. epub ahead of print DOI 10.1007/s00586-011-2012-2
67. Heliovaara M, Impivaara O, Sievers K, et al. Lumbar disc syndrome in Finland. *J Epidemiol Community Health*. 1987;41:251-258.
68. Heikkilä JK, Heikkilä K, Rita H, et al. Genetic and environmental factors in sciatica evidence from a nationwide panel of 9365 adult twin pairs. *Ann Med*. 1989;21:393-398.
69. Konstantinou K, Dunn KM. Sciatica: Review of epidemiological studies and prevalence estimates. *Spine*. 2008;33:2464-2472.
70. Lavy C, James A, Wilson-MacDonald J, et al. Cauda equina syndrome. *BMJ*. 2009;338:b936.
71. McCarthy MJ, Aylott CE, Grevitt MP, et al. Cauda equina syndrome: Factors affecting long-term functional and sphincteric outcome. *Spine*. 2007;32:207-216.
72. Patrick DL, Deyo RA, Atlas SJ, et al. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995;20:1899-1908; discussion 1909.
73. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: Summary and general recommendations. *Spine*. 2000;25:3100-3103.

74. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use.[erratum appears in spine 1999 feb 15;24(4):418]. [review] [17 refs]. *Spine*. 1998;23:2003-2013.
75. Grotle M, Brox JI, Vollestad NK. Functional status and disability questionnaires: What do they assess? A systematic review of back-specific outcome questionnaires. *Spine*. 2005;30:130-140.
76. Kamper SJ, Stanton TR, Williams CM, et al. How is recovery from low back pain measured? A systematic review of the literature. *Eur Spine J*. 2011;20:9-18.
77. Nykvist F, Hurme M, Alaranta H, et al. Severe sciatica: A 13-year follow-up of 342 patients. *Eur Spine J*. 1995;4:335-338.
78. Peul WC, van Houwelingen HC, van den Hout WB, et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med*. 2007;356:2245-2256.
79. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: The Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA*. 2006;296:2451-2459.
80. Atlas SJ, Deyo RA, Keller RB, et al. The Maine lumbar spine study, part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine*. 1996;21:1777-1786.
81. Balague F, Nordin M, Sheikhsadeh A, et al. Recovery of severe sciatica. *Spine*. 1999;24:2516-2524.
82. Osterman H, Seitsalo S, Karppinen J, et al. Effectiveness of microdiscectomy for lumbar disc herniation: A randomized controlled trial with 2 years of follow-up. *Spine*. 2006;31:2409-2414.
83. Weber H, Holme I, Amlie E. The natural course of acute sciatica with nerve root symptoms in a double-blind placebo-controlled trial evaluating the effect of piroxicam. *Spine*. 1993;18:1433-1438.
84. Thomas KC, Fisher CG, Boyd M, et al. Outcome evaluation of surgical and nonsurgical management of lumbar disc protrusion causing radiculopathy. *Spine*. 2007;32:1414-1422.
85. Jensen TS, Albert HB, Sorensen JS, et al. Magnetic resonance imaging findings as predictors of clinical outcome in patients with sciatica receiving active conservative treatment. *J Manipulative Physiol Ther*. 2007;30:98-108.
86. Selim AJ, Ren XS, Fincke G, et al. The importance of radiating leg pain in assessing health outcomes among patients with low back pain. results from the veterans health study. *Spine*. 1998;23:470-474.
87. Andersson GB, Svensson HO, Oden A. The intensity of work recovery in low back pain. *Spine*. 1983;8:880-884.
88. Hagen KB, Thune O. Work incapacity from low back pain in the general population. *Spine*. 1998;23:2091-2095.
89. Crook J, Milner R, Schultz IZ, et al. Determinants of occupational disability following a low back injury: A critical review of the literature. *J Occup Rehabil*. 2002;12:277-295.
90. Wolfe F, Pincus T. Listening to the patient: A practical guide to self-report questionnaires in clinical care. *Arthritis Rheum*. 1999;42:1797-1808.
91. Cherkin DC, Deyo RA, Loeser JD, et al. An international comparison of back surgery rates. *Spine*. 1994;19:1201-1206.
92. Jansson KA, Nemeth G, Granath F, et al. Surgery for herniation of a lumbar disc in Sweden between 1987 and 1999. an analysis of 27,576 operations. *J Bone Joint Surg Br*. 2004;86:841-847.
93. Rasmussen C, Nielsen GL, Hansen VK, et al. Rates of lumbar disc surgery before and after implementation of multidisciplinary nonsurgical spine clinics. *Spine*. 2005;30:2469-2473.
94. Sherman J, Cauthen J, Schoenberg D, et al. Economic impact of improving outcomes of lumbar discectomy. *Spine J*. 2010;10:108-116.
95. Peul WC, van den Hout WB, Brand R, et al. Prolonged conservative care versus early surgery in patients with sciatica caused by lumbar disc herniation: Two year results of a randomised controlled trial. *BMJ*. 2008;336:1355-1358.
96. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: The Spine Patient Outcomes Research Trial (SPORT): A randomized trial. *JAMA*. 2006;296:2441-2450.

97. Atlas SJ, Keller RB, Chang Y, et al. Surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: Five-year outcomes from the Maine lumbar spine study. *Spine*. 2001;26:1179-1187.
98. Atlas SJ, Keller RB, Wu YA, et al. Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the Maine lumbar spine study. *Spine*. 2005;30:927-935.
99. Gibson JN, Grant IC, Waddell G. The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine*. 1999;24:1820-1832.
100. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse: Updated Cochrane review. *Spine*. 2007;32:1735-1747.
101. Brox JI. The contribution of RCTs to quality management and their feasibility in practice. *Eur Spine J*. 2009;18 Suppl 3:279-293.
102. Jensen TT, Asmussen K, Berg-Hansen EM, et al. First-time operation for lumbar disc herniation with or without free fat transplantation. prospective triple-blind randomized study with reference to clinical factors and enhanced computed tomographic scan 1 year after operation. *Spine*. 1996;21:1072-1076.
103. Ronnberg K, Lind B, Zoega B, et al. Peridural scar and its relation to clinical outcome: A randomised study on surgically treated lumbar disc herniation patients. *Eur Spine J*. 2008;17:1714-1720.
104. Spangfort EV. The lumbar disc herniation. A computer-aided analysis of 2,504 operations. *Acta Orthop Scand Suppl*. 1972;142:1-95.
105. Turner JA, Deyo RA, Loeser JD, et al. The importance of placebo effects in pain treatment and research. *JAMA*. 1994;271:1609-1614.
106. Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med*. 2009;361:557-568.
107. Korhonen T, Karppinen J, Paimela L, et al. The treatment of disc-herniation-induced sciatica with infliximab: One-year follow-up results of FIRST II, a randomized controlled trial. *Spine*. 2006;31:2759-2766.
108. Genevay S, Finckh A, Zufferey P, et al. Adalimumab in acute sciatica reduces the long-term need for surgery: A 3-year follow-up of a randomised double-blind placebo-controlled trial. *Ann Rheum Dis*. 2011. Online First, 10.1136/annrheumdis-2011-200373
109. Luijsterburg PA, Verhagen AP, Ostelo RW, et al. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: A systematic review. *Eur Spine J*. 2007. 16:881-899.
110. Chou R, Atlas SJ, Stanos SP, et al. Nonsurgical interventional therapies for low back pain: A review of the evidence for an American pain society clinical practice guideline. *Spine*. 2009;34:1078-1093.
111. Iversen T, Solberg TK, Romner B, et al. Effect of caudal epidural steroid or saline injection in chronic lumbar radiculopathy: Multicentre, blinded, randomised controlled trial. *BMJ*. 2011;343:d5278.
112. Benoist M, Boulu P, Hayem G. Epidural steroid injections in the management of low-back pain with radiculopathy: An update of their efficacy and safety. *Eur Spine J*. 2011 Epub ahead of print. DOI: 10.1007/s00586-011-2007-z.
113. Hasenbring M, Marienfeld G, Kuhlendahl D, et al. Risk factors of chronicity in lumbar disc patients. A prospective investigation of biologic, psychologic, and social predictors of therapy outcome. *Spine*. 1994;19:2759-2765.
114. Nykvist F, Alaranta H, Hurme M, et al. Clinical findings as outcome predictors in rehabilitation of patients with sciatica. *Int J Rehabil Res*. 1991;14:131-144.
115. Nykvist F, Knuts LR, Alaranta H, et al. Clinical, social, and psychological factors and outcome in a 5-year follow-up study of 276 patients hospitalized because of suspected lumbar disc herniation. *Int Disabil Stud*. 1990;12:107-112.
116. Carragee EJ, Kim DH. A prospective analysis of magnetic resonance imaging findings in patients with sciatica and lumbar disc herniation: Correlation of outcomes with disc fragment and canal morphology. *Spine*. 1997;22:1650-1660.

117. Deyo RA, Andersson G, Bombardier C, et al. Outcome measures for studying patients with low back pain. *Spine*. 1994;19:2032S-2036S.
118. Vroomen PC, de Krom MC, Knottnerus JA. Predicting the outcome of sciatica at short-term follow-up. *Br J Gen Pract*. 2002;52:119-123.
119. Beauvais C, Wybier M, Chazerain P, et al. Prognostic value of early computed tomography in radiculopathy due to lumbar intervertebral disk herniation. A prospective study. *Joint Bone Spine*. 2003;70:134-139.
120. Atlas SJ, Chang Y, Keller RB, et al. The impact of disability compensation on long-term treatment outcomes of patients with sciatica due to a lumbar disc herniation. *Spine*. 2006;31:3061-3069.
121. Edwards RR, Klick B, Buenaver L, et al. Symptoms of distress as prospective predictors of pain-related sciatica treatment outcomes. *Pain*. 2007;130:47-55.
122. Sintonen H. The 15D instrument of health-related quality of life: Properties and applications. *Ann Med*. 2001;33:328-336.
123. Peul WC, Brand R, Thomeer RT, et al. Influence of gender and other prognostic factors on outcome of sciatica. *Pain*. 2008;138:180-191.
124. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonoperative treatment for lumbar disc herniation: Four-year results for the Spine Patient Outcomes Research Trial (SPORT). *Spine*. 2008;33:2789-2800.
125. Atlas SJ, Tosteson TD, Blood EA, et al. The impact of workers' compensation on outcomes of surgical and nonoperative therapy for patients with a lumbar disc herniation: SPORT. *Spine*. 2010;35:89-97.
126. Rihn JA, Hilibrand AS, Radcliff K, et al. Duration of symptoms resulting from lumbar disc herniation: Effect on treatment outcomes: Analysis of the Spine Patient Outcomes Research Trial (SPORT). *J Bone Joint Surg Am*. 2011;93:1906-1914.
127. Bergquist-Ullman M, Larsson U. Acute low back pain in industry. A controlled prospective study with special reference to therapy and confounding factors. *Acta Orthop Scand*. 1977;(170):1-117.
128. Graver V, Loeb M, Rasmussen F, et al. Clinical overall score: Outcome evaluation after lumbar disc surgery, assessments of reliability and validity. *Scand J Rehabil Med*. 1998;30:227-233.
129. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000;25:2940-52; discussion 2952.
130. Roland M, Morris R. A study of the natural history of back pain. part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine*. 1983;8:141-144.
131. Stauffer RN, Coventry MB. Anterior interbody lumbar spine fusion. analysis of Mayo clinic series. *J Bone Joint Surg Am*. 1972;54:756-768.
132. Ware JE, Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. conceptual framework and item selection. *Med Care*. 1992;30:473-483.
133. Von Korff M, Jensen MP, Karoly P. Assessing global pain severity by self-report in clinical and health services research. *Spine*. 2000;25:3140-3151.
134. Hurme M, Alaranta H. Factors predicting the result of surgery for lumbar intervertebral disc herniation. *Spine*. 1987;12:933-938.
135. Sorensen LV, Mors O, Skovlund O. A prospective study of the importance of psychological and social factors for the outcome after surgery in patients with slipped lumbar disk operated upon for the first time. *Acta Neurochir (Wien)*. 1987;88:119-125.
136. Little KB, Fisher J. Two new experimental scales of the MMPI. *J Consult Psychol*. 1958;22:305-306.
137. Haaland AK, Graver V, Ljunggren AE, et al. Fibrinolytic activity as a predictor of the outcome of prolapsed intervertebral lumbar disc surgery with reference to background variables: Results of a prospective cohort study. *Spine*. 1992;17:1022-1027.
138. Graver V, Ljunggren AE, Malt UF, et al. Can psychological traits predict the outcome of lumbar disc surgery when anamnestic and physiological risk factors are controlled for? results of a prospective cohort study. *J Psychosom Res*. 1995;39:465-476.

139. Graver V, Ljunggren AE, Loeb M, et al. Background variables (medical history, anthropometric and biological factors) in relation to the outcome of lumbar disc surgery. *Scand J Rehabil Med.* 1998;30:221-225.
140. Graver V, Haaland AK, Magnaes B, et al. Seven-year clinical follow-up after lumbar disc surgery: Results and predictors of outcome. *Br J Neurosurg.* 1999;13:178-184.
141. Kjellby-Wendt G, Styf JR, Carlsson SG. The predictive value of psychometric analysis in patients treated by extirpation of lumbar intervertebral disc herniation. *J Spinal Disord.* 1999;12:375-379.
142. Woertgen C, Holzschuh M, Rothoerl RD, et al. Does the choice of outcome scale influence prognostic factors for lumbar disc surgery? A prospective, consecutive study of 121 patients. *Eur Spine J.* 1997:173-180.
143. Woertgen C, Rothoerl RD, Breme K, et al. Variability of outcome after lumbar disc surgery. *Spine.* 1999;24:807-811.
144. Prolo DJ, Oklund SA, Butcher M. Toward uniformity in evaluating results of lumbar spine operations. A paradigm applied to posterior lumbar interbody fusions. *Spine.* 1986;11:601-606.
145. Schade V, Semmer N, Main CJ, et al. The impact of clinical, morphological, psychosocial and work-related factors on the outcome of lumbar discectomy. *Pain.* 1999;80:239-249.
146. Nygaard OP, Kloster R, Solberg T. Duration of leg pain as a predictor of outcome after surgery for lumbar disc herniation: A prospective cohort study with 1-year follow up. *J Neurosurg.* 2000;92:131-134.
147. Rothoerl RD, Woertgen C, Brawanski A. When should conservative treatment for lumbar disc herniation be ceased and surgery considered? *Neurosurg Rev.* 2002;25:162-165.
148. Asch HL, Lewis PJ, Moreland DB, et al. Prospective multiple outcomes study of outpatient lumbar microdiscectomy: Should 75 to 80% success rates be the norm? *Journal of Neurosurgery: Spine.* 2002;96:34-44.
149. Carragee EJ, Han MY, Suen PW, et al. Clinical outcomes after lumbar discectomy for sciatica: The effects of fragment type and anular competence. *J Bone Joint Surg Am.* 2003;85-A:102-108.
150. Ng LC, Sell P. Predictive value of the duration of sciatica for lumbar discectomy. A prospective cohort study. *J Bone Joint Surg Br.* 2004;86:546-549.
151. Kohlboeck G, Greimel KV, Piotrowski WP, et al. Prognosis of multifactorial outcome in lumbar discectomy: A prospective longitudinal study investigating patients with disc prolapse. *Clin J Pain.* 2004;20:455-461.
152. Lonne G, Solberg TK, Sjaavik K, et al. Recovery of muscle strength after microdiscectomy for lumbar disc herniation: A prospective cohort study with 1-year follow-up. *Eur Spine J.* 2011.epub ahead of print DOI 10.1007/s00586-011-2122-x
153. Radloff L. The CES-D scale: A self-report depression scale for research in the general population, applied psychological methods 1 (1977). *Applied Psychological Measurement.* 1977;3:385-401.
154. Solberg TK, Olsen JA, Ingebrigtsen T, et al. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J.* 2005;14:1000-1007.
155. Zigmond AS, Snaith R. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* 1983;67:361-370.
156. Kohlmann T, Raspe H. Die patientennache diagnostik von funktionseinschränkungen im alltag. *Psychomed.* 1994:621-627.
157. Greenough CG, Fraser RD. Assessment of outcome in patients with low-back pain. *Spine.* 1992;17:36-41.
158. Hanvik LJ. MMPi profiles in patients with low-back pain. *J Consult Psychol.* 1951;15:350-353.
159. Main CJ. The modified somatic perception questionnaire (MSPQ). *J Psychosom Res.* 1983;27:503-514.
160. Tunturi T, Kataja M, Keski-Nisula L, et al. Posterior fusion of the lumbosacral spine. evaluation of the operative results and the factors influencing them. *Acta Orthop Scand.* 1979;50:415-425.

161. Derogatis LR, Lipman RS, Covi L. SCL-90: An outpatient psychiatric rating scale--preliminary report. *Psychopharmacol Bull.* 1973;9:13-28.
162. Ashworth J, Konstantinou K, Dunn KM. Prognostic factors in non-surgically treated sciatica: A systematic review. *BMC Musculoskelet Disord.* 2011;12:208.
163. Atlas SJ, Deyo RA, van den Ancker M, et al. The Maine-Seattle back questionnaire: A 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: Results of a derivation and validation cohort analysis. *Spine.* 2003;28:1869-1876.
164. Vroomen PC, de Krom MC, Wilmink JT, et al. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. *J Neurol Neurosurg Psychiatry.* 2002;72:630-634.
165. Grovle L, Haugen AJ, Keller A, et al. Reliability, validity, and responsiveness of the Norwegian versions of the Maine-Seattle back questionnaire and the sciatica bothersomeness and frequency indices. *Spine.* 2008;33:2347-2353.
166. Deville WL, van der Windt DA, Dzaferagic A, et al. The test of Lasegue: Systematic review of the accuracy in diagnosing herniated discs. *Spine.* 2000;25:1140-1147.
167. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. report of the multicenter criteria committee. *Arthritis Rheum.* 1990;33:160-172.
168. Clark ME, Kori SH, Brockel J. Kinesiophobia and chronic pain: Psychometric characteristics and factor analysis of the Tampa scale. 1996. *American pain society annual meeting.*
169. Keller RB, Atlas SJ, Singer DE, et al. The Maine lumbar spine study, part I: Background and concepts. *Spine.* 1996;21:1769-1776.
170. Grotle M, Brox JI, Vollestad NK. Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: Methodological aspects of the Norwegian version. *J Rehabil Med.* 2006;38:346-353.
171. Loge JH, Kaasa S. Short form 36 (SF-36) health survey: Normative data from the general Norwegian population. *Scand J Soc Med.* 1998;26:250-258.
172. Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behav Sci.* 1974;19:1-15.
173. Sandanger I, Moum T, Ingebrigtsen G, et al. Concordance between symptom screening and diagnostic procedure: The Hopkins Symptom Checklist-25 and the composite international diagnostic interview I. *Soc Psychiatry Psychiatr Epidemiol.* 1998;33:345-354.
174. Eriksen HR, Ihlebaek C, Ursin H. A scoring system for subjective health complaints (SHC). *Scand J Public Health.* 1999;27:63-72.
175. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: What, why, and how? *BMJ.* 2009;338:b375.
176. Altman DG. *Practical Statistics for Medical Research.* London: Chapman & Hall/CRC, 1991.
177. Pallant J. *SPSS Survival Manual.* 2nd ed. Open university Press, 2005.
178. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42.
179. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135-160.
180. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther.* 2005;85:257-268.
181. Bland JM, Altman DG. Cronbach's alpha. *BMJ.* 1997;314:572.
182. Stratford PW, Riddle DL. Assessing sensitivity to change: Choosing the appropriate change coefficient. *Health Qual Life Outcomes.* 2005;3:23.
183. de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59:1033-1039.
184. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29-36.
185. O'brien RM. A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity.* 2007;41:673-690.

186. DeCoster J. Overview of factor analysis 1998. Retrieved October 10 2011 from [www.stat-help.com/notes.html](http://www.stat-help.com/notes.html)
187. Harvill LM. Standard error of measurement. *Educational Measurement: Issues and Practice*. 1991;10:33-41.
188. Cortina JM. What is coefficient alpha? an examination of theory and applications. *J Appl Psychol*. 1993;78:98-104.
189. Bland JM, Altman DG. Statistics notes: Validating scales and indexes. *BMJ* 2002;324:606-607.
190. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine & Rehabilitation*. 2000;81:S15-20.
191. Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. *Qual Life Res*. 2003;12:349-362.
192. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Second ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., 1988.
193. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25:3186-3191.
194. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342:1878-1886.
195. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *BMJ*. 2007;335:806-808.
196. MacMahon S, Collins R. Reliable assessment of the effects of treatment on mortality and major morbidity, II: Observational studies. *Lancet*. 2001;357:455-462.
197. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342:1887-1892.
198. Solberg TK, Sorlie A, Sjaavik K, et al. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthop*. 2011;82:56-63.
199. Feinstein AR. Epidemiologic analyses of causation: The unlearned scientific lessons of randomized trials. *Journal of clinical epidemiology* 1989;42:481-489.
200. Peul WC, van Houwelingen HC, van der Hout WB, et al. Prolonged conservative treatment or 'early' surgery in sciatica caused by a lumbar disc herniation: Rationale and design of a randomized trial [ISRCT 26872154]. *BMC Musculoskelet Disord*. 2005;6:8.
201. Genevay S, Atlas SJ, Katz JN. Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: A structured literature review. *Spine*. 2010;35:803-811.
202. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: An American College of Rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis*. 2010;69:1580-1588.
203. Birkmeyer NJ, Weinstein JN, Tosteson AN, et al. Design of the Spine Patient Outcomes Research Trial (SPORT). *Spine*. 2002;27:1361-1372.
204. 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;28:1913-1921.
205. Brage S, Sandanger I, Nygard JF. Emotional distress as a predictor for low back disability: A prospective 12-year population-based study. *Spine*. 2007;32:269-274.
206. Storheim K, Brox JI, Holm I, et al. Predictors of return to work in patients sick listed for sub-acute low back pain: A 12-month follow-up study. *J Rehabil Med*. 2005;37:365-371.
207. Hellum C, Johnsen LG, Storheim K, et al. Surgery with disc prosthesis versus rehabilitation in patients with low back pain and degenerative disc: Two year follow-up of randomised study. *BMJ*. 2011;342:d2786.
208. Ihlebaek C, Brage S, Eriksen HR. Health complaints and sickness absence in Norway, 1996-2003. *Occup Med (Lond)*. 2007;57:43-49.



209. Ihlebaek CM, Ødegaard A, Vikne J, et al. Subjective health complaints in patients with chronic whiplash associated disorders (WAD). relationships with physical, psychological, and collision associated factors. *Nor J Epidemiol.* 2006;16:119-126.
210. Hagen EM, Svensen E, Eriksen HR, et al. Comorbid subjective health complaints in low back pain. *Spine.* 2006;31:1491-1495.
211. Gøtestam KG, Svebak S, Jensen EN. The role of personality, mood, subjective health, and stress in depressive symptoms among high school students. *Eur.J.Psychiat.* 2008;22:121-129.
212. Buitenhuis J, Jaspers JP, Fidler V. Can kinesiophobia predict the duration of neck symptoms in acute whiplash? *Clin J Pain.* 2006;22:272-277.
213. Sieben JM, Vlaeyen JW, Portegijs PJ, et al. A longitudinal study on the predictive validity of the fear-avoidance model in low back pain. *Pain.* 2005;117:162-170.
214. Grotle M, Foster NE, Dunn KM, et al. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain.* 2010;151:790-797.
215. Miranda H, Viikari-Juntura E, Martikainen R, et al. Individual factors, occupational loading, and physical exercise as predictors of sciatic pain. *Spine.* 2002;27:1102-1109.
216. Shiri R, Karppinen J, Leino-Arjas P, et al. Cardiovascular and lifestyle risk factors in lumbar radicular pain or clinically defined sciatica: A systematic review. *Eur Spine J.* 2007;16:2043-2054.
217. Schade V, Semmer N, Main CJ, et al. The impact of clinical, morphological, psychosocial and work-related factors on the outcome of lumbar discectomy. *Pain.* 1999;80:239-249.
218. Silverplatt K, Lind B, Zoega B, et al. Clinical factors of importance for outcome after lumbar disc herniation surgery: Long-term follow-up. *Eur Spine J.* 2010;19:1459-1467.
219. Aina A, May S, Clare H. The centralization phenomenon of spinal symptoms--a systematic review. *Man Ther.* 2004;9:134-143.
220. Skytte L, May S, Petersen P. Centralization: Its prognostic value in patients with referred symptoms and sciatica. *Spine.* 2005;30:E293-9.
221. Fayers PM, Machin D. *Quality of Life : Assessment, Analysis and Interpretation.* Chischester: John Wiley, 2000.
222. Chavance M. Handling missing items in quality of life studies. *Commun Statist -Theory Meth.* 2004;33:1371-1383.
223. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25:3186-3191.
224. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46:1417-1432.
225. Ware JE, Jr. SF-36 health survey update. *Spine.* 2000;25:3130-3139.
226. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-310.
227. Swinkels-Meewisse EJ, Swinkels RA, Verbeek AL, et al. Psychometric properties of the Tampa scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Man Ther.* 2003;8:29-36.
228. de Souza FS, Marinho Cda S, Siqueira FB, et al. Psychometric testing confirms that the Brazilian-Portuguese adaptations, the original versions of the fear-avoidance beliefs questionnaire, and the Tampa scale of kinesiophobia have similar measurement properties. *Spine.* 2008;33:1028-1033.
229. Lundberg MKE, Styf J, Carlsson SG. A psychometric evaluation of the Tampa scale for kinesiophobia - from a physiotherapeutic perspective. *Physiotherapy Theory & Practice.* 2004;20:121-133.
230. Bunketorp L, Carlsson J, Kowalski J, et al. Evaluating the reliability of multi-item scales: A non-parametric approach to the ordered categorical structure of data collected with the Swedish version of the Tampa scale for kinesiophobia and the self-efficacy scale. *J Rehabil Med.* 2005;37:330-334.
231. Cleland JA, Fritz JM, Childs JD. Psychometric properties of the fear-avoidance beliefs questionnaire and Tampa scale of kinesiophobia in patients with neck pain. *Am J Phys Med Rehabil.* 2008;87:109-117.

232. Mintken PE, Cleland JA, Whitman JM, et al. Psychometric properties of the fear-avoidance beliefs questionnaire and Tampa scale of kinesiophobia in patients with shoulder pain. *Arch Phys Med Rehabil*. 2010;91:1128-1136.
233. Monticone M, Giorgi I, Baiardi P, et al. Development of the Italian version of the Tampa scale of kinesiophobia (TSK-I): Cross-cultural adaptation, factor analysis, reliability, and validity. *Spine*. 2010;35:1241-1246.
234. Woby SR, Roach NK, Urmston M, et al. Psychometric properties of the TSK-11: A shortened version of the Tampa scale for kinesiophobia. *Pain*. 2005;117:137-144.
235. Svensson E. Guidelines to statistical evaluation of data from rating scales and questionnaires. *J Rehabil Med*. 2001;33:47-48.
236. Ostelo RW, Swinkels-Meewisse IJ, Knol DL, et al. Assessing pain and pain-related fear in acute low back pain: What is the smallest detectable change? *Int J Behav Med*. 2007;14:242-248.
237. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: A critical review and recommendations. *J Clin Epidemiol*. 2000;53:459-468.
238. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol*. 2010;10:22.
239. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737-745.
240. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: A comparison of different instruments. *Pain*. 1996;65:71-76.
241. Lauridsen HH, Manniche C, Korsholm L, et al. What is an acceptable outcome of treatment before it begins? methodological considerations and implications for patients with chronic low back pain. *Eur Spine J*. 2009;18:1858-1866.
242. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113:9-19.
243. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: The lesson of cronbach. *J Clin Epidemiol*. 1997;50:869-879.
244. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: A review of strengths and weaknesses and considerations for design. *The Journal of manual & manipulative therapy* 2010;17:163-170.
245. Ong BN, Hooper H, Jinks C, et al. 'I suppose that depends on how I was feeling at the time': Perspectives on questionnaires measuring quality of life and musculoskeletal pain. *Journal of health services research & policy* 2006;11:81-88.
246. 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;27:1213-1222.
247. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: A choice of methods using the Oswestry Disability Index, Medical Outcomes Study Questionnaire Short Form 36, and pain scales. *Spine J*. 2008;8:968-974.
248. Mannion AF, Porchet F, Kleinstuck FS, et al. The quality of spine surgery from the patient's perspective: Part 2. minimal clinically important difference for improvement and deterioration as measured with the core outcome measures index. *Eur Spine J*. 2009;18 Suppl 3:374-379.
249. Gatchel RJ, Lurie JD, Mayer TG. Minimal clinically important difference. *Spine*. 2010;35:1739-1743.
250. de Vet HC, Terwee CB, Ostelo RW, et al. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*. 2006;4:54.
251. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56:395-407.

252. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine*. 2008;33:90-94.
253. Farrar JT, Young JP, Jr, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94:149-158.
254. Spratt KF. Patient-level minimal clinically important difference based on clinical judgment and minimally detectable measurement difference: A rationale for the SF-36 physical function scale in the SPORT intervertebral disc herniation cohort. *Spine*. 2009;34:1722-1731.
255. Damsgard E, Fors T, Anke A, et al. The Tampa scale of kinesiophobia: A Rasch analysis of its properties in subjects with low back and more widespread pain. *J Rehabil Med*. 2008;39:672-678.
256. Goubert L, Crombez G, Van Damme S, et al. Confirmatory factor analysis of the Tampa scale for kinesiophobia: Invariant two-factor model across low back pain patients and fibromyalgia patients. *Clin J Pain*. 2004;20:103-110.
257. Roelofs J, Goubert L, Peters ML, et al. The Tampa scale for kinesiophobia: Further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *Eur J Pain*. 2004;8:495-502.
258. Heuts PH, Vlaeyen JW, Roelofs J, et al. Pain-related fear and daily functioning in patients with osteoarthritis. *Pain*. 2004;110:228-235.
259. Houben RMA, Leeuw M, Vlaeyen JWS, et al. Fear of movement/injury in the general population: Factor structure and psychometric properties of an adapted version of the Tampa scale for kinesiophobia. *J Behav Med*. 2005;28:415-424.
260. Gomez-Perez L, Lopez-Martinez AE, Ruiz-Parraga GT. Psychometric properties of the Spanish version of the Tampa scale for kinesiophobia (TSK). *J Pain*. 2011;12:425-435.
261. Wong WS, Kwok HY, Luk KDK, et al. Fear of movement/(re) injury in chinese patients with chronic pain: Factorial validity of the Chinese version of the Tampa scale for kinesiophobia. *J Rehabil Med*. 2010;42:620-629.
262. French DJ, Roach PJ, Mayes S. Peur du mouvement chez des accidentés du travail: L'échelle de kinésiophobie de Tampa (EKT). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*. 2002;34:28-33.
263. Geisser ME, Haig AJ, Theisen ME. Activity avoidance and function in persons with chronic back pain. *J Occup Rehabil*. 2000;10:215-227.
264. Cook AJ, Brawer PA, Vowles KE. The fear-avoidance model of chronic pain: Validation and age analysis using structural equation modeling. *Pain*. 2006;121:195-206.
265. George SZ, Lentz TA, Zeppieri G, et al. Analysis of shortened versions of the Tampa scale for kinesiophobia and pain catastrophizing scale for patients after anterior cruciate ligament reconstruction. *Clin J Pain*. 2011 Epub ahead of print.
266. American Psychiatric Association, American Psychiatric Association. Task Force on DSM-IV. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. American Psychiatric Publishing, Inc., 2000.
267. Vlaeyen JWS, De Jong J, Sieben J, et al. Graded exposure in vivo for pain-related fear. In: Turk DC, Gatchel RJ, eds. *Psychological Approaches to Pain Management: A Practitioner's Handbook*. 2nd ed. New York: The Guilford Press, 2002:210-33.
268. <http://en.wikipedia.org/wiki/Success>.
269. <http://en.wikipedia.org/wiki/Recovery>.
270. Mannion AF, Porchet F, Kleinstuck FS, et al. The quality of spine surgery from the patient's perspective. part 1: The core outcome measures index in clinical practice. *Eur Spine J*. 2009;18 Suppl 3:367-373.
271. Smolen JS, Aletaha D. Monitoring rheumatoid arthritis. *Curr Opin Rheumatol*. 2011;23:252-258.
272. Simon LS, Evans C, Katz N, et al. Preliminary development of a responder index for chronic low back pain. *J Rheumatol*. 2007;34:1386-1391.

273. Hestbaek L, Leboeuf-Yde C, Manniche C. Is low back pain part of a general health pattern or is it a separate and distinctive entity? A critical literature review of comorbidity with low back pain. *J Manipulative Physiol Ther.* 2003;26:243-252.
274. Von Korff M, Crane P, Lane M, et al. Chronic spinal pain and physical-mental comorbidity in the united states: Results from the national comorbidity survey replication. *Pain.* 2005;113:331-339.
275. Grovle L, Haugen AJ, Ihlebaek CM, et al. Comorbid subjective health complaints in patients with sciatica: A prospective study including comparison with the general population. *J Psychosom Res.* 2011;70:548-556.
276. Heliovaara M, Makela M, Knekt P, et al. Determinants of sciatica and low-back pain. *Spine.* 1991;16:608-614.
277. Linton SJ. A review of psychological risk factors in back and neck pain. *Spine.* 2000;25:1148-1156.
278. Chou R, Baisden J, Carragee EJ, et al. Surgery for low back pain: A review of the evidence for an American pain society clinical practice guideline. *Spine.* 2009;34:1094-1109.
279. Dandy WE. Loose cartilage from intervertebral disk simulating tumor of the spinal cord. *Archives of Surgery.* 1929;19:660-672

## Errataliste

- side 10 linje 5 Dandy<sup>4</sup> er endret til Dandy<sup>279</sup>
- side 17 linje 22 cauda equine rettet til cauda equina
- side 18 linje 30 patients rettet til patients'
- side 34 linje 6 referanse 79 er nevnt to ganger, den ene er fjernet
- side 71 Referanse 223. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. *Spine*. 2000;25:3192-3199 er erstattet med: Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25:3186-3191.
- side 74 Tillegg av referanse:  
279. Dandy WE. Loose cartilage from intervertebral disk simulating tumor of the spinal cord. *Archives of Surgery*. 1929;19:660-672.



# Papers

Paper I-III













# Estimates of success in patients with sciatica due to lumbar disc herniation depend upon outcome measure

Anne Julsrud Haugen · Lars Grøvle ·  
Jens Ivar Brox · Bård Natvig · Anne Keller ·  
Dag Soldal · Margreth Grotle

Received: 26 October 2010 / Revised: 9 February 2011 / Accepted: 12 April 2011 / Published online: 24 April 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** The objectives were to estimate the cut-off points for success on different sciatica outcome measures and to determine the success rate after an episode of sciatica by using these cut-offs. A 12-month multicenter observational study was conducted on 466 patients with sciatica and lumbar disc herniation. The cut-off values were estimated by ROC curve analyses using *Completely recovered* or *Much better* on a 7-point global change scale as external criterion for success. The cut-off values

(references in brackets) at 12 months were leg pain VAS 17.5 (0–100), back pain VAS 22.5 (0–100), Sciatica Bothersomeness Index 6.5 (0–24), Maine-Seattle Back Questionnaire 4.5 (0–12), and the SF-36 subscales bodily pain 51.5, and physical functioning 81.7 (0–100, higher values indicate better health). In conclusion, the success rates at 12 months varied from 49 to 58% depending on the measure used. The proposed cut-offs may facilitate the comparison of success rates across studies.

**Keywords** Sciatica · Outcome measures · Pain · Disability

A. J. Haugen (✉) · L. Grøvle  
Department of Rheumatology, Østfold Hospital Trust,  
1603 Fredrikstad, Norway  
e-mail: annhau@so-hf.no

J. I. Brox  
Section for Back Surgery and Physical Medicine and  
Rehabilitation, Orthopaedic Department,  
Oslo University Hospital, Rikshospitalet, Oslo, Norway

B. Natvig · M. Grotle  
National Resource Centre for Rehabilitation in Rheumatology,  
Diakonhjemmet Hospital, Oslo, Norway

A. Keller  
Department of Physical Medicine and Rehabilitation,  
Oslo University Hospital, Ullevaal, Oslo, Norway

D. Soldal  
Department of Rheumatology, Sørlandet Hospital Health  
Enterprise, Kristiansand, Norway

M. Grotle  
FORMI (Communication Unit for Musculoskeletal Disorders),  
Oslo University Hospital, Ullevaal, Oslo, Norway

B. Natvig  
Department of General Practice and Community Medicine,  
ASAM, University of Oslo, Oslo, Norway

## Introduction

Sciatica, defined as nerve root pain—or radicular pain in the leg below the knee, is most commonly caused by lumbar disc herniation. The natural course of sciatica is often described as favourable, and it is presumed that the symptoms of the majority of patients improve [11, 18, 22]. A recent 1-year follow-up study of patients with sciatica randomised to early surgery or to prolonged conservative treatment found recovery rates of 95% for both groups [17]. On the other hand, in the SPORT study, only 44.7% of the patients who did not undergo surgery and 77.1% of the surgically treated patients were satisfied with their symptoms after 1 year [23]. These divergent results on sciatica patients are challenging [4], and the different outcomes available make it difficult to compare the results of published studies. Primary patient-reported outcomes commonly used in clinical studies on sciatica are patient-rated global change [1, 8, 14, 17, 23], leg pain [15, 17, 22], back pain [22], Medical Outcomes Study Short Form 36 (SF-36) bodily pain and physical functioning [23], North

American Spine Society (NASS) neurogenic symptoms score (NSS) [20], Oswestry Disability Index (ODI) [23] and the Roland Morris disability questionnaire (RMDQ) [17, 22]. Few studies have used sciatica-specific outcomes such as the Sciatica Bothersomeness and Frequency Indices, and then only as secondary outcomes [1, 17, 23].

Clinical findings and imaging have shown little relevance for change of symptoms in patients with sciatica. Therefore, patient's self-reports are important. The minimum clinically important difference (MCID) represents the smallest improvement or change in score based on an external anchoring question about the patient's perceived improvement [5, 13]. However, although there is a large improvement on a measure, the patient might not have had a good outcome or success. A cut-off differing between success and non-success makes it possible to estimate the rate of patients with a good outcome. As long as no consensus exists on the definition of success, there is no agreement on the cut-off values for success for various outcome measures of sciatica. A 7-point Likert scale has been used as a recovery scale (complete or nearly complete disappearance of sciatica symptoms) [17] and as a global perceived effect scale (completely recovered and much improved) [12] in patients with sciatica.

The aims of this study were to estimate the cut-off points for success on different outcome measures and to determine the success rate after an episode of sciatica using the outcome cut-off values.

## Materials and methods

### Design

We conducted a prospective multicentre observational study of patients with sciatica. The sample size was calculated based on an expectation that surgical treatment would be necessary for 30% of the patients and that 70% of those who were surgically treated and 50% of those who were not surgically treated would experience a good outcome [1]. Power analysis indicated a sample size on 300 patients requiring 90% power with a two-sided exact Fisher test of 5% significance level. Taking loss-to-follow-up into consideration we ended up with a study sample of 400 patients. The study protocol was approved by the Regional Committee for Medical Research Ethics and The Ombudsman for Privacy in Research at the Norwegian Social Science Data Services.

### Patients

Patients were recruited from specialty back clinics at four hospitals in southeast Norway (Sykehuset Østfold,

Sørlandet Sykehus, Oslo Universitetssykehus Ullevål and Sykehuset Innlandet). Patients were referred to the specialty back clinics from the primary health care service. Assessment and treatment were conducted as usual in the back clinics. A physician or physiotherapist informed consecutive eligible patients about the study. Inclusion criteria were age  $\geq 18$  years and radiating pain below the knee level, and/or paresis and a disc herniation at the corresponding level and side according to magnetic resonance imaging (MRI) or computed tomography (CT). Exclusion criteria were prior surgery at the same disc level, fracture, infection, malignancy, pregnancy, and lack of fluency in Norwegian. All patients signed an informed consent form after receiving oral and written information.

### Procedure

At the first visit, a questionnaire on sociodemographic factors (age, gender, education length, smoking status and work status) was completed. Patients also reported their history of previous sciatica and the duration of relevant sciatica episodes. A clinical examination was carried out by trained physiotherapists or physicians at the departments of rheumatology, physical medicine and rehabilitation, or orthopaedics.

The assessment included motor function of the muscles of the knee, leg and toe, the Trendelenburg test, sensory loss in the leg, reflexes of the Achilles tendon and patella (all deemed abnormal if reduced) and the straight leg raising test (deemed abnormal if  $<60^\circ$ ). Number of tender-points were recorded. Standardisation of testing procedures was done in meetings with the participating centres.

A follow-up questionnaire and a prepaid envelope were sent to the patients after 3, 6 and 12 months. The date was recorded for patients who underwent back surgery. A study nurse sent a reminder after 2 weeks if no reply was obtained.

### Measures

Leg pain and low back pain were measured using the visual analogue scale (VAS) 0–100. The Sciatica Bothersomeness Index (SBI), which is a composite score of four questions (each score ranging from 0–6) that include elements of leg pain and sensory and motor disturbances, was used [7, 16]. Total score ranges from 0 to 24 and higher scores indicate worse symptoms. The Maine-Seattle Back Questionnaire (MSBQ) [2, 7] which was developed from the RMDQ and modified for sciatica patients, was used to measure functional status. The scale is composed of 12 items and the score ranges from 0 to 12, higher scores indicating greater disability. The MSBQ mainly assesses disability, but also includes questions about pain. The Norwegian versions of

the SBI and the MSBQ were recently validated [7]. The bodily pain and physical functioning subscales of SF-36 [21] are parts of a generic quality of life questionnaire in widespread use, including studies on sciatica [1, 23, 24]. Scores range from 0 to 100 and higher values indicate better health.

The follow-up questionnaires included a 7-item global change scale for leg and back pain with the following categories: *Completely recovered*, *Much better*, *Better*, *A little better*, *No change*, *A little worse* and *Much worse*. Patients were categorised as success if they responded with *Completely recovered* or *Much better*. Satisfaction to spend the rest of life in the current state was registered using a 5-item Likert scale with the wording *Very satisfied*, *A little satisfied*, *Neither satisfied nor not*, *A little dissatisfied*, and *Very Dissatisfied*.

#### Statistical analyses

Baseline and follow-up comparisons between groups were analysed using an independent *t*-test for continuous data and  $\chi^2$  for categorical data. Floor or ceiling effects were defined as >15% of scores in the lowest or highest parts of the scales, respectively [19].

A receiver operating characteristic (ROC) curve was obtained by plotting every possible cut-off score's sensitivity on the *y*-axis against 1-specificity on the *x*-axis. Sensitivity was defined as the proportion of patients who were correctly classified in the success group (*Completely recovered* or *Much better*). Specificity was defined as the proportion of patients who were correctly classified in the non-success group. The value on the curve closest to the upper left corner shows the highest sensitivity and specificity and was considered the best cut-off score for distinguishing between success and non-success. The area under the ROC curve (AUC) reflects the scoring system's ability in differentiating between success and non-success. An AUC value >0.70 was considered satisfactory [19].

McNemar's test was used when comparing sensitivity of two diagnostic measures where the cut-off point was set to specificity 80% for each measure.

The Statistical Package for Social Sciences Version 17.0 (SPSS Inc., Chicago, IL) was used for data analysis.

## Results

Four hundred sixty-six patients were included in the study; 268 (57.5%) were men. The mean age was  $43.6 \pm 11.5$  years. Patient characteristics are presented in Table 1. Fifty-seven (12.2%) patients, six of whom underwent surgery, were lost to follow-up at 1 year. At baseline these patients were significantly younger, had a higher leg pain

score and more often tested positive on the straight leg raising test than patients who completed the study (Table 1).

Patients were classified as *acute* if duration of current sciatica episode was less than 3 months and *chronic* if duration was 3 months and more. The acute patients ( $n = 192$ , 41.4%) represented more males, higher score on SBI, lower scores on SF-36 bodily pain and physical functioning, and higher frequency of reduced motor function, sensory loss and reflex depletion at the clinical examination. Baseline leg pain intensity score was >90 in 100 patients. The SF-36 bodily pain had a ceiling effect as 99 (21.2%) patients reported a score of 0 (the worst degree of pain).

The 3-, 6- and 12-month questionnaires were answered by 434 (93.1%), 423 (90.8%) and 409 (87.8%) patients, respectively. Mean leg pain score decreased from  $63.2 \pm 28.2$  at baseline to  $26.7 \pm 28.8$  at 12 months. Over the same period, the mean SBI score decreased from  $14.2 \pm 5.0$  to  $7.2 \pm 6.1$  and the MSBQ score decreased from  $8.1 \pm 2.6$  to  $4.2 \pm 3.4$ . Those who worked full time were 93 (20.0%) at baseline, 183 (42.4%) at 3 months, 220 (52.0%) at 6 months and 250 (61.1%) at 12 months.

By the 12-month follow-up, 126 patients had received surgical treatment. Patients selected for surgery were younger, had more frequently positive straight leg raising test and reported significantly more pain and disability at baseline than those who did not undergo surgery. At 12 months, the surgically treated patients had significantly better outcomes in terms of leg pain ( $P = 0.001$ ) and SBI ( $P < 0.001$ ) than non-surgically treated patients, whereas outcomes such as back pain, MSBQ and SF-36 bodily pain and physical functioning scores did not differ between surgically and non-surgically treated patients (Fig. 1).

Table 2 presents the mean scores for the outcome variables at 12 months according to the global change scale. All outcome scores increased correspondingly with the categories in the global change scale, except the items *No change*, *A little worse* and *Much worse* for SBI and VAS in leg and back. Using the global change scale 222 (54.4%) of the patients reported success; 83 were of the surgically treated group (69.2%) and 139 were of the non-surgically treated group (48.3%). Overall, 188 (46.1%) of the patients indicated they were satisfied to spend the rest of their lives in their current state; 72 were of the surgically treated group (60.0%) and 116 were of the non-surgically treated group (40.4%).

The ROC analyses showed a high AUC (0.76–0.89) for leg and back pain scores, SBI, MSBQ and the SF-36 bodily pain and physical functioning scores at 3, 6 and 12 months. Similar results were observed using satisfied to spend the rest of life in the current state as an external criterion (AUC 0.74–0.87). Cut-off points with the highest sensitivity and

**Table 1** Baseline characteristics for 409 patients with complete data and 57 non-responders at the 1-year follow-up

	Total cohort n = 409	Non-responders n = 57	P
Age years, mean (SD)	44.2 (11.4)	39.2 (10.9)	0.002
Males	232 (56.7)	36 (63.2)	0.357
Current smoker	171 (42.1)	29 (51.8)	0.171
Education >12 years	197 (49.6)	30 (53.6)	0.580
Working status			0.440
Working full time	83 (20.3)	10 (17.5)	
Partly sick leave	48 (11.7)	6 (10.5)	
Total sick leave	200 (48.9)	27 (47.4)	
Rehabilitation, disability pension	45 (11.0)	8 (14.1)	
Other	33 (8.1)	6 (10.4)	
Previous disc surgery	17 (4.2)	1 (1.8)	0.374
First sciatica episode	183 (44.9)	27 (48.2)	0.636
Duration current sciatica episode			0.586
<3 months	172 (42.3)	20 (35.1)	
3–6 months	135 (33.2)	21 (36.8)	
>6 months	100 (24.6)	16 (28.1)	
VAS low back pain (0–100), mean (SD)	41.4 (30.0)	51.6 (28.6)	0.015
VAS leg pain (0–100), mean (SD)	62.6 (28.1)	68.0 (28.1)	0.180
SBI (0–24), mean (SD)	14.2 (5.0)	14.5 (5.2)	0.740
MSBQ (0–12), mean (SD)	8.1 (2.6)	8.8 (2.5)	0.050
SF-36 (0–100 <sup>a</sup> ) bodily pain, mean (SD)	22.9 (17.7)	21.2 (17.4)	0.480
SF-36 (0–100 <sup>a</sup> ) physical functioning, mean (SD)	50.4 (25.4)	45.4 (25.7)	0.172
Clinical examination			
Straight leg raising <60°	224 (55.6)	43 (75.4)	0.004
Sensory loss	236 (58.1)	37 (64.9)	0.330
Muscle weakness	177 (44.1)	26 (47.3)	0.661
Reflex weakness	190 (47.0)	22 (40.0)	0.327
>10 tenderpoints	34 (8.9)	5 (9.4)	0.894

Data are presented as number (percentage) unless otherwise indicated

VAS Visual Analogue Scale, SBI Sciatica Bothersomeness Index, MSBQ Maine-Seattle Back Questionnaire, SF-36 Medical Outcomes Study Short-Form Health Survey

<sup>a</sup> Higher score indicates better health

specificity that distinguished between patients with and without success of sciatica at 12 months were 17.5 for leg pain score, 22.5 for back pain score, 6.5 for SBI score and 4.5 for MSBQ score at 12 months. The cut-off points for the SF-36 bodily pain and physical functioning scores at 12 months were 51.5 and 81.7, respectively. Table 3 presents the cut-off values at 3 and 12 months for the total cohort and for the subgroups of gender, duration of symptoms and surgical status. The proportion of patients achieving success according to the calculated cut-off values at 12 months varied from 49% (SF-36 bodily pain) to 58% (MSBQ) (Fig. 2).

With a fixed specificity on 80%, the sensitivity for the outcome measures varied from 69% (leg pain VAS) to 86% (MSBQ). The MSBQ showed a significantly higher sensitivity (86%) with the locked specificity at 80% compared with all the other outcome measures ( $P \leq 0.013$ ). There were no significant differences in sensitivity comparing leg and back pain scores, SBI and the SF-36 bodily pain and physical functioning scores where each measure's cut-off point was set to specificity 80%.

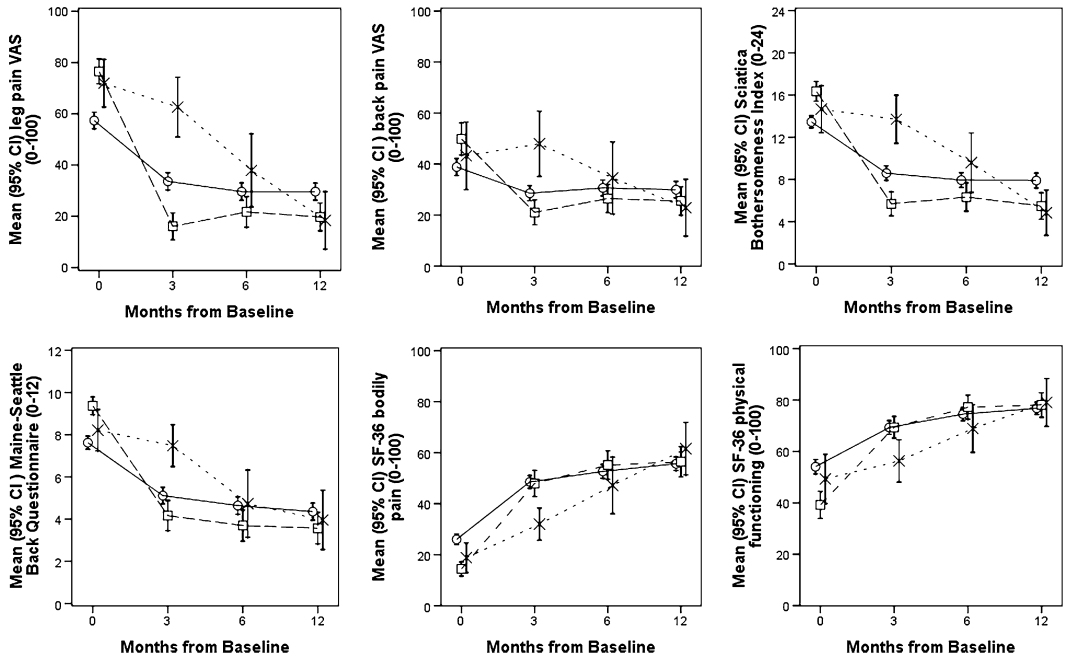
## Discussion

The main finding of this sciatic cohort study was that the success rates among the different outcome measures at 1 year varied from 49 to 58%. This indicates that success rates cannot be used for comparison between studies unless the success criterion is exactly defined.

We used a strict definition of success using *Completely recovered* or *Much better* as the success criterion. In the non-success group, there were patients with both deterioration and some improvement of the symptoms. After 1 year, 112 (27.5%) patients reported their condition as *Better* or *A little better*, but only 14 (12.5%) of them were satisfied with their symptoms.

The ROC cut-off was constructed using the 12-month absolute score for each variable and *Completely recovered* or *Much better* as the external criterion. When the same external criterion was used for the 12-month change scores, the results were similar (AUC 0.72–0.87). We decided to use cut-off values for the absolute follow-up scores in stead





**Fig. 1** Time course for scores for leg pain (VAS), back pain (VAS), Sciatica Bothersomeness Index, Maine-Seattle Back Questionnaire, SF-36 bodily pain and SF-36 physical functioning (means with 95% CIs) in patients treated without surgery ( $n = 289$ ) (circles), surgery within 3 months of registration ( $n = 98$ ) (squares) and surgery between 3 and 12 months of registration ( $n = 28$ ) (multi symbols)

**Table 2** Mean (SD) scores at 12 months for leg pain (VAS), back pain (VAS), Sciatica Bothersomeness Index (SBI), Maine-Seattle Back Questionnaire (MSBQ), SF-36 bodily pain, SF-36 physical functioning registered at every category of the 7-point global change scale

Global change of sciatica symptoms	$n^a$	Leg pain (0–100)	Back pain (0–100)	Sciatica Bothersomeness Index (0–24)	Maine-Seattle Back Questionnaire (0–12)	SF-36 bodily pain (0–100) <sup>b</sup>	SF-36 physical functioning (0–100) <sup>b</sup>
Completely recovered	61	1.9 (6.0)	6.3 (18.1)	0.9 (2.0)	0.7 (1.5)	80.6 (22.9)	92.4 (11.3)
Much better	161	14.4 (16.7)	16.1 (15.0)	4.6 (3.6)	2.6 (2.1)	64.9 (18.3)	85.6 (14.2)
Better	59	28.9 (22.5)	30.6 (22.1)	8.7 (5.2)	5.2 (3.0)	50.5 (18.0)	72.9 (19.8)
A little better	53	40.0 (27.6)	49.8 (25.9)	10.3 (4.3)	6.5 (3.0)	41.3 (16.8)	65.2 (22.1)
No change	44	62.9 (27.3)	55.3 (28.3)	14.6 (5.1)	7.7 (2.2)	36.5 (19.0)	60.8 (20.3)
A little worse	20	52.5 (28.3)	61.0 (21.4)	13.3 (5.3)	8.2 (1.7)	29.7 (18.3)	59.4 (22.1)
Much worse	10	82.1 (14.6)	55.9 (32.5)	18.3 (5.2)	9.3 (2.0)	21.5 (16.0)	44.0 (25.1)

<sup>a</sup> 408 completed the questionnaire

<sup>b</sup> Higher values indicate better health

of change scores in order to enhance feasibility. Cut-off values of follow-up scores can easily be used in clinical practice as a guideline for evaluating outcome of sciatica.

In the ROC-cut-off analyses all outcomes showed satisfactory ability to discriminate between success and non-success. However, the scores for the MSBQ at 12 months had the highest AUCs and a significantly higher sensitivity

when specificity was locked at 80% and might be the most preferable score in evaluating pain and function during sciatica. Over time, the SBI and the MSBQ showed the most stable cut-off points.

Patients selected for surgery had a higher baseline score and a more rapid decline in symptoms than those who did not undergo surgery. However, the outcomes at 1-year

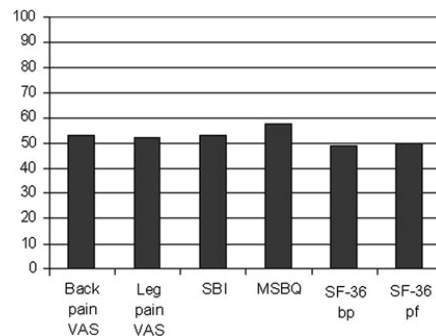
**Table 3** Cut-off values with sensitivity and specificity using “Completely recovered or Much better” as external criterion

	3 months		12 months	
	Cut-off value	Sens, spec	Cut-off value	Sens, spec
<b>VAS leg (0–100)</b>				
Total population	20.5	78, 79	17.5	78, 77
Acute	19.5	75, 67	18.5	78, 74
Chronic	21.5	84, 85	16.5	76, 79
Surgery	11.5	80, 80	23.0	91, 81
No surgery	22.5	75, 76	18.5	73, 75
Male	19.5	76, 74	13.5	75, 75
Female	21.5	83, 85	24.0	86, 83
<b>VAS back (0–100)</b>				
Total population	19.5	71, 72	22.5	79, 78
Acute	20.5	77, 73	23.5	76, 77
Chronic	19.5	70, 70	19.5	78, 80
Surgery	28.5	89, 65	24.5	82, 84
No surgery	20.5	75, 70	22.5	78, 77
Male	20.5	73, 70	21.5	79, 77
Female	19.5	74, 74	25.5	80, 81
<b>Sciatica Bothersomeness Index (0–24)</b>				
Total population	6.5	77, 81	6.5	78, 75
Acute	7.5	78, 75	6.5	76, 73
Chronic	6.5	82, 82	6.5	80, 76
Surgery	6.5	82, 80	5.5	86, 87
No surgery	7.5	80, 78	6.5	70, 73
Male	6.5	76, 76	5.5	80, 87
Female	8.5	84, 83	6.5	72, 81
<b>Maine-Seattle Back Questionnaire (0–12)</b>				
Total population	4.5	74, 81	4.5	86, 78
Acute	4.5	71, 77	3.5	76, 85
Chronic	4.5	76, 83	4.5	87, 81
Surgery	5.5	79, 85	4.5	87, 84
No surgery	4.5	75, 81	3.5	77, 83
Male	4.5	77, 79	3.5	82, 81
Female	5.5	77, 76	4.5	87, 88
<b>SF-36 bodily pain (0–100)</b>				
Total population	41.5	71, 72	51.5	74, 80
Acute	41.5	70, 70	51.5	74, 74
Chronic	41.5	70, 73	56.5	74, 85
Surgery	36.5	79, 75	46.5	82, 71
No surgery	46.5	75, 73	61.5	74, 84
Male	46.5	70, 67	56.5	77, 76
Female	41.5	68, 82	46.5	81, 73
<b>SF-36 physical functioning (0–100)</b>				
Total population	72.5	69, 68	81.7	75, 78
Acute	72.5	68, 63	81.7	75, 74
Chronic	72.5	70, 71	82.5	75, 80
Surgery	67.5	73, 80	82.5	73, 84

**Table 3** continued

	3 months		12 months	
	Cut-off value	Sens, spec	Cut-off value	Sens, spec
No surgery	72.5	72, 66	81.7	77, 76
Male	77.5	66, 68	81.7	85, 74
Female	67.5	78, 65	77.5	75, 78

Cut-off values are presented at 3- and 12-month follow-up for the total population and the subgroups: sciatica duration <3 months at inclusion (acute), sciatica duration  $\geq 3$  months at inclusion (chronic), surgery registered at 3- respectively, 12-month follow-up, no surgery, male and female



**Fig. 2** Per cent of patients achieving success according to calculated cut-off values for back and leg pain VAS, Sciatica Bothersomeness Index (SBI), Maine-Seattle Back Questionnaire (MSBQ), and SF-36 bodily pain (bp) and physical functioning (pf)

follow-up were similar, which is in agreement with the Cochrane Review and other observational studies [1, 6, 23].

The global change scores correspond to the results of the MLSS study [1]. This shows that many patients still experience pain and disability after 1 year.

Definition of recovery in terms of a cut-off point enables comparison of results between studies. To our knowledge, two previous studies proposed a defined score for evaluating recovery from sciatica. Both studies used composite scores and presented strict definitions. One study defined recovery from sciatica as  $ODI \leq 20 + VAS \text{ pain} \leq 15 + \text{muscle strength } 5/5$  [3], and another used the definition absence of sciatic leg pain  $0/10 + RMDQ \leq 3$  [9]. Low back pain recovery has in other studies been defined using cut-off points, but there is lack of consistency among the measures [10].

There are limitations to the study. Number and characteristics of eligible patients who did not enter the study were not registered. Clinical examination was carried out by different physiotherapists and physicians; this might

give an assessment error. The ODI and the RMDQ were not used for comparison with the sciatica specific questionnaires.

In conclusion, using the cut-off scores which best discriminated the patients with and without success at 1 year, the success rates varied from 49 to 58% depending on the measure used.

**Acknowledgments** Supported by a grant from the South-Eastern Norway Regional Health Authority. We thank Eli Molde Hagen, Knut Morten Huneide, Anett Bjørnødegård and Bjarte Justnæs for their help with the data collection and Leiv Sandvik for statistical advice.

**Conflict of interest** None.

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

## References

- Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, Singer DE (1996) The Maine lumbar spine study, part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine* 21:1777–1786
- Atlas SJ, Deyo RA, van den Ancker M, Singer DE, Keller RB, Patrick DL (2003) The Maine-Seattle back questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: results of a derivation and validation cohort analysis. *Spine* 28:1869–1876
- Balagué F, Nordin M, Sheikhzadeh A, Echegoyen AC, Brisby H, Hoogewoud HM, Fredman P, Skovron ML (1999) Recovery of severe sciatica. *Spine* 24:2516–2524
- Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 25:3100–3103
- Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY (2008) Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 8:968–974
- Gibson JN, Waddell G (2007) Surgical interventions for lumbar disc prolapse: updated cochrane review. *Spine* 32:1735–1747
- Grogle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M (2008) Reliability, validity, and responsiveness of the Norwegian versions of the Maine-Seattle back questionnaire and the Sciatica Bothersomeness and frequency indices. *Spine* 33:2347–2353
- Hoogland T, Schubert M, Miklitz B, Ramirez A (2006) Transforaminal posterolateral endoscopic discectomy with or without the combination of a low-dose chymopapain: a prospective randomized study in 280 consecutive cases. *Spine* 31:E890–E897
- Jensen TS, Albert HB, Sorensen JS, Manniche C, Leboeuf-Yde C (2007) Magnetic resonance imaging findings as predictors of clinical outcome in patients with sciatica receiving active conservative treatment. *J Manipulative Physiol Ther* 30:98–108
- Kamper SJ, Stanton TR, Williams CM, Maher CG, Hush JM et al (2011) How is recovery from low back pain measured? A systematic review of the literature. *Eur Spine J* 20:9–18
- Koes BW, van Tulder MW, Peul WC (2007) Diagnosis and treatment of sciatica. *BMJ* 334:1313–1317
- Luijsterburg PAJ, Verhagen AP, Ostelo RWJG, van den Hoogen HJMM, Peul WC, Avezaat CJJ, Koes BW (2008) Physical therapy plus general practitioners' care versus general practitioners' care alone for sciatica: a randomised clinical trial with a 12-month follow-up. *Eur Spine J* 17:509–517
- Mannion AF, Porchet F, Kleinstuck FS, Lattig F, Jeszenszky D, Bartanusz V, Dvorak J, Grob D (2009) The quality of spine surgery from the patient's perspective: part 2. Minimal clinically important difference for improvement and deterioration as measured with the Core Outcome Measures Index. *Eur Spine J* 18(Suppl 3):374–379
- Nykvist F, Hurme M, Alaranta H, Einola S (1989) A prospective 5-year follow-up study of 276 patients hospitalized because of suspected lumbar disc herniation. *Int Disabil Stud* 11:61–67
- Osterman H, Seitsalo S, Karppinen J, Malmivaara A (2006) Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. *Spine (Phila Pa 1976)* 31:2409–2414
- Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB (1995) Assessing health-related quality of life in patients with sciatica. *Spine* 20:1899–1908; discussion 1909
- Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, Thomeer RT, Koes BW, Leiden-The Hague Spine Intervention Prognostic Study Group et al (2007) Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 356:2245–2256
- Saal JA, Saal JS (1989) Nonoperative treatment of herniated lumbar intervertebral disc with radiculopathy. An outcome study. *Spine* 14:431–437
- Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42
- Thomas KC, Fisher CG, Boyd M, Bishop P, Wing P, Dvorak MF (2007) Outcome evaluation of surgical and nonsurgical management of lumbar disc protrusion causing radiculopathy. *Spine* 32:1414–1422
- Ware JE Jr, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30:473–483
- Weber H, Holme I, Amlie E (1993) The natural course of acute sciatica with nerve root symptoms in a double-blind placebo-controlled trial evaluating the effect of piroxicam. *Spine* 18:1433–1438
- Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA (2006) Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA* 296:2451–2459
- Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA (2006) Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 296:2441–2450







# Prognostic Factors for Non-Success in Patients with Sciatica and Disc Herniation

**Anne Julsrud Haugen**, Department of Rheumatology, Østfold Hospital Trust, Fredrikstad, Norway  
[annhau@so-hf.no](mailto:annhau@so-hf.no)

**Jens Ivar Brox**, Section for Back Surgery and Physical Medicine and Rehabilitation, Orthopaedic Department, Oslo University Hospital, Rikshospitalet, Oslo, Norway  
[j.i.brox@medisin.uio.no](mailto:j.i.brox@medisin.uio.no)

**Lars Grøvle**, Department of Rheumatology, Østfold Hospital Trust, Fredrikstad, Norway  
[largro@so-hf.no](mailto:largro@so-hf.no)

**Anne Keller**, Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Ullevål, Norway  
[akel0008@regionh.dk](mailto:akel0008@regionh.dk)

**Bård Natvig**, National Resource Centre for Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway;  
Department of General Practice and Community Medicine, ASAM, University of Oslo, Norway  
[bard.natvig@medisin.uio.no](mailto:bard.natvig@medisin.uio.no)

**Dag Soldal**, Department of Rheumatology, Sørlandet Hospital Health Enterprise, Kristiansand, Norway  
[dag.magnar.soldal@sshf.no](mailto:dag.magnar.soldal@sshf.no)

**Margreth Grotle**, FORMI (Communication Unit for Musculoskeletal Disorders), Division of Neuroscience, Oslo University Hospital, Ullevål, Norway;  
National Resource Centre for Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway  
[margreth.grotle@medisin.uio.no](mailto:margreth.grotle@medisin.uio.no)

Corresponding author: Anne Julsrud Haugen [annhau@so-hf.no](mailto:annhau@so-hf.no)

# **Abstract**

## **Background**

Few studies have investigated prognostic factors for patients with sciatica, especially for patients treated without surgery. The aim of this study was to identify factors associated with non-success after 1 and 2 years of follow-up.

## **Methods**

The study was a prospective multicentre observational study including 466 patients with sciatica and lumbar disc herniation. Potential prognostic factors were sociodemographic characteristics, back pain history, kinesiophobia, emotional distress, pain, subjective health complaints and clinical examination findings. Study participation did not alter treatment considerations for the patients in the clinics. Patients reported on the questionnaires if surgery of the disc herniation had been performed. Uni- and multivariate logistic regression analyses were used to evaluate factors associated with non-success, defined as Maine–Seattle Back Questionnaire score of  $\geq 5$  (0–12).

## **Results**

At 1 and 2 year follow-ups, the rates of non-success were 44% and 39%, respectively. Approximately 1/3 of the patients were treated surgically. In the final multivariate model, non-success at 1 year was significantly associated with being male (OR1.7 [95% CI;1.1–2.7]), smoker (2.1 [1.3–3.3]), more back pain (1.0 [1.01–1.02]), more subjective health complaints (1.1 [1.0–1.2]), reduced tendon reflex (1.6 [1.01–1.02]), and not treated surgically (3.0 [1.8–5.0]). Further, factors significantly associated with non-success at 2 years were duration of back problems >1 year (1.9 [1.1–3.3]), duration of sciatica >3 months (2.3 [1.4–3.8]), more subjective health complaints (1.1 [1.0–1.1]) and kinesiophobia (1.04 [1.00–1.08]).

## **Conclusions**

Prognosis of sciatica was only weakly associated with sciatica-specific clinical findings, and more strongly associated with pain, disability, comorbid subjective health complaints and kinesiophobia. This calls for a broader assessment of patients with



sciatica than the traditional clinical assessment in which mainly the physical symptoms and signs are investigated.

## **Key words**

sciatica; disc herniation; prognostic factors; non-success

## **Background**

Sciatica, also known as nerve root pain or radicular pain, is defined as radiating pain of the leg below knee level. A lumbar disc herniation is the most common cause of sciatica. Success rates and prognosis for sciatica vary between studies. A Dutch study on primary care patients indicated a good prognosis for sciatica where approximately 75% of the patients recovered after 3 months [1]. However, in a Finnish study on patients with sciatica who were referred to hospital, nearly 70% had persistent sciatica symptoms 13 years later [2].

Most studies on prognostic factors for sciatica have been performed on patients who received surgical treatment and have assessed the influence of sociodemographic, work-related, psychological, imaging, pain-related, surgery-related, and clinical factors [3, 4]. In a recent randomised study on early surgery versus prolonged conservative treatment for sciatica, female gender was found to be a strong predictor of an unsatisfactory outcome [5].

Limited knowledge exists about the prognosis of patients who are non-surgically treated for sciatica; a recent review did not find any strong or consistent predictor of persistent disability in non-surgically treated patients with sciatica [6]. Symptoms of depression and anxiety were predictors for a worsening of pain and function after up to 3 years of follow-up in the Maine Lumbar Spine Study (MLSS) where both surgically treated and not surgically treated patients were enrolled [7]. In a cohort of French workers, psychosomatic symptoms, long-lasting duration of sciatica, carrying heavy loads and driving at least 2 hours per day predicted persistent sciatica after 2 years [8]. Clinical findings of radiculopathy are important for the diagnosis of sciatica, but little is known about how clinical signs and symptoms influence the prognosis of sciatica [9].

The objective of the present study was to identify prognostic factors associated with non-success after 1 and 2 years of follow-up for sciatica and disc herniation in patients referred for secondary care.

# Methods

## Design

A prospective observational multicenter cohort study was conducted. The Regional Committee for Medical Research Ethics and The Ombudsmann for Privacy in Research at the Norwegian Social Science Data Services approved the study protocol.

## Cohort selection and recruitment

Patients were recruited from specialty back clinics at four hospitals in southeast Norway (Sykehuset Østfold, Sørlandet Sykehus, Oslo Universitetssykehus Ullevål and Sykehuset Innlandet). Consecutive eligible patients were invited to participate in the study by clinic staff.

Inclusion criteria were age  $\geq 18$  years, radiating pain and/or paresis below knee level and a disc herniation at the corresponding level and side that had been verified by magnetic resonance imaging (MRI) or computed tomography (CT). Exclusion criteria were prior surgery at the same disc level, fracture, infection, malignancy, pregnancy and lack of fluency in Norwegian. All participants received oral and written information about the study and gave informed consent.

Study participation did not involve any specific type of intervention or alter treatment considerations for the patients in the clinics. Hence, patients received the usual consultations including information about sciatica and disc herniation, back exercises, physical therapy, and pain medication. Surgery was performed for patients with severe symptoms after discussion with an orthopaedic surgeon at each center.

## Procedure

At the day of inclusion patients completed a comprehensive questionnaire. Clinical examination was conducted by a physician or physiotherapist.

A follow-up questionnaire and a prepaid envelope were sent to the patients after 12 and 24 months. A reminder was sent after 2 weeks if no reply was obtained.

## Potential prognostic factors

At inclusion, the following sociodemographic factors were recorded: age, sex, education (years of schooling), smoking status and work status. Patients also reported their history of back pain (dichotomised to  $< 1$  year or  $\geq 1$  year), previous sciatica

episodes (0 or  $\geq 1$ ), and duration of the index episode (dichotomised to  $< 3$  months or  $\geq 3$  months). Furthermore, the following patient-reported variables were recorded.

Pain intensity in the back and in the leg (sciatica) during the previous week according to a horizontal visual analogue scale (VAS) ranging from 0 (no pain) to 100 (worst pain ever).

The Sciatica Bothersomeness Index (SBI) was used to assess sciatic symptoms [10]. The SBI is a composite of scores for four symptoms: leg pain (sciatica); numbness or tingling in the leg, foot or groin; weakness in the leg or foot; and back or leg pain while sitting. Scores were in the range of 0-6 for each item, providing a total score of 0-24, where higher scores indicate worse symptoms.

The subjective health complaints questionnaire contained 29 common somatic and psychological complaints [11]. Patients graded the intensity of each complaint during the previous month as not at all (0), a little (1), some (2) or severe (3). Each item was dichotomised to absent (0) or present (1, 2, 3). Two of the items, low back pain and leg pain during exercise, were excluded, resulting in a score in the range of 0-27.

Patients also completed the Tampa scale for Kinesiophobia (TSK) [12, 13], a 13-item, four-point questionnaire. Scores range from 13 to 52, where higher scores indicate increased kinesiophobia.

Emotional distress was assessed using the Hopkins Symptom Check List-25 (HSCL-25) [14]. The questionnaire includes 25 items on depression, anxiety and somatisation during the previous week and ranges from 1 (not at all) to 4 (extremely). The total score is calculated by adding all scores and dividing the sum by the number of completed items. An average item score  $\geq 1.75$  were found to be a good predictor of current help-seeking in a Norwegian epidemiologic study and is commonly used to define cases with emotional distress [15].

At inclusion, a clinical examination was carried out by physiotherapists or physicians. The assessment included motor function (deemed abnormal if the extension or flexion of the knee or ankle or the extension of the big toe was reduced, if the Trendelenburg test was positive or if abnormal tiptoe or heel walking was present), sensibility (deemed abnormal if tactile sensibility was reduced), reflexes of the Achilles tendon or patella (deemed abnormal if reduced or not elicited) and the straight-leg-raising test (deemed abnormal if  $< 60^\circ$ ).

In order to assess the effect of surgery of the herniated disc, patients who had undergone surgery during the observational period reported the date of surgery at follow-up questionnaires.

### **Outcome measure and definition of non-success**

The Maine-Seattle Back Questionnaire (MSBQ) was the main outcome measure [10]. The MSBQ is a shortened version of the Roland-Morris Disability Questionnaire that was modified for patients with sciatica and spinal stenosis [16]. The scale is composed of 12 items, each with the answer yes (1) or no (0), achieving a score range of 0-12. The MSBQ assesses disability and functional limits due to sciatic and back pain, and higher scores indicate worse symptoms. We have previously reported that the MSBQ is the best measure for distinguishing between success and non-success in sciatica at 1 year of follow-up [17]. Non-success was defined as an MSBQ score  $\geq 5$ .

### **Statistical analysis**

The required sample size was calculated based on the expectation that surgical treatment would be necessary for 30% of the patients and that 30% of those who were surgically treated and 50% of those who were not surgically treated would not experience a successful outcome after 1 year [18]. Estimation of sample size was based on 90% power to detect a difference at the 5% significance level in good outcome between patients who were surgically treated and patients who were not surgically treated using a two-sided exact Fisher test. Power analysis indicated that a sample size of 300 patients was required. Considering loss-to-follow-up, we decided to include at least 400 patients.

Baseline prognostic variables were analysed using logistic regression with MSBQ  $\geq 5$  as dependent variable. Missing items in questionnaires were substituted with mean values. A chi-square test and a Student's t-test were used to compare responders and non-responders. The significance level was set at 0.05. The variance inflation factor (VIF) was used to check for multicollinearity and variables with a VIF  $<5$  were accepted for inclusion in the model [19]. The results are presented as odds ratios (ORs) with 95% confidence intervals (CIs). Baseline variables from the univariate analysis with a p-value  $< 0.10$  were included in a multivariate model. The final model controlled for age, sex and baseline MSBQ. In a backward approach, the non-significant variable with the

highest p-value was removed in a stepwise manner until all variables had  $p < 0.05$ . In the last step, in order to assess the effect of surgery on the outcome, the variable surgery (yes/no) was added to the model. Variables remaining in the final model were tested for possible interaction effects. Data were analysed using the SPSS package (version 18.0) for Windows (SPSS Inc., Chicago, IL).

## **Results**

### **Study sample**

In total, 466 patients were included. Four hundred nine patients (88%) responded to the 1-year follow-up questionnaire and 380 (82%) responded to the 2-year follow-up questionnaire. Among the responders, 120 (29%) had received surgical treatment during the first year of follow-up and nine patients had received surgical treatment during the second year of follow-up. Because of some loss to follow-up, 120 (32%) were registered as surgically treated at 2 years.

### **Missing data**

The participants who did not respond to the questionnaire at 1 year were significantly younger, had higher back pain scores and were more frequently positive for the straight-leg-raising test at baseline. The non-responders at 2 years were younger, more often current smokers, had higher back pain scores, more emotional distress and were more frequently positive for the straight-leg-raising test at baseline.

### **Baseline characteristics**

The baseline characteristics are presented in Table 1. The mean age at baseline was 43.6 years and 58% were males. Most of the patients were employed although only approximately 20% were in work at inclusion. Furthermore, 45% of the sample reported a sciatica episode for the first time. In general, the patients reported high scores on the self-reported questionnaires, in particular for leg pain, disability and emotional distress.

### **Non-success**

Patients with non-success (MSBQ  $\geq 5$ ) numbered 178 patients (44%) at 1 year and 145 (39%) at 2 years. Among the surgically treated patients, 42 (35%) had non-

success at the 1-year follow-up, and 47 (39%) had non-success at the 2-year follow-up. In the non-surgical group, 136 (47%) and 98 (39%) patients had non-success at 1 and 2 years, respectively.

### **Prognostic indicators**

Table 2 presents the associations between non-success and all baseline factors. Table 2 also includes the association between non-success and surgical treatment during follow-up. Higher scores for all self-reported health status variables, shorter education length and longer duration of back problems were significantly associated with non-success at both 1 and 2 years according to the univariate analyses. Smoking was also related to non-success at 1 year, OR=1.99 (1.34–2.98), but was not significant at 2 years, OR=1.51 (0.99–2.30). Non-surgical treatment was associated with non-success at 1 year, OR=1.66 (1.07–2.58), but not at 2 years, OR=0.98 (0.63–1.52). When adjusting each baseline variable for surgical treatment during follow-up, only a positive straight-leg-raising test was borderline significantly associated with non-success at 1 year (OR=1.47 [0.98–2.23],  $p = 0.07$ ).

The final multivariable regression models are presented in Table 3. Subjective health complaints were the only prognostic factor significantly associated with non-success at both the 1- and 2-year follow-ups. Other variables independently and significantly associated with non-success at 1 year were male sex, higher intensity of back pain and abnormal reflexes in the clinical examination. Prognostic factors for non-success at 2 years included higher kinesiophobia and longer duration of back pain and sciatica. Adding surgical status to the final models showed non-surgical treatment to be significantly associated with non-success at 1 year (OR=2.97 [1.75–5.04]), but not at 2 years (OR=1.32 [0.78–2.23]). The adjustment variable, the baseline scores of the MSBQ, remained significant in the multivariate models both at 1 and 2 year follow-up. For the final models, the explained variance assessed by Nagelkerke  $R^2$  was 26% at both 1 and 2 years. No multicollinearity or interactions were detected for the selected baseline variables. At baseline the correlations assessed by Spearman's Rho were: Subjective health complaints and kinesiophobia 0.19; subjective health complaints and emotional distress 0.59; and kinesiophobia and emotional distress 0.41.

## Discussion

This study shows that 44% of the patients with sciatica who were referred for secondary care had a non-successful outcome at 1 year and 39% at 2 years. A high score for subjective health complaints was the only variable that predicted non-success at both 1 and 2 years. Males, smokers, patients with higher scores for low back pain and patients who had not undergone surgery had an independently higher risk of non-success at 1 year, but not at 2 years of follow-up. A long duration of back pain and sciatica symptoms and a high level of kinesiophobia were associated with non-success at 2 years. No sciatica-specific clinical findings were associated with non-success, except for a weak association with abnormal tendon reflex at 1 year.

The main strength of this study is the large sample size, which was considered representative of sciatica patients in Norway, the high response rate and the use of imaging to confirm the diagnosis of disc herniation. We have used the most precise outcome measure for the current cohort, which in a previous study showed the highest sensitivity and specificity to discriminate between successful outcome or not for sciatica patients [17]. A broad range of potential prognostic variables including several clinical findings, psychological variables and subjective health complaints were investigated. There is a need for a prospective observation study, particularly on conservatively treated patients where the few studies have shown divergent results [6]. There are also few studies with more than one year follow up [20]. Additionally, the conclusions of previously published randomised controlled studies are weakened by high rates of cross-over [21-23].

A limitation to the internal validity of the study was an incomplete recording of patients who according to the inclusion and exclusion criteria were eligible but for some reason either were not invited or declined to participate. Other potential prognostic factors were not investigated in the current study, for example details of imaging findings [24-27] and the phenomenon of centralization [28, 29] may become important. Lastly, the final model should be tested in another sciatica sample.

The long-term prognosis in the present study was in accordance with the findings of other studies, but the prognosis for the surgically treated patients in the current study was poorer than that of other studies [18, 30]. Additionally, the differences between surgically treated and non-surgically treated patients were smaller than in comparable studies [30, 31].

We have no clear explanation as to why more subjective health complaints was the only variable associated with non-success at both 1 and 2 years, but a lower threshold for reporting bodily discomfort, including sciatica may have had an influence. The subjective health complaints inventory includes muscular pain, headaches, stomach pain and discomfort, hot flushes, extra heart beats, sleep problems, tiredness, dizziness, anxiety and depressed thoughts during the previous month [32]. In two Norwegian studies, high scores were associated with reduced function and more complaints in patients with nonspecific low back pain [33] and whiplash [34]. We have previously reported that the patients in the current cohort at baseline reported more subjective health complaints than the general population and that the number of subjective health complaints nearly doubled in those with persisting sciatica at the 1 year follow-up [35]. When subjective health complaints were entered into the regression model, emotional distress lost significance. Emotional distress has been found to be related to pain and disability in sciatica patients [7, 36], and is reported to be an important prognostic factor in nonspecific low back pain [37, 38]. However, none of those studies tested the influence of subjective health complaints.

In the present study, females had better outcomes than males at the 1-year follow-up. This contrasts with the results of the study of Peul [5], in which female sex was a strong predictor for an unsatisfactory outcome at 1 year for patients with sciatica and disc herniation. One possible explanation for the divergent results is that our data were adjusted for subjective health complaints, emotional distress and kinesiophobia. The exclusion criteria in the study of Peul were a duration of sciatica symptoms of more than 12 weeks, similar complaints during the previous year, and severe comorbidity. The current study is probably more representative of the majority of patients with sciatica and disc herniation who are referred for secondary care.

The poor prognosis among smokers is in agreement with the results of some studies on surgically treated patients [39, 40], but conflicting results have also been published [4].

Kinesiophobia was an independent prognostic variable for non-success at 2 years. The fear-of-movement/(re)injury model states that the reaction to pain may consist of confrontation in patients with a non-catastrophizing behaviour, and of avoidance in patients with a high catastrophizing behaviour. Pain-related fear may lead to prolonged chronic pain and disability [41] and was prognostic for patients with



nonspecific low back pain after 6 months in two Dutch populations, one of which was a population-based survey [42] and one of which included army workers [43]. In a cross-sectional study of a Swedish population with specific low back pain (defined as disc herniation, isthmic spondylolisthesis or spinal stenosis) attending an orthopaedic clinic, high scores for kinesiophobia were associated with high disability scores [44]. Another cross-sectional Swedish study on patients treated with surgery for disc herniation found that half of the patients suffered from kinesiophobia 10–34 months after the operation and that patients with kinesiophobia were more affected in several other variables as pain, disability and symptoms of depression [45]. Contrary to these findings, fear of movement and pain catastrophizing were not associated with recovery among patients with residual complaints at 3 and 12 months following lumbar disc surgery [46].

The final models showed that patients who were not treated surgically were nearly three times more likely to have non-success at 1 year, but no significant association was identified between surgical treatment and outcome at 2 years. Most operations were performed during the first year of follow-up. The benefits of surgical treatment decreased with time, which is similar to the results of other studies [20, 23, 47].

Most ORs in the final models were <2 which means sparse associations with the variables. However, patients who at inclusion had longer duration of back pain and sciatica were about twice more likely to have non-success at 2 years. This is in line with some studies on surgically treated patients [39, 40, 48], but not in line with conservatively treated patients [25, 49]. None of these studies had follow-up more than 1 year and cannot be properly compared with the results of the current study.

## **Conclusions**

Prognosis of sciatica was only weakly associated with sciatica-specific clinical findings, and more strongly associated with pain, disability, comorbid subjective health complaints and kinesiophobia. The absence of surgical treatment during the observation period was associated with non-success at the 1-year follow-up, but not at the 2-year follow-up. This calls for a broader assessment of patients with sciatica than the traditional clinical assessment in which mainly the physical symptoms and signs are investigated. The results of the present study may be used to identify subgroups of patients referred to hospital with an increased risk of a poor prognosis for sciatica.

## List of abbreviations used

CI	Confidence Interval
CT	Computed Tomography
HSCL-25	Hopkins Symptom Check List-25
MLSS	Maine Lumbar Spine Study
MRI	Magnetic Resonance Imaging
MSBQ	Maine-Seattle Back Questionnaire
OR	Odds Ratio
SBI	Sciatica Bothersomeness Index
SPSS	Statistical Package for the Social Sciences
TSK	Tampa scale for Kinesiophobia
VAS	Visual Analogue Scale
VIF	Variance Inflation Factor

## Competing interests

There are no conflicts of interest to report interest in regard to the present work.

## Authors' contributions

Conception and design: AJH, LG, MG. Acquisition and data: AJH, LG, AK, DS. Analysis and interpretation of data: AJH, JIB, LG, AK, BN, MG. Drafting the manuscript: AJH. Critical revision of the manuscript: AJH, JIB, LG, AK, DS, BN, MG. Statistical analysis: AJH, LG, MG. All authors read and approved the manuscript.

## Acknowledgements

Acknowledgement (1): Supported by a grant from the Eastern Norway Regional Health Authority.

Acknowledgment (2): We thank Eli Molde Hagen, Knut Morten Huneide, Anett Bjørnødegård and Bjarte Justnæs for their help with the data collection and Leiv Sandvik for statistical advice. We also thank Camilla Ihlebæk for contributions to the manuscript.

## References

1. Vroomen PC, de Krom MC, Knottnerus JA: **Predicting the outcome of sciatica at short-term follow-up.** Br J Gen Pract 2002, **52**(475):119-123.
2. Nykvist F, Hurme M, Alaranta H, Kaitsaari M: **Severe sciatica: a 13-year follow-up of 342 patients.** Eur Spine J 1995, **4**(6):335-338.
3. den Boer JJ, Oostendorp RA, Beems T, Munneke M, Oerlemans M, Evers AW: **A systematic review of bio-psychosocial risk factors for an unfavourable outcome after lumbar disc surgery.** Eur Spine J 2006, **15**(5):527-536.
4. Mannion AF, Elfering A: **Predictors of surgical outcome and their assessment.** Eur Spine J 2006, **15 Suppl 1**:S93-108.
5. Peul WC, Brand R, Thomeer RT, Koes BW: **Influence of gender and other prognostic factors on outcome of sciatica.** Pain 2008, **138**(1):180-191.
6. Ashworth J, Konstantinou K, Dunn KM: **Prognostic Factors in Non-Surgically Treated Sciatica: A Systematic Review.** BMC Musculoskelet Disord 2011, **12**(1):208.
7. Edwards RR, Klick B, Buenaver L, Max MB, Haythornthwaite JA, Keller RB, Atlas SJ: **Symptoms of distress as prospective predictors of pain-related sciatica treatment outcomes.** Pain 2007, **130**(1-2):47-55.
8. Tubach F, Beaute J, Leclerc A: **Natural history and prognostic indicators of sciatica.** J Clin Epidemiol 2004, **57**(2):174-179.
9. Koes BW, van Tulder MW, Peul WC: **Diagnosis and treatment of sciatica.** BMJ 2007, **334**(7607):1313-1317.
10. Grovle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M: **Reliability, validity, and responsiveness of the Norwegian versions of the Maine-Seattle Back Questionnaire and the Sciatica Bothersomeness and Frequency Indices.** Spine 2008, **33**(21):2347-2353.
11. Eriksen HR, Ihlebaek C, Ursin H: **A scoring system for subjective health complaints (SHC).** Scand J Public Health 1999, **27**(1):63-72.
12. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H: **Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance.** Pain 1995, **62**(3):363-372.
13. Haugen AJ, Grovle L, Keller A, Grotle M: **Cross-cultural adaptation and validation of the Norwegian version of the Tampa scale for kinesiophobia.** Spine 2008, **33**(17):E595-601.
14. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L: **The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory.** Behav Sci 1974, **19**(1):1-15.
15. Sandanger I, Moum T, Ingebrigtsen G, Dalgard OS, Sørensen T, Bruusgaard D: **Concordance between symptom screening and diagnostic procedure: the Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview I.** Soc Psychiatry Psychiatr Epidemiol 1998, **33**(7):345-354.
16. Atlas SJ, Deyo RA, van den Ancker M, Singer DE, Keller RB, Patrick DL: **The Maine-Seattle back questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: results of a derivation and validation cohort analysis.** Spine 2003, **28**(16):1869-1876.
17. Haugen AJ, Grovle L, Brox JI, Natvig B, Keller A, Soldal D, Grotle M: **Estimates of success in patients with sciatica due to lumbar disc herniation depend upon outcome measure.** Eur Spine J 2011, **20**(10):1669-1675.

18. Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, Singer DE: **The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica.** Spine 1996, **21**(15):1777-1786.
19. O'Brien RM: **A caution regarding rules of thumb for variance inflation factors.** Quality and Quantity 2007, **41**(5):673-690.
20. Gibson JN, Waddell G: **Surgical interventions for lumbar disc prolapse: updated cochrane review.** Spine 2007, **32**(16):1735-1747.
21. Osterman H, Seitsalo S, Karppinen J, Malmivaara A: **Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up.** Spine 2006, **31**(21):2409-2414.
22. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA: **Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial.** JAMA 2006, **296**(20):2441-2450.
23. Weber H: **Lumbar disc herniation. A controlled, prospective study with ten years of observation.** Spine 1983, **8**(2):131-140.
24. Jensen TS, Albert HB, Sorensen JS, Manniche C, Leboeuf-Yde C: **Magnetic resonance imaging findings as predictors of clinical outcome in patients with sciatica receiving active conservative treatment.** J Manipulative Physiol Ther 2007, **30**(2):98-108.
25. Balague F, Nordin M, Sheikzadeh A, Echevoyen AC, Brisby H, Hoogewoud HM, Fredman P, Skovron ML: **Recovery of severe sciatica.** Spine 1999, **24**(23):2516-2524.
26. Carragee EJ, Kim DH: **A prospective analysis of magnetic resonance imaging findings in patients with sciatica and lumbar disc herniation: correlation of outcomes with disc fragment and canal morphology.** Spine 1997, **22**(14):1650.
27. Beauvais C, Wybier M, Chazerain P, Harboun M, Lioté F, Roucoules J, Koeger AC, Bellaiche L, Orcel P, Bardin T: **Prognostic value of early computed tomography in radiculopathy due to lumbar intervertebral disk herniation. A prospective study.** Joint Bone Spine 2003, **70**(2):134-139.
28. Aina A, May S, Clare H: **The centralization phenomenon of spinal symptoms--a systematic review.** Man Ther 2004, **9**(3):134-143.
29. Skytte L, May S, Petersen P: **Centralization: its prognostic value in patients with referred symptoms and sciatica.** Spine 2005, **30**(11):E293-9.
30. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA: **Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort.** JAMA 2006, **296**(20):2451-2459.
31. Atlas SJ, Keller RB, Chang Y, Deyo RA, Singer DE: **Surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: five-year outcomes from the Maine Lumbar Spine Study.** Spine 2001, **26**(10):1179-1187.
32. Ihlebaek C, Brage S, Eriksen HR: **Health complaints and sickness absence in Norway, 1996-2003.** Occup Med (Lond) 2007, **57**(1):43-49.
33. Hagen EM, Svensen E, Eriksen HR, Ihlebaek CM, Ursin H: **Comorbid subjective health complaints in low back pain.** Spine 2006, **31**(13):1491-1495.
34. Ihlebaek CM, Ødegaard A, Vikne J, Eriksen HR, Lærum E: **Subjective health complaints in patients with chronic Whiplash Associated**

- Disorders (WAD). Relationships with physical, psychological, and collision associated factors.** *Nor J Epidemiol* 2006, **16**(2):119-126.
35. Grovle L, Haugen AJ, Ihlebaek CM, Keller A, Natvig B, Brox JI, Grotle M: **Comorbid subjective health complaints in patients with sciatica: A prospective study including comparison with the general population.** *J Psychosom Res* 2011, **70**(6):548-556.
  36. Grovle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M: **The bothersomeness of sciatica: patients' self-report of paresthesia, weakness and leg pain.** *Eur Spine J* 2010, **19**(2):263-269.
  37. Grotle M, Vollestad NK, Brox JI: **Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II.** *Spine* 2006, **31**(9):1038-1046.
  38. Costa Lda C, Maher CG, McAuley JH, Hancock MJ, Herbert RD, Refshauge KM, Henschke N: **Prognosis for patients with chronic low back pain: inception cohort study.** *BMJ* 2009, **339**:b3829.
  39. Rothoerl RD, Woertgen C, Brawanski A: **When should conservative treatment for lumbar disc herniation be ceased and surgery considered?** *Neurosurg Rev* 2002, **25**(3):162-165.
  40. Woertgen C, Holzschuh M, Rothoerl RD, Brawanski A: **Does the choice of outcome scale influence prognostic factors for lumbar disc surgery? A prospective, consecutive study of 121 patients.** *Eur Spine J* 1997, (3):173-180.
  41. Vlaeyen JW, Linton SJ: **Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art.** *Pain* 2000, **85**(3):317-332.
  42. Picavet HS, Vlaeyen JW, Schouten JS: **Pain catastrophizing and kinesiophobia: predictors of chronic low back pain.** *Am J Epidemiol* 2002, **156**(11):1028-1034.
  43. Helmhout PH, Staal JB, Heymans MW, Harts CC, Hendriks EJ, de Bie RA: **Prognostic factors for perceived recovery or functional improvement in non-specific low back pain: secondary analyses of three randomized clinical trials.** *Eur Spine J* 2010, **19**(4):650-659.
  44. Lundberg M, Frennered K, Hagg O, Styf J: **The impact of fear-avoidance model variables on disability in patients with specific or nonspecific chronic low back pain.** *Spine* 2011, **36**(19):1547-1553.
  45. Svensson GL, Lundberg M, Ostgaard HC, Wendt GK: **High degree of kinesiophobia after lumbar disc herniation surgery.** *Acta Orthop* 2011, **82**(6):732-736.
  46. Ostelo RW, Vlaeyen JW, van den Brandt PA, de Vet HC: **Residual complaints following lumbar disc surgery: prognostic indicators of outcome.** *Pain* 2005, **114**(1-2):177-185.
  47. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD: **Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline.** *Spine* 2009, **34**(10):1094-1109.
  48. Nygaard OP, Kloster R, Solberg T: **Duration of leg pain as a predictor of outcome after surgery for lumbar disc herniation: a prospective cohort study with 1-year follow up.** *J Neurosurg* 2000, **92**(2 Suppl):131-134.
  49. Hasenbring M, Marienfeld G, Kuhlendahl D, Soyka D: **Risk factors of chronicity in lumbar disc patients. A prospective investigation of biologic, psychologic, and social predictors of therapy outcome.** *Spine* 1994, **19**(24):2759.

## Tables

**Table 1**

**Baseline characteristics. Values are n (%) unless otherwise stated.**

	<b>N = 466</b>	<b>(%)</b>
<b>Sociodemographic variables</b>		
Mean (SD) age years	43.6	(11.5)
Males	268	(57.5)
Current smoker	200	(43.3)
Education > 12 years	227	(50.1)
Working status		
Working full time	92	(19.8)
Partly sick leave	51	(11.0)
Complete sick leave, rehabilitation	234	(50.4)
Disability pension	33	(7.1)
Other	54	(11.7)
<b>Back pain/sciatica history</b>		
First sciatica episode	210	(45.3)
Duration back problems < 1 year	145	(31.1)
Duration current sciatica episode < 3 months	192	(41.4)
<b>Self-reported health status</b>		
Mean (SD) Maine-Seattle Back Questionnaire (0–12)	8	(3)
Mean (SD) Pain intensity* back (0–100)	42.6	(30.0)
Mean (SD) Pain intensity* leg (0–100)	63.2	(28.2)
Mean (SD) Sciatica Bothersomeness Index (0–24)	14	(5)
Mean (SD) Subjective health complaints (0–27)	7.5	(4.5)
Mean (SD) Kinesiophobia† (13–52)	27	(7)
Mean (SD) Emotional distress‡ (1–4)	1.58	(0.43)
<b>Clinical finding</b>		
Muscular weakness	203	(44.5)
Reflex reduced or absent	212	(46.2)
Sensory loss	273	(59.0)
Straight-leg-raising test < 60°	267	(58.0)

\*Visual analogue scale; †Tampa Scale for Kinesiophobia; ‡Hopkins Symptom Check List-25

**Table 2****Factors associated with non-success at 1 and 2 years, OR (95% CI).****Results from univariate analyses. ORs with p< 0.1 in bold text.**

	1 year		2 years	
	OR	(95%CI)	OR	(95%CI)
<b>Sociodemographic variables</b>				
Age	1.01	(0.99–1.03)	<b>1.02</b>	(1.00–1.04)
Male sex (ref female)	1.03	(0.70–1.53)	0.85	(0.56–1.30)
Smoking (ref nonsmoking)	<b>1.99</b>	(1.34–2.98)	<b>1.51</b>	(0.99–2.30)
Education, years (continuous)	<b>0.93</b>	(0.87–0.99)	<b>0.92</b>	(0.85–0.98)
Partly sick leave (ref working)	0.84	(0.38–1.86)	1.50	(0.66–3.42)
Complete sick leave, rehabilitation (ref working)	<b>1.90</b>	(1.11–3.24)	<b>2.18</b>	(1.21–3.95)
Disability pension (ref working)	<b>2.94</b>	(1.23–7.01)	<b>3.82</b>	(1.48–9.82)
Other (ref working)	<b>1.99</b>	(0.96–4.14)	<b>2.08</b>	(0.94–4.60)
<b>Back pain/sciatica history</b>				
Previous sciatica episodes > 0 (ref 0)	1.23	(0.83–1.82)	1.16	(0.76–1.76)
Duration back problems ≥ 1 years (ref <1 year)	<b>2.08</b>	(1.33–3.26)	<b>2.48</b>	(1.52–4.06)
Duration of index episode ≥ 3 months (ref <3 months)	1.21	(0.81–1.80)	<b>1.69</b>	(1.11–2.60)
<b>Self-reported health status</b>				
Maine-Seattle Back Questionnaire (continuous)	<b>1.26</b>	(1.16–1.38)	<b>1.30</b>	(1.19–1.44)
Pain intensity* back (continuous)	<b>1.02</b>	(1.01–1.03)	<b>1.02</b>	(1.01–1.02)
Pain intensity* leg (continuous)	<b>1.01</b>	(1.01–1.02)	<b>1.01</b>	(1.01–1.02)
Sciatica Bothersomeness Index (continuous)	<b>1.11</b>	(1.06–1.16)	<b>1.11</b>	(1.06–1.16)
Subjective health complaints (continuous)	<b>1.13</b>	(1.07–1.82)	<b>1.16</b>	(1.10–1.23)
Kinesiophobia† (continuous)	<b>1.06</b>	(1.03–1.09)	<b>1.07</b>	(1.04–1.10)
Emotional distress‡ (continuous)	<b>3.40</b>	(2.08–5.56)	<b>3.31</b>	(1.95–5.59)
<b>Clinical finding</b>				
Muscular weakness (ref no weakness)	1.37	(0.92–2.03)	1.27	(0.84–1.93)
Reflex reduced or absent (ref normal reflexes)	<b>1.41</b>	(0.95–2.09)	1.07	(0.70–1.62)
Sensory loss (ref intact sensibility)	1.10	(0.74–1.64)	1.24	(0.82–1.90)
Straight-leg-raising test <60° (ref >60°)	1.31	(0.88–1.95)	1.15	(0.76–1.75)
<b>No surgery (ref surgery)</b>	<b>1.66</b>	(1.07–2.58)	0.98	(0.63–1.52)

\*Visual analogue scale; †Tampa Scale for Kinesiophobia; ‡ Hopkins Symptom Check List-25

**Table 3**  
**Factors associated with non-success at 1 and 2 years, OR (95% CI). Results from the final multivariate regression analyses.**

	OR	1 year (95% CI)	p	OR	2 years (95% CI)	p
<b>Sociodemographic variables</b>						
Age	1.00	(0.99–1.02))	0.689	1.02	(0.99–1.04)	0.161
Male sex (ref female)	1.70	(1.06–2.73)	0.029	1.29	(0.78–2.15)	0.326
Smoking (ref nonsmoking)	2.06	(1.31–3.25)	0.002			
<b>Back pain/sciatica history</b>						
Duration back problems ≥1 years (ref <1 year)				1.92	(1.11–3.32)	0.020
Duration of index episode ≥ 3 months (ref <3 months)				2.30	(1.40–3.80)	0.001
<b>Self-reported health status</b>						
Maine-Seattle Back Questionnaire (continuous)	1.24	(1.11–1.38)	<0.001	1.28	(1.14–1.43)	<0.001
Pain intensity* back (continuous)	1.01	(1.01–1.02)	0.002			
Subjective health complaints (continuous)	1.09	(1.03–1.15)	0.003	1.10	(1.03–1.17)	0.003
Kinesiophobia† (continuous)				1.04	(1.00–1.08)	0.033
<b>Clinical finding</b>						
Reflex reduced or absent (ref normal reflexes)	1.62	(1.03–2.56)	0.037			
<b>No surgery (ref surgery)</b>	2.97	(1.75–5.04)	<0.001	1.32	(0.78–2.23)	0.308

\*Visual analogue scale; †Tampa Scale for Kinesiophobia