The diagnosis and conservative treatment of sciatica

Citation for published version (APA):

Vroomen, P. C. A. J. (1998). *The diagnosis and conservative treatment of sciatica*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.19980708pv>

Document status and date:

Published: 01/01/1998

DOI:

[10.26481/dis.19980708pv](https://doi.org/10.26481/dis.19980708pv)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

The diagnosis and conservative treatment of sciatica

The studies presented in this thesis were performed at the department of Neurology of the Maastricht University Hospital and the department of General Practice of the Maastricht University. Both departments participated in the Netherlands School of Primary Care Research (CaRe), which was acknowledged in 1995 by the Royal Dutch Academy of Science (KNAW). The studies were supported by a grant from the Maastricht University Hospital. Financial support for the printing of this thesis has kindly been provided by CaRe, the Maastricht University, CIBA, Yamanouchi, and Roche.

ISBN: 90-9011770-9

Cover: Patrick Vroomen, Maastricht
Cover from Domenico Cotugno's *De ischiade Nervosa Commentarius*, Naples, 1764, Biblioteca nazionale di San Marco
Cover from J.J. Forst's thesis *Contribution à l'étude clinique de la sciatique*, Paris, 1881 (with dr. Lasègue as the president of the committee).

Printing: Datawyse/Krips repro

©1998, P.C.A.J. Vroomen

No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage or retrieval system, without the permission in writing of the author, or, when appropriate, of the publishers of the publications.

The diagnosis and conservative treatment of sciatica

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Universiteit Maastricht,
op gezag van de Rector Magnificus
Prof. Dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen
in het openbaar te verdedigen
op woensdag 8 juli 1998 om 16.00 uur

door

Patrick Carolus Andreas Johannes Vroomen

geboren op 11 juli 1969 te Brunssum

Promotoren: Prof. Dr. J.A. Knottnerus
Prof. Dr. J. Troost

Co-promotor: Dr. M.C.T.F.M. de Krom

Beoordelingscommissie:

Prof. Dr. H.F.J.M. Crebolder (voorzitter)
Prof. Dr. E.A.M. Beuls
Prof. Dr. P.G. Knipschild
Prof. Dr. J. Stam (Universiteit van Amsterdam)
Dr. G.H.I.M. Walenkamp

Voor mijn grootouders

Voor mijn ouders

Preface

On the one hand, science is a lonely business with a great deal of *monologue intérieur* and sheer endless dialogues with a personal computer. I know now why my computer was named Trust® (no advertisement intended).

On the other hand, medical science is really all about people and none of this thesis would have come about, if I hadn't had the opportunity to put my trust in a great many people. Credit where credit is due; this thesis features a big thanks to these people in the preface (and not in a postscript).

It was inspiring to see the enthusiasm and commitment to the study of most of the 305 patients that participated. Elaborately informing patients of the objectives of the study has turned out to be a rewarding activity. I am very grateful for their time and effort despite their painful affliction; they also made the task of a standardized investigation much less tedious. The contribution by the 50 general practitioners has been crucial and surely has lived up to expectations. I hope you appreciate the results.

A thanks to the MRI staff for always managing to squeeze in another patient, to the people from the outpatient clinic (the "Great Randomizers") for always knowing where to find me, to the people that printed and posted my newsletters and to numerous others in the Academic Hospital that helped in some way or another. Thanks, "neurological" colleagues, for words of encouragement in times of deadlines and adversity. Now, you all go for it (Ariane, I'll bring you coffee any time)!

To the data typists at MEMIC for not letting me enter 386000 fields of data myself.

To Arnold Kester; for showing that statistics is not just science, but craft also. The discovery was an unexpected pleasure.

To Jan Wilmink, the image tutor, for a great deal of contemporary inspiration and information regarding sciatica. Can I now enter the Sciatology* Society?

To Jaap Troost, for your doubtless advice and your strong impact on my career.

To Andre Krottnerus, my regular sessions with you were of great diagnostic and prognostic value. It has greatly helped in letting science run a favourable course.

To Marc de Krom. You asked me not to thank you, so I won't. Anyway, working with you has been a pleasure that I hope will continue well into the next millenium! *La bêtise est à l'homme, s'il ne le demand pas à de Krom.*

I thank my friends for the support, Rob and Vincent for backing me up.

To my parents; my gratitude is beyond words!

I thank Inez more than I'd like to commit to this cold paper.

In the end, the greatest challenges are not in science.

---

*Not to be confused with Scientology

Figure 1. The token of the research project ISCHIAS

C

ontents “The diagnosis and conservative treatment of sciatica”

	Preface	7
Chapter 1	Introduction	11
Chapter 2	A systematic review of the value of history and physical examination	19
Chapter 3	Technique and consistency of the history and physical examination	29
Chapter 4	Diagnostic value of the history and the physical examination	41
Chapter 5	Patho-anatomy of the clinical findings in the sciatica patient A Magnetic Resonance Imaging study	53
Chapter 6	A systematic review of the value of the conservative treatment	67
Chapter 7	A randomized controlled trial of 14 days of bed rest versus watchful waiting for sciatica	79
Chapter 8	Clinical findings predicting the prognosis of sciatica	93
Chapter 9	When should the sciatica patient undergo lumbosacral discectomy?	103
Chapter 10	General discussion	113
	References	123
	Summary	131
	Samenvatting	135
	Walk of life	138

It is a hard task to treat what is common in a way of your own
Horatius

This thesis focusses on the topic of sciatica. Unfortunately, there is little consensus in the literature as to what “sciatica” exactly refers to¹. In this thesis the following definitions were used:

Sciatica

a syndrome composed of signs and symptoms supposedly indicating pathological involvement of one or more lumbosacral nerve roots. The prerequisite is the presence of irradiating pain into the leg.

Sciatica due to disc herniation

a disease with disc herniation as the likely mechanical cause of sciatica.

Nerve root compression

the anatomical/radiological observation indicating “sciatica due to disc herniation”.

Lumbosacral radicular syndrome

A complex of signs and symptoms involving the lumbosacral nerve roots and consisting of one or more of the following signs and symptoms: irradiating pain into one or more lumbar or sacral dermatomes (*sciatica proper*), nerve tension signs, and/or neurological deficit (sensory loss, decreased strength, absent/decreased tendon reflexes, urinary incontinence). It is regarded as the near equivalent of sciatica with the remark that the radicular syndrome may occur with or without pain.

As a term “lumbosacral radicular syndrome” should actually be preferred over sciatica. Nonetheless, the term “sciatica” is used in this thesis, as it is the term predominantly used in the Anglosaxian language.

The main purpose of this introduction is to describe the research objectives of the thesis. It contains the motivation for the research that is described in the following chapters. Yet, at the start of the work for this thesis, one question was of even greater relevance to me than the research objectives themselves: why should I participate in research on sciatica?

At first glance, sciatica seemed to have a high prevalence. The literature supported this impression, yet the estimates varied greatly. A varied terminology indicating nerve root involvement and its causes contribute to these widely varying estimations. In approximation of the literature, the following may be stated:

One third of all people will complain of low back pain in the course of a month² and 40-80% will have some degree of low back pain during their lifetime³. The incidence reported for sciatica ranges from 2 to 20 percent³. Forty percent of people with back complaints report to have ever had sciatica, half of whom in the past month⁴. The incidence estimates of sciatica range from 0.6% (data from the Transition project, a Dutch registry of illness categories in primary care) to 2.7%^{5 6}. Strikingly, of all Worker's Compensation claims (absenteeism longer than 52 weeks), eight percent are classified as due to diseases of the intervertebral disc⁷.

Though sciatica is a frequent complaint in the neurological outpatient clinic, it is definitely not a hot topic in neurological research! In a recent european neurological congress, only one of more than 600 abstracts dealt with sciatica. The particular abstract was derived from one of the chapters of this thesis. It had the effect of sobering my expectations and brought me to refocus my perspective. Is it because there is nothing fancy or technological about irradiating low back pain and disc herniations. Or do we already know all we need to know? At second glance, I concluded there were still challenges left. The management of sciatica did not seem satisfactory. An improvement in the management would lead to a benefit in a large number of patients (albeit small for individual cases). The doctor might also be better off. An improved patient management is, of course, always satisfying, but seeing the diversity in common complaints can also improve coping with daily routine. Despite the sobering effects of the congress attendance, I decided to tackle some of the challenges, in the hope the validity of the challenges could be demonstrated along the way. Hence, the work for this thesis got underway. Any work thrives on inspiration by the previous scientific endeavours of others. Where contemporary inspiration is lacking, history may provide solace. Fortunately, sciatica is as old as medicine.

Already in primitive times, sciatica may have been referred to as "hexenschuß" by the early Germans or as "elf's arrow" by the early British⁸. Also, Jacob's Injury in Genesis 32:25-32 may have been a biblical reference to irradiating low back pain. The word sciatica stems from the ancient Greek word ισχιας (ischias), meaning "from the hip". Several Greek philosophers/medical doctors (the ancient Greek were very aware of the continuum between philosophy and medicine) such as Philistion, Herophilus and Erasistratus wrote passages on sciatica. Generally though, the disease was poorly delineated. The father of medicine, Hippocrates (in the Fifth Century BC) came close to present day definitions. Interestingly, he believed an excess of horseback riding to be a major predisposing factor. Bed rest, massage, heat, dietary alterations, and music "to pipe away the pain" were common treatments at the time. In his treaty "*De ischiade et psoadice*", Caelius Aurelianus (Fourth Century AD) very adequately described sciatica as follows: "a strong, severe pain emanating from the lower back and radiating into the buttocks, perineum, and even to the popliteal fossa, calf, foot and toes. The pain was accompanied by severe low back spasm, sensory disturbances,

and muscle wasting. The pain was associated with painful defaecation (even in the toes), a crooked posture and an inability to bend forward. Sudden movements, lifting heavy objects, unaccustomed digging and prolonged sexual intercourse were thought to be etiological”.

There was little advance in the way of diagnosis and therapy of sciatica in the Dark Ages, but the renaissance brought new and frequent references to sciatica. The age of enlightenment also produced the most prominent patient case. The Sun king, Louis XIV (figure 1), was bled and purged to death for his frequent episodes of sciatica (with neurolues as an important differential diagnostic consideration). Domenico Cotugno's book *De Ischiade Nervosa Commentarius* (figure 2) has become famous for providing the first thorough description of the cerebrospinal fluid⁹, yet the major subject of this 1764 work was sciatica. He concluded sciatica to be the result of irritation by cerebrospinal fluid and subsequent inflammation of its coverings. A further stage would be a dropsy ultimately leading to a semi-paralysis. Proposed treatments varied from blood letting by phlebotomy or by haemorrhoid bleeding to purges. Massage was thought to be advantageous as well as olive oil and opium. His personal contribution to treating sciatica was the use of caustics to form blisters particularly below the knee and on the foot. He proudly stated that not a single recurrence was known to him as a result of this treatment. This observation might have been biased by a considerable verification “scare”.

Figure 1. King Louis XIV clearly before or between his bouts of sciatica

DOMINICVS COTVNNIVS

Figure 2. Domenico Cotugno
Taken from the cover of *Elogio Storico del Cavalier D. Domenico Cotugno*, Napoli, 1823

The advances of the nineteenth century were primarily in the description of the history and physical examination of sciatica^{10 11 12}. Many new tests were described such as the *points douloureux* along the sciatic nerve by Valleix, the Lasègue sign (see figure 3) by Lasègue's pupil Forst^{13 14 15}, and Bragard sign and the crossed straight leg raising test by Fajerstajn^{16 17}.

Figure 3. Dr. Ernest Lasègue (1816-1883)

Finally, in their 1934 study¹¹⁶ Mixter and Barr pointed out the association between sciatica and disc herniations as well as the potential for surgical intervention. This study and the development of increasingly sophisticated radiological procedures for visualising nerve root and intervertebral disc may well be the major contributions of the twentieth century.

In summary, the major advances since the ancient Greek and Roman writings on diagnosis and therapy of sciatica have been the description of several clinical signs in the nineteenth century, and the development of the radiological diagnosis and the surgical therapy of sciatica in the twentieth. From Greek Antiquity onwards, bed rest has been a mainstay of therapy.

The questions the historical synopsis raises are practically the same as those that come to mind when managing sciatica patients. When a patient presents to his physician* for the first time with pain irradiating into the leg, many physicians will look for the signs and symptoms of sciatica. Which signs and symptoms may be referred to as sciatica? A definite answer cannot be given as sciatica denotes a syndrome of lumbosacral nerve root involvement, due to a variety of causes, and for which no definite gold standard exists. The definition of such a syndrome is essentially a consensus statement such as the statement on the lumbosacral radicular syndrome on page 11.

Disc herniations are a major cause of sciatica. Considering that sciatica due to disc herniations may have specific characteristics with regard to diagnosis, prognosis and therapy, it seems desirable to establish an early diagnosis of this particular etiology. It raises the question how such an early diagnosis can be made. In other words, **which signs and symptoms indicate sciatica due to disc herniation; how firmly is their value supported by the literature?** Chapter two offers a description of a systematic review of the literature. It summarizes the diagnostic properties of the history and physical examination with regard to sciatica (due to disc herniation).

Before setting out to investigate the diagnostic accuracy of such clinical findings in a new diagnostic study, it has to be determined **how a consistent history and physical examination in the patient suspected of sciatica can be carried out.** It seems appropriate to choose a technique not merely based on the original descriptions in the literature, but also on practicality and on the basis of widely distributed opinions. In chapter three, a synthesis of the technique for assessing history and physical examination in the patient with pain irradiating into the leg is given. The inter-observer consistency for these clinical findings is described.

A diagnosis of sciatica due to disc herniation may, in time, necessitate additional diagnostic procedures and invasive therapy. Therefore, it seems relevant to determine **how accurately an initial diagnosis of sciatica due to disc herniation can be made.** Chapter four highlights a cross-sectional diagnostic study to investigate the properties of history and physical examination in the primary care sciatica patients with Magnetic Resonance Imaging as gold standard. Its results should show the value of history and physical examination in the diagnosis of sciatica due to disc herniation.

When a detailed and standardized clinical and MRI diagnosis is established in the entire sample of primary care patients with irradiating pain to the leg, **what can we learn from the associations between MRI anatomy and the occurrence of particular clinical findings?** Chapter five features an exploration of the potential patho-anatomy underlying the signs and symptoms of sciatica.

When a clinical syndrome of sciatica of sufficient severity is present, conservative

---

* to enhance readability of the text, "his" is used where "his/her" is intended. The choice for "his" instead of "her" is not arbitrary, but based on the higher prevalence of sciatica among male subjects.

therapies may be implemented to alter the spontaneous course of the disease. A systematic review was undertaken to show **which conservative therapies are proven to be effective for sciatica**. It is reported on in chapter 6.

Particularly in the initial stages of sciatica, bed rest seems to be a prominent therapeutic option. Evidence in support of bed rest seems to be lacking, so the question remains **whether a period of bed rest cures sciatica?** To answer this question the most common bed rest regimen, bed rest with toilet and shower facilities for a duration of 14 days, was compared with watchful waiting in a randomized controlled trial. Chapter seven highlights the background, methods and results of this trial and offers a discussion.

A doctor faces three challenges, when confronted with a sciatica patient. Can the cause be established, can it be treated, and can we predict the outcome? The latter becomes even more relevant, when the foremost determinant of the outcome is the natural course. **Can clinical findings predict the outcome in sciatica patients?** was the elementary question of the prognostic study reported in chapter 8. The outcome was considered in terms of subjective improvement over two and twelve weeks, functional status, low back and leg pain and work absenteeism.

Eventual discectomy might be regarded as another type of outcome in the sciatica patient. Chapter 9 features an investigation of the possibilities to predict eventual operation for sciatica due to disc herniation at an early stage. If a prediction would be possible this might have direct consequences for patient management, especially in view of the final question: **When should the sciatica patient undergo lumbosacral discectomy?**

A general discussion (chapter 10), a reference list and a summary will conclude this thesis.

The diagnosis of sciatica

2

A systematic review of the value of history and physical examination

---

ABSTRACT

The literature was systematically reviewed to investigate the value of history and physical examination in the diagnosis of sciatica due to disc herniation. Also, population characteristics and study design features affecting the diagnostic properties of the history and physical examination items were reviewed.

The initial clinical diagnosis of lumbar nerve root involvement due to disc herniation is based on a combination of diagnostic tests: the history and the physical examination. The literature provides several reports on the diagnostic value of these tests. Not all statements, however, are directly relevant to clinical practice and, more importantly, a number of population characteristics and study design features may greatly influence the values reported.

The literature for the period 1965-1994 was searched. 37 studies meeting the selection criteria were retrieved and were systematically and independently read by three readers to determine diagnostic test properties using a standard scoring list to determine the methodological quality of the diagnostic information. A meta-analysis was performed when study results were sufficiently homogeneous to allow statistical pooling.

Few studies investigated the value of the history. Pain distribution seemed to be the only useful history item for determination of the level of disc herniation. Of the physical examination signs the SLR was the only sign consistently reported to be sensitive for sciatica due to disc herniation. However, the sensitivity values varied greatly, the pooled sensitivity and specificity being 0.85 and 0.52, respectively. The crossed SLR was the only sign shown to be specific; the pooled sensitivity and specificity were 0.30 and 0.84 respectively. There was considerable disagreement on the specificity of the other neurological signs (paresis, sensory loss, reflex loss). All studies were biased in one way or another with three major types of bias being predominant: sequential-ordering bias, spectrum bias, and diagnostic/test review bias. Several other methodological drawbacks limited the generalizability of the study results. In general, test sensitivity is most likely to have been overestimated and test specificity underestimated.

The literature on the diagnostic value of history and physical examination in the diagnosis of sciatica due to lumbar disc herniation is subject to a sequential-ordering bias while most studies suffer from other types of bias as well. Furthermore, the literature on numerous signs and symptoms is scarce or absent. Additional diagnostic research involving sciatica due to disc herniation is recommended, especially in patients not selected by referral.

INTRODUCTION

In any patient with pain irradiating into the leg, the presence of lumbosacral nerve root involvement will be considered by the general practitioner. In the event of nerve root involvement the pain is referred to as sciatica. The initial diagnosis will be based on the history and physical examination¹⁸. Many studies describe the diagnostic value of history and physical examination findings for nerve root involvement and more specifically nerve root compression mostly in terms of the sensitivity and specificity. Concise overviews have been compiled regarding the value of clinical findings in the low back pain patient. Andersson and Deyo reported on the history and physical examination in patients with herniated lumbar discs¹. Yet, a systematic review of the value of clinical findings in patients suspected of sciatica due to disc herniation was lacking. A review was considered systematic if (1) the literature was acquired in a reproducible fashion with minimalisation of selection bias, (2) the diagnostic studies' methodology was assessed in a standardised way and (3) the results could be summarized quantitatively (=meta-analysis).

Therefore, our study was designed to systematically review the value of such findings. Moreover, the methodological difficulties in diagnostic studies of clinical findings were focussed on by systematically reviewing population characteristics and study design features.

MATERIAL AND METHODS

The literature was searched for eligible studies using Medline (the period 1965-June 1994) employing the search terms "diagnosis" in combination with "radiculitis or root compression or disc/disk herniation or herniated disc/disk or disc/disk prolapse or hernia nuclei pulposi or sciatica" (n=686) and thesaurus-terms "Intervertebral-Disk-Displacement/diagnosis" and "Sciatica/diagnosis" (n=1511). Also the automated indexes of the *Nederlands Tijdschrift voor Geneeskunde* and *Huisarts en Wetenschap* were searched and reference lists and experts were consulted.

Only reports written in English, French, German or Dutch, that dealt with sciatica and/or disc herniation and that reported tests results of various history and physical examination items were included. Excluded were studies dealing exclusively with other diseases than sciatica (e.g. low back pain) or with other diagnostic tests than the components of history and physical examination or that lacked any presumed gold standard.

A final selection of studies meeting the criteria was read and assessed independently by three assessors (AK, MdK, PV) using a standard form (see table 1 and the appendix). If there was any disagreement between the readers, a consensus was reached through discussion.

The scoring system was based on well accepted methodological criteria in the literature¹⁹. Major determinants of the generalizability of study results (the "external validity") were considered to be the spectrum of patients included (reflected in items

A, B and C; 30 points)^{20 21 22 23} and the technical details of the test (item D; 10 points)^{23 20}. A third requirement was the statistical power of the study (item H; 10 points). Major determinants of diagnostic accuracy (the "internal validity") were the appropriateness and description of the reference test (item E), the extent of verification bias (item F but also reflected in the spectrum, especially item C) and the independence of index and reference test interpretation (item G) (from Irwig *et al.*²³ and Begg²⁴). Apart from valid and generalizable results an informative report requires complete and correct data presentation²⁰ (item I; 10 points).

The test properties were statistically pooled only if more than 3 studies reported on the properties of a particular test. The Dersimonian and Laird random effects method was used (with calculation of the 95 percent confidence interval (95% CI) using the Diamond formula)^{25 26}.

RESULTS

Table 1 shows the diagnostic characteristics of the history and physical examination items for those studies for which all test characteristics (see item I) could be calculated.

The history

The only history items of which the diagnostic value was reported in the literature are the presence and distribution of pain and the presence of sensory symptoms. The reported sensitivity of the pain description varied from 0.38 in a study of patients with disc herniation verified at operation²⁷ to 0.99 in non-surgical myelography-verified patients. The specificity of these descriptions was low (0.06-0.50). Using pain drawings to register the pain was found to have a higher specificity (0.84)²⁸. Sensory symptoms were found to be neither sensitive (0.30) nor specific (0.58)²⁹.

The physical examination

Eight studies provided adequate test characteristics of the **straight leg raising test (SLR)**. However, in all studies the patient inclusion was based on the availability of the gold standard. Moreover, five of eight studies focussed on surgical candidates with surgical findings as the gold standard. The pooled sensitivity of the straight leg raising test for the non-surgical studies was 0.85 (95% CI 0.38-0.98) with a specificity of 0.52 (95% CI 0.26-0.76). For study populations of surgically treated patients the pooled sensitivity was 0.91 (95% CI 0.78-0.97) and the specificity 0.32 (0.17-0.52). One study included a specification of the angle (60 degrees) in their requirements for a positive SLR test resulting in a specificity of 0.84 and a sensitivity of 0.30³⁰.

Table 1. Characteristics of the history and the physical examination

Diagnostic test	gold	Population	Sens	Spec	PV+	Pr	Study	Sum
HISTORY								
Presence and distribution of pain								
pain drawing	myelo	surgical	46	84	71	46	²⁸	90
sciatic pain	myelo	non-surgical	99	6	74	74	²⁹	55
pain in groin or iliac crest	myelo	surgical	11	46	15	46	²⁸	90
<i>pain below knee</i>	<i>surgical</i>	<i>surgical</i>	38	50	86	89	²⁷	56
Sensory symptoms	myelo	non-surgical	30	58	57	74	²⁹	55
PHYSICAL EXAMINATION								
Straight leg raising								
	myelo	surgical	86	32	51	46	²⁸	90
	CT	non-surgical	45	78	81	67	³⁰	71
	combination	surgical	70	70	87	74	³¹	61
	myelo	non-surgical	98	44	83	74	²⁹	55
straigh leg raising < 60	CT	non-surgical	30	84	80	67	³⁰	71
<i>Typical Lasègue's sign</i>	<i>surgical</i>	<i>surgical</i>	79	88	67	58	³²	81
	<i>surgical</i>	<i>surgical</i>	97	26	87	86	³³	71
<i>Typical Lasègue's sign</i>	<i>surgical</i>	<i>surgical</i>	89	17	87	87	³⁴	63
	<i>surgical</i>	<i>surgical</i>	93	17	79	77	³⁵	62
	<i>surgical</i>	<i>surgical</i>	96	10	90	89	²⁷	56
Crossed straight leg								
	combination	surgical	23	98	97	74	³¹	61
	myelo	non-surgical	43	97	98	74	²⁹	55
	<i>surgical</i>	<i>surgical</i>	23	88	92	86	³³	70
	<i>surgical</i>	<i>surgical</i>	24	100	100	87	³⁴	63
	<i>surgical</i>	<i>surgical</i>	33	95	98	89	²⁷	56
Paresis iliopsoas								
	myelo	non-surgical	10	97	91	74	²⁹	55
tibialis anterior	myelo	non-surgical	54	89	93	74	²⁹	55
Gastrocnemius	myelo	non-surgical	13	100	100	74	²⁹	55
	<i>surgical</i>	<i>surgical</i>	47	52	57	58	³⁴	63
<i>extensor hall. long.</i>	<i>surgical</i>	<i>surgical</i>	62	50	91	89	²⁷	56
Atrophy calf atrophy								
	myelo	non-surgical	29	97	97	74	²⁹	55
	<i>surgical</i>	<i>surgical</i>	38	50	86	89	²⁷	56
Sensory deficits								
	<i>surgical</i>	<i>surgical</i>	60	57	66	58	³⁴	63
	<i>surgical</i>	<i>surgical</i>	28	65	87	89	²⁷	56
Depressed reflexes								
APR / KPR	myelo	surgical	54	68	40	46	²⁸	90
APR	myelo	non-surgical	48	89	92	74	²⁹	55
	<i>surgical</i>	<i>surgical</i>	54	60	92	89	²⁷	56
KPR	<i>surgical</i>	<i>surgical</i>	4	92	72	86	³³	70
KPR	<i>surgical</i>	<i>surgical</i>	14	65	77	89	²⁷	56
CLUSTERS								
"Radiculopathy"	CT	non-surgical	37	88	86	67	³⁰	71
Probable diagnosis	myelo	surgical	49	78	98	68	³⁶	49
Definite diagnosis	myelo	surgical	64	65	58	48	³⁶	49
Probable diagnosis	combination	surgical	53	92	62	20	³⁷	44
Cluster of at least two tests	combination	surgical	27	97	96	76	³⁸	35
"Root compression"	<i>surgical</i>	<i>surgical</i>	71	100	100	58	³⁴	63
"neurological disturbances"	<i>surgical</i>	<i>surgical</i>	96	15	79	77	³⁵	62
Probable diagnosis	<i>surgical</i>	<i>surgical</i>	68	0	99	68	³⁶	49
Definite diagnosis	<i>surgical</i>	<i>surgical</i>	66	29	54	48	³⁶	49

Note to table 1- If studies used operative findings as the reference test, this was considered to introduce a selection bias, as subjects were likely to have been clinically selected on the basis of disease severity, resulting in an indication for operation. Because this bias may be considerable, the test characteristics collected from studies subject to this bias are presented in italics.

The order in the studies of each test reflects the methodological score; first in line is the study with the highest score.

In the non-surgical population the pooled sensitivity and specificity of the crossed straight leg raising were 0.32 (95%CI 0.16-0.54) and 0.98 (95%CI 0.94-0.99), respectively. The test characteristics of the reversed SLR or femoral nerve stretching test could not be adequately calculated. There are no diagnostic studies determining the value of tender points on pressure along the sciatic nerve, typical pain or paraesthesias on compression of jugular veins, typical pain on passive hyperextension and lateroflexion of the spine, or typical pain on neck flexion.

Decreased muscle strength is reported as both a non-sensitive and a non-specific sign in a study of a surgical population, while another study reports decreased strength of the iliopsoas, anterior tibial or gastrocnemius muscle to be specific (over 0.89). This pattern is reflected in the variance in the reported specificity of muscle atrophy.

Sensory deficits are both non-sensitive and non-specific. No distinction was made, however, between the vital and gnostic qualities of the sensibility.

Depressed reflexes are non-sensitive signs; especially depressed knee tendon reflexes are rarely present in patients suffering from herniated discs. There is considerable disagreement regarding the high specificity of the knee and ankle tendon reflexes.

The value of signs and symptoms of the cauda equina syndrome remains incompletely understood.

The value of combined tests

In clinical practice, combinations of tests rather than single test results are used to obtain greater sensitivity and/or specificity. Various combinations have been investigated and their diagnostic properties varied greatly. For example, Hirsch³⁵ defines a set of neurological disturbances that has a 0.96 sensitivity with a 0.15 specificity, while Hudgins³⁸ defines a set with 0.27 sensitivity with a 0.97 specificity.

The methodology

37 studies were retrieved. Their methodological scores are presented in table 2. The methodological scores varied greatly (mean 51, standard deviation 16, range 28-90). All studies were carried out in secondary or tertiary care patient populations; 69% in surgical and 31% in non-surgical departments. Thirty percent of the studies provided a description of the clinical patient characteristics, while demographic details such as age and sex were described by 54%. Forty-one percent used a prospective design. 43% of studies gave adequate details on selection criteria. The spectrum of disease was thoroughly described in 51% of the studies, but in 16 of these 19 studies the spectrum was considered too narrow. The intended use of the index test was described.

Table 2. Methodological scores per study

Study	A	B	C	D	E	F	G	H	I	gold standard
Uden ²⁸	10	10	10	10	10	10	10	10	10	myelography
Kosteljanetz ³²	7	8	10	9	12	10	5	10	10	operation
Haldeman ³⁰	7	5	10	4	10	10	5	10	10	CT
Spangfort ³³	10	3	10	2	10	10	5	10	10	operation
Jensen ³⁹	5	8	10	8	7	10	5	5	10	combination
Jensen ⁴⁰	5	8	10	9	5	10	5	5	10	operation
Morris ⁴¹	5	4	10	9	7	10	10	10	2	operation
Hakelius ⁴²	2	1	10	6	10	10	5	10	10	operation
Kosteljanetz ³⁴	5	8	10	13	7	0	5	5	10	operation
Hirsch ³⁵	7	1	10	4	5	10	5	10	10	operation
Hudgins ³¹	0	8	10	3	5	10	5	10	10	combination
Hakelius ⁴³	5	5	10	5	10	10	5	10	0	operation
Kortelainen ⁴⁴	4	6	10	3	5	10	5	5	10	operation
Meylemans ⁴⁵	10	8	10	2	12	10	5	0	0	CT myelo
Knutsson ²⁷	10	1	10	0	5	10	5	5	10	operation
Kerr ²⁹	4	3	10	3	5	10	0	10	10	myelography
Ljunggren ⁴⁶	10	8	0	6	10	10	5	5	0	operation
Guardjian ⁴⁷	7	1	10	0	5	10	5	5	10	operation
Weise ⁴⁸	7	6	0	9	5	0	10	5	10	CT myelo
Wei ⁴⁹	7	5	0	4	9	10	5	0	10	operation
Aejmelaeus ³⁶	7	5	10	2	5	0	5	5	10	myelography
Portnoy ⁵⁰	5	4	0	5	5	10	5	5	10	operation
Mann ⁵¹	0	0	10	4	0	10	10	0	10	combination
Postacchini ⁵²	2	4	0	5	5	0	5	10	10	operation
van Damme ⁵³	0	8	10	0	0	10	5	5	0	combination
Jonsson ⁵⁴	10	4	0	2	5	10	0	5	2	operation
Aho ⁵⁵	10	3	0	5	5	0	5	5	2	operation
Christodoulides ⁵⁶	2	5	0	1	5	10	5	5	2	operation
Hudgins ³⁸	0	6	10	4	0	10	5	0	0	combination
Albert ⁵⁷	10	1	0	0	5	10	0	5	2	operation
Aronson ⁵⁸	4	1	0	0	5	10	5	5	2	operation
Borovich ⁵⁹	7	3	0	0	5	10	5	0	2	operation
Woodhall ¹⁷	0	0	10	0	7	10	0	5	0	operation
Katznelson ⁶⁰	2	1	0	2	5	10	5	5	0	operation
Keegan ⁶¹	0	1	0	9	5	10	5	0	0	operation
Kostuiks ⁶²	7	1	0	0	7	10	5	0	0	operation
Wiberg ⁶³	0	2	0	1	5	10	5	5	0	operation

Criteria for methodological evaluation of the selected studies (see appendix I)

A	patient description	F	application of reference test
B	study population	G	independence of interpretation
C	diseased and non-diseased included	H	sample size
D	technical quality of index test	I	data presentation
E	technical quality of reference test		

in 16% of the studies, with information on normality, standardization and reproducibility in 38%, 30% and 8%, respectively. The contribution of single test results to the overall cluster of tests was established in eight percent. The reference test was appropriate in 89% of studies and performed in all patients in 87% of the studies. However, details on normality, standardization and reproducibility of the reference test were missing in 76%, 100% and 8% respectively. In the majority (89%) of studies it was unclear whether the clinical findings and the gold standard were assessed independently. Only 30% of the studies included more than 35 patients in diseased and non-diseased groups. Almost half of the studies (46%) did not allow the reader to calculate the necessary diagnostic test characteristics.

DISCUSSION

Very little attention has been paid to the investigation of the value of the history. The reported diagnostic value of the pain distribution greatly varies from study to study, but it seems to be the only sensitive sign as well as a useful sign of the level of disc herniation. The diagnostic value of a thorough history remains unclear, however.

Nerve root tension signs have been more thoroughly evaluated. The primary gain of the SLR was its sensitivity, even though it varied greatly across studies. A work-up bias may have lead to overestimation of this sensitivity. The value of all other nerve root tension signs is unknown. There is disagreement in the literature with regard to the value of decreased muscle strength, sensory loss or depressed reflexes as signs of nerve root involvement. The value might well be their high specificity. Likewise it is unclear whether the physical examination adds much to the diagnostic information of the history. Requiring a certain angle at which the pain is provoked for the test to be positive seemed to add little to the specificity but diminished the sensitivity. The variance of the SLR properties across studies may partly be explained by the great variance in the definition of normal and abnormal test results and the technique of straight leg raising or Lasègue's sign. In different settings the so called positivity criterion²⁴ may be differently applied. The value of requiring typical dermatomal pain distribution, pain below the knee, or hip fixation during SLR has not been studied. These limitations would only be desirable if they would cause a significant increase in specificity.

Omissions in the literature retrieval may produce biased review results. To assure comprehensiveness of the review results a thorough automated literature search was attempted with perusal of reference lists and suggestions by experts in the field as a means of detecting omissions in the automated retrieval (e.g. in the field of physical therapy). Searches for diagnostic reviews were troubled by the fact that studies supplying relevant diagnostic information were not always labelled as "diagnostic". This is in contrast with the more straightforward searches for reviews of randomized controlled trials (RCTs). Such trials seem more clearly labelled and are hence more easily detected by automated searches.

A second point of concern in the review may have been the methodological

assessment. Because some of the studies retrieved were not specifically designed as diagnostic studies, the methodological quality of these studies is likely to be inferior to the scores of "truly" diagnostic studies. Our explicit intention was not to assess the quality of the study or report but to assess the quality of the diagnostic information presented. The results of the poorer studies were not omitted to illustrate the variance both in quality and outcome of the studies that have created our state of knowledge. Though the heterogeneity of the studies complicated the scoring, the problems encountered in the standardised comparison of studies could generally be resolved by a consensus of the three readers. The order of the results in table 1 and 2 is determined by the methodological score. A different weighing system might have led to a different order. The application of a differently weighted scoring system focusing on determinants of test accuracy²³ indeed produced a different order, but it resulted in the same twelve studies at the top of the table. The weights applied to each item and the resulting sum scores were not used to exclude studies. Moreover, because of the small number of studies of each test it was considered inappropriate to relate the quality of the studies (sum scores) to the test characteristics. The effect of the particular weighing system used will be limited, therefore.

Magnetic Resonance (MR) imaging, computed tomography, myelography and operative verification were all considered adequate gold standards. Though MR imaging seems to have become the state of the art radiological diagnostic procedure for nerve root compression, there is no definitive evidence of its superiority to the other procedures mentioned^{64 65}. Though surgical verification was the predominant gold standard, little is known about the accuracy and consistency of operative findings. Unanswered questions are whether observations during surgical procedures indeed reflect particular pathologic states of nerve root and whether specific operative techniques with varying degrees of exposure of the operative field influence the validity and accuracy of the observations. Also, operation invariably leads to sequential-ordering bias^{21 66} (also called work-up or verification bias) occurring when positive test results are more likely to lead to verification by a reference test than negative test results^{22 67}. It is also present if patients with the disease are more likely to undergo the radiological gold standard test. Only a few studies avoided this type of bias. Even more importantly, all studies reveal another cause for a sequential-ordering bias partly associated with the former type: all populations consisted of secondary (or tertiary) care patients, only. It seems logical that patients with a "positive" history and physical examination were preferentially referred to secondary care. A consequence is, however, that the interpretation of a set of diagnostic tests by the primary care physician determined which patient eventually ended up in the secondary care study of the same diagnostic test. As a result a profound impact of the referral pattern on the test properties in secondary care may be expected. The sensitivity will then be overestimated while the specificity will be underestimated^{22 24}. A related type of bias is spectrum bias occurring when the test performs differently in the research population than it would in customary care⁶⁸. As a consequence, sensitivity and specificity values cannot be applied directly to other populations. The majority of studies reviewed included surgical candidates only (69%). This has led to a spectrum

bias because subjects are likely to have become surgical candidates because of a "positive" history and examination. Thus, the patients had more extreme and severe manifestations than in customary care. The effect of this spectrum bias might be reflected in the higher pooled sensitivity and lower pooled specificity of the SLR in the studies with operated patients. The frequently encountered inclusion of diseased patients only leads to such a bias as well. Other less overt details of the spectrum were often poorly described. This is unjustified when considering that, for example, ignoring a demographic characteristic like age may cause spectrum bias as well. With increasing age both prevalence and spectrum of sciatica change as well as the interpretation of test results (more difficult in the youngest³³ and the oldest of the spectrum). Only a few studies adequately described the spectrum, but in all it was too much narrowed down to a more severely affected population.

A **diagnostic or test review bias** may have occurred in those 89% of the studies that did not report on an independent assessment of the index and reference test. Either the observations were unblinded indeed or the authors thought blinding of observations was such a logical requirement that they felt it unnecessary to mention it²⁰.

The **precision and consistency** of the test property estimations may be unclear. None of the studies described the statistical precision of the estimations using variability measures such as standard deviation and confidence intervals, notwithstanding the fact that small sample sizes were a frequent finding. The lack of standardised testing, a design feature frequently encountered, negatively affects the test consistency or reliability. The consistency of the tests was poorly described and only one study investigated it in conjunction with the test accuracy³⁴. Because both accuracy and consistency of diagnostic tests can be population specific it is advocated that both be determined in a given research population. Also, standardised testing and an adequate definition of normality should be a requirement. Studies of clusters of tests may be valuable but the clusters in this review differed greatly in sensitivity and specificity and were subject to the same biases as the other investigations.

The evaluation of the diagnostic value of the history (particularly the value of specific pain descriptions/qualities, and of aggravating circumstances and the overall value of the history) and the physical examination in sciatica due to disc herniation warrants further investigation. Such investigations are to be preferably carried out in patients not selected on the basis of history and examination findings. Diagnostic tests should be applied to all patients in a standardised and well described manner. The reproducibility can be determined concurrently. It is important that appropriate gold standard procedures are applied to all patients irrespective of diagnostic test results. Multivariable analysis can be helpful in determining both the independent and combined value of diagnostic tests. Clusters might be constructed by first determining the tests with significant, independent diagnostic value, include these in a clinical prediction rule, and evaluate this rule in another, clinically similar, population⁶⁹.

APPENDIX

A Patient description (maximum 10 points). If both demographic (at least age and sex) and clinical patient characteristics (at least two of the following: previous episode(s) of sciatica, duration of disease, neurological dysfunction) were adequately described, 5 points per category were scored, 2 points for inadequate description and 0 points for no description.

B Study population (maximum 10 points). 3 points were scored for a prospective design, 3 for an adequate description of selection criteria and 4 for an adequate description of patient spectrum. As there are no generally accepted pathologic stages of sciatica due to disc herniation, the clinical and the co-morbid components of the spectrum but not the pathologic component of the spectrum was assessed. If the selection criteria were regarded as inadequately reproducible, 1 point was scored. If the spectrum was regarded as too homogeneous to assure generalization, 2 points were given. 0 points were given for "no description".

C Diseased and non-diseased included (maximum 10 points). 10 points were given if both diseased and non-diseased subjects were included, 0 for inclusion of only diseased subjects.

D Technical quality of index test (maximum 15 points). Description of the use of the index test had to specify both the intended population and the diagnostic goal to score 2 points. If the testing was standardised, 3 points were scored. If the description of normal and abnormal test results allowed reproduction of the study and if it was according to professional standards, 5 points were scored. For an inadequate description, 2 points were scored. If the observer variability was assessed or described 3 and 2 points were scored respectively. An additional 2 points could be scored if a description of the contribution of test results to the overall cluster of tests was provided.

E Technical quality of reference test (maximum 15 points). Either operative findings, myelography, CT-myelography or MRI were regarded as appropriate reference tests scoring 5 points. If there was an adequate description of standardized testing, the definition of normality and the reproducibility 2, 5 and 3 points were scored respectively.

F Application of reference test (maximum 10 points). If the same reference test was applied in all patients 10 points were scored. If the reference test was not the same for all patients 5 points were scored.

G Independence of interpretation (maximum 10 points). If it could be definitely established that index and reference test were independently assessed 10 points were given. If there was a degree of uncertainty 5 points were given.

H Sample size (maximum 10 points). 10 points were scored if both diseased and non-diseased groups contained 35 patients or more, 5 points if only one of the groups contained 35 patients or more and 0 points if there were less than 35 patients per group.

I Data presentation (maximum 10 points). 10 points were given if all fields of a diagnostic test matrix could be filled and 2 points if only sensitivity and prevalence could be calculated.

3

Technique and consistency of the history and the physical examination in patients with irradiating pain into the leg

---

ABSTRACT

This section features a description of the technique of the history and physical examination in patients with irradiating pain into the leg. Subsequently it offers a report of a cross-sectional study to investigate the consistency of signs and symptoms of sciatica in primary care patients with irradiating pain into the leg.

The literature does not report on all the clinical tests for nerve root involvement. All investigations were in low back pain patients, only. Often, little information on either examination technique, proportion of positive test results or clinical patient characteristics is provided.

A random selection of 91 patients was investigated by a neurologist-resident couple. Agreement percentages, proportions of positive test results, and kappas were calculated.

The kappa of the overall conclusion after the history was 0.40, increasing to 0.66 after physical examination. Kappas were good for decreased muscle strength and sensory loss (0.57-0.82), intermediate for reflex changes (0.42-0.53), and poor for the examination of the lumbar spine (0.16-0.33). The straight leg raising (SLR), crossed SLR, Bragard's sign and Naffziger's sign were the most consistent nerve root tension signs (> 0.66).

Two clinicians disagreed on the presence of nerve root involvement in 1 of 4 patients after history and in 1 of 5 patients after physical examination. Asking for an increase of pain on coughing/straining/sneezing, a feeling of coldness in the legs and urinary incontinence and the investigation of muscle strength, sensory loss, reflex changes, SLR and Bragard's sign provide the most consistent results.

INTRODUCTION

In a comprehensive review⁷⁰ the efficacy of surgical therapy (successful outcome) for lumbar disc herniation was estimated at 75%. It has been suggested that poor patient selection for surgery explains a fair deal of the poor patient outcome⁷¹. Apart from financial motives⁷¹, poor patient selection for surgery may be explained by diagnostic difficulties. A fairly definitive morphological diagnosis of a lumbar disc herniation can be achieved⁷². The (functional) diagnosis of nerve root disorders (often but not solely due to disc herniations) is much more troublesome, not merely because a gold standard for nerve root involvement is lacking. It may partly explain the possible variation between clinicians' diagnostic decisions concerning nerve root involvement and disc herniations¹⁹. As mentioned, the variations in the outcome of diagnostic procedures may also affect the outcome of therapy and the results of trials in the sciatica patient. Hence, it is important to know the magnitude of and the reasons for such variations. The purpose of this paper is, firstly, to describe the initial history and physical examination in the sciatica patient with specification of positive and negative test results (often imprecise in previous studies¹ and, secondly, to document the variation in the test results attained. The study goal in other words was to investigate to the consistency (i.e. reproducibility) of the signs and symptoms of lumbosacral nerve root involvement?

Several papers have reported on the consistency of diagnostic test results used in the evaluation of patients with low back pain and sciatica. Nelson *et al.*⁷³ described the overall agreement of the history and physical examination, but information on diagnostic testing technique and consistency measures like kappa (κ) values was lacking. Such information was provided in subsequent reports by Waddell *et al.*⁷⁴ and McCombe *et al.*⁷⁵ but they investigated low back pain patients. (generally without pain in the leg). Also, proportions of positive test results, important in the interpretation of agreement percentages and kappa values were not described. In this paper, a consistency study is reported in a research population with a spectrum of disease relevant for patients presenting to the general practitioner.

PATIENTS AND METHODS

The study was performed in Maastricht and surrounding villages. Patients were referred by 50 general practitioners for inclusion in a diagnostic study of the value of history and physical examination in patients suspected of lumbar nerve root involvement. Inclusion required a new episode of pain irradiating into the leg below the gluteal fold, of sufficient intensity to justify 14 days of bed rest as a therapy. Exclusion criteria were contra-indications for Magnetic Resonance Imaging, previous spinal surgery, and previous bed rest therapy for the current episode. Beforehand, a standardized testing technique had been agreed on and had been practised together. In a first random sample of the patients referred, the history and physical examination were performed independently by a neurologist (MdK) and a neurological resident

(PV). The order of who was first and second observer was varied at random. Both observers consecutively examined the patients in the same office room. A second sample of patients was assessed by the neurologist (MdK) and a third neurological resident (AdW), who had not had a special training in the testing technique.

The history preceded the physical examination. The order in which the tests were performed was the same for both observers. The testing technique was as follows:

Technique of history taking

- * Point out the distribution of the pain with one finger. The observer assessed whether the distribution was "typically radicular" (yes/no/not applicable).
- * How does the pain in the leg change on coughing/sneezing/straining, sitting/standing or walking/lying (increase/no change/decrease/not applicable)?
- * Have you noticed a weakness or decreased muscle strength/a numbness or sensory loss/a coldness/paraesthesias in the leg (yes/no/sometimes/not anymore)?
- * Do you have a disturbed feeling of micturition/any urinary incontinence (yes/no)?
- * How many previous episodes of pain in the back/pain in the leg have you had?

Technique of the physical examination.

With the patient standing erect and firmly on both feet with the arms along the body:

- * Decreased lumbosacral lordosis (yes/no): assessment of a decrease in lumbosacral sagittal curvature.
- * Antalgic posture (yes/no): while the patient attained the most comfortable standing posture, it was determined whether there was a convex scoliosis and a flexed knee on the afflicted side⁷⁶.
- * Lumbar scoliosis (yes/no): if there was a deviation in the coronal plane, the side of convexity of the scoliosis was recorded.
- * Lumbar anteflexion: the level of the sacroiliac joint and the level 10 cm rostral to the sacroiliac joint were marked. Then, with the patient trying to reach the floor with the extended fingers of both hands without bending the knees, the increase in distance between the two marks was noted in millimetres. This was labeled the Schöber (P4). The distance between the floor and the third digit was measured (finger floor distance).
- * Tenderness on percussion over the lumbar vertebrae: (yes with radicular radiation/yes without radicular radiation/no): the vertebrae were firmly percussed with the fist and pain distribution was recorded.
- * Paravertebral hypertonia (yes, painful with radicular radiation/yes, painful without radicular radiation/yes, but not painful/no): the paravertebral musculature was palpated for indurations.
- * Kemp's sign (no pain/ipsilateral pain to the leg/contralateral pain to the leg/ipsilateral pain in the back/contralateral pain in the back): the back was hyper extended and simultaneously lateroflexed by the observer. The localization of eventual pain was registered^{77 78}.
- * Naffziger's sign (typical pain in the leg/non typical pain/no pain): both internal jugular veins were manually compressed for a minimum duration of ten seconds.

The patient was asked whether a blushing or pressure sensation in the head occurred. If so, the patient was asked for any other sensations^{79 76 78}. A typically dermatomal pain was recorded as a positive test.

- * Head flexion symptom (no pain/pain in the back/pain in the leg): typically dermatomal pain in the leg upon head flexion is noted (also referred to as Hyndham sign⁷⁸).
- * Walking on heels (normal/abnormal): this was abnormal if any other part of the foot than the heel touched the floor when the patient walked on his heels.
- * Walking on the toes (normal/abnormal): this was abnormal if the heel touched the floor when the patient walked on his toes.
- * Knee-bending (normal/abnormal): the patient was asked to stand on one leg and bend the knee (90°). Any failure of flexion was registered as abnormal.

With the patient supine on a flat investigation table with a single pillow:

- * Decreased or painful mobility of the hip: The leg was flexed (90°) in the knee and the hip and then passively endo- and exorotated in the hip. Pain and/or decreased mobility of the hip was noted.
- * Straight leg raising test = SLR (negative/pain in the back only/pain in the leg in a non-dermatomal distribution/pain in the leg in a dermatomal distribution): the extended leg was gently flexed in the hip by the observer. The Plurimeter-V, placed on the caudal part of the tibial bone, was used to measure the angle at which pain first occurs⁸⁰. Additionally, it was noted whether the pain was typically dermatomal and whether fixation in the hip joint occurred.
- * Pain on leg extension with a 90° flexion of the knee: (yes/no): in addition to the straight leg raising, "bent" leg raising was assessed (as proposed by Lasègue's pupil Forst¹³ see figures 1 and 2).

Figure 1. Lasègue's sign
(taken from J.J. Forst's thesis, 1881)

Figure 2. "Bent" leg raising sign

- * Bragard's sign: (pain/no pain): when the SLR became painful or restricted, the leg was lowered ($\pm 5^\circ$) and the ankle was dorsiflexed¹⁷. An increase of the pain on SLR was noted. The first description of this sign has been attributed to others than Bragard such as Fajerstajn and Roch⁸¹.

- * Crossed straight leg raising (negative/pain in the back only/pain in the leg in a non-dermatomal distribution/pain in the leg in a dermatomal distribution): the same procedure as for the SLR. The angle of hip flexion at which pain in the contralateral leg occurred was measured¹⁶.
- * Decreased strength (5-4-3-2-1): the extensor hallucis longus, anterior tibial, long and short peronei, gastrocnemius, hamstring, quadriceps, iliopsoas, and gluteus maximus et minimus muscles were tested adopting the technique and the rating scale described by the Medical Research Council⁸².
- * Atrophy: the circumference of the leg was measured (in centimeters), only if inspection raised the suspicion of atrophy. This was done halfway between lateral epicondyle of the femur and the great trochanter, and halfway between the fibula head and the lateral malleolus.
- * Sensory loss - light touch (disturbed/not disturbed): a dermatomal map was used and dermatomal keypoints were tested⁸³. The dermatomes were tested bilaterally and simultaneously by softly striking the skin. The patient (with eyes closed) was asked if the feeling clearly differed between left and right.
- * Sensory loss - pain (disturbed/not disturbed): the patient with eyes closed was asked whether the gentle pressure exerted by the observer through the Neurosens (a plastic stick with a sharp and a blunt side) was sharp or blunt. The sharp and blunt ends were applied at random.
- * Reflexes (+4/+3/+2/+1/0/-1/-2/-3/-4): The reflexes were tested as described by Bronisch⁸⁴. The Mayo clinic grading system was used. The adductor reflex was tested by hitting the index finger placed on the insertion of adductor tendon at the medial femoral epicondyle. The knee tendon reflex was tested with both knees flexed 30°. The ankle tendon reflex was tested with the observer's back to the patient and holding the leg (hip and knee flexed 90°) in the armpit. The patient was asked to cough or perform Jendrassik's maneuver if there was any difficulty in eliciting the reflexes.

With the patient in the lateral recumbent position on a flat investigation table

- * Femoral nerve stretch test (no pain/pain in the back/pain in the leg): After a 15 degree retroflexion of the straight leg in the hip the knee was bent. If the latter procedure intensified or provoked the patient's complaints in the leg, the test was considered positive⁸⁵.
- * Tenderness of sciatic nerve (no pain/local tenderness/typical radiating tenderness at 1 point/typical radiating tenderness at two points): the sciatic nerve was palpated along its course in the buttock and upper leg ("Valleix points").

Data analysis

The percentage of positive test results for both observers, the percentage of agreement and the kappa for inter-observer consistency (agreement beyond chance - with asymptotic standard deviations) was determined using the Statistical Package for the Social Sciences (SPSS^R).

RESULTS

338 patients were referred by the 50 participating general practitioners. Upon ascertainment of the selection criteria by the investigators 284 patients could be included in the diagnostic study. In a one in three random sample (n= 91) paired observations were made. The average age was 46 years (SD= 11.2) ranging from 16 to 77 years. 53% were male and 48% female. 57% of the patients had a duration of pain in the leg of less than 30 days, 20% of 31-89 days and 23% of 90 days or more.

Table 1. The inter-observer variability of the history. (n=91)

Item	percentage agreement	percentage of positive test results (observer1/observer2)	kappa (SD)
typical radiation pattern of pain	80	(84/96)	0.24 (0.10)
increase of pain when coughing/ sneezing/ straining	81	(43/46)	0.64 (0.08)
increase of pain on sitting	73	(56/48)	0.49 (0.09)
increase of pain on walking	77	(51/59)	0.56 (0.08)
increase of pain on lying down	67	(59/64)	0.41 (0.09)
decreased strength	63	(35/25)	0.14 (0.10)
sensory loss/ numbness	70	(34/51)	0.41 (0.09)
coldness	82	(32/23)	0.56 (0.10)
paraesthesias	71	(56/57)	0.47 (0.08)
disturbed urinary passage	90	(7/8)	0.26 (0.17)
urinary incontinence	96	(8/10)	0.79 (0.13)
previous back pain episodes	91	(80/89)	0.67 (0.10)
previous leg pain episodes	68	(52/38)	0.44 (0.08)
positive history			
yes/doubtful/no	68	(70/60)	0.33 (0.09)
definitely yes	73	(70/60)	0.40 (0.10)

Note- The second column shows the proportion of negative and positive test results agreed on with all observations as the denominator, the third column (in brackets) the percentage of patients with a positive test result for first and second observer respectively, and the fourth column the kappa values with the standard deviation (SD).

Note to table 2- The 5th column mentions the consistency values in the literature followed by the reference number in brackets followed sometimes by a specification of the test (please see legend table 1 for the other columns). For 2 tests the kappas were based on 75 paired observations (in italics).

Table 2. The inter-observer variability of the physical examination (n=91)

Item	percentage agreement	percentage of positive test results (observer1/observer2)	kappa (SD)	literature (reference number)
decreased lordosis	76	(29/18)	0.33 (0.11)	
antalga	78	(14/30)	0.38 (0.11)	
vertebral percussion tenderness	69	(20/36)	0.32 (0.09)	
paravertebral hypertonia	79	(7/21)	0.16 (0.11)	0.11 (6)
abnormal walking on heels	88	(19/14)	0.57 (0.11)	-0.03 (6)
abnormal walking on toes	98	(5/8)	0.82 (0.12)	0.00 (6)
abnormal kneebending	97	(2/3)	0.66 (0.32)	
paresis	88	(24/21)	0.69 (0.09)	
extensor hallucis longus	96	(16/12)	0.82 (0.09)	0.65 (6)
peronei	93	(10/8)	0.59 (0.15)	1.00 (6)
anterior tibial	95	(16/11)	0.68 (0.12)	1.00 (6)
gastrocnemius	99	(3/2)	0.80 (0.20)	0.65 (6)
sensory loss	81	(35/36)	0.59 (0.09)	0.68 (6)
hypesthesia	88	(30/29)	0.71 (0.08)	
hypalgesia	78	(22/18)	0.57 (0.12)	
reflex differences				
ankle tendon	87	(13/20)	0.53 (0.12)	0.39 (6)
knee tendon	95	(5/4)	0.42 (0.21)	0.23 (6)
adductor tendon reflex	86	(0/7)	0.0	
absent reflexes				
ankle tendon reflex	86	(9/14)	0.52 (0.14)	
knee tendon reflex	100	(0/0)	-	
adductor tendon reflex	86	(13/3)	0.09 (0.12)	
positive straight leg raising				
typically dermatomal pain	85	(57/63)	0.68 (0.07)	0.36 (6);
any pain in the leg	79	(75/85)	0.36 (0.11)	0.66 (6)
SLR < than 45 degrees	82	(23/14)	0.43 (0.12)	0.56 (5); < 75° 0.33 (11) [†] 0.44 (11) ^{††}
positive Bragard	84	(55/63)	0.66 (0.08)	
positive crossed SLR	89	(10/14)	0.49 (0.14)	0.74 (4)
if SLR is positive	91	(17/17)	0.70 (0.14)	
positive reversed SLR	79	(24/30)	0.47 (0.10)	72% agreement (4)
positive Valleix pressure points	77	(7/21)	0.14 (0.11)	
positive Kemp sign	68	(27/37)	0.28 (0.10)	
positive Naffziger sign	85	(19/15)	0.71 (0.11)	
positive history & examination				
yes/doubtful/no (68%)	79	(67/63)	0.59 (0.08)	irritation 0.45 compression
definitely yes	84	(67/63)	0.66 (0.08)	0.62 (5)

A random sample of 75 patients was examined by the couple MdK-PV and 16 patients by MdK-AdW. For most tests the combined kappa for both couples did not differ from the kappa for the couple MdK-PV by more than 0.05. Only for the Valleix pressure points, Naffziger's sign and disturbed knee-bending was there a greater difference in kappas (MdK and PV's kappa was better by 0.07, 0.21 and 0.28, respectively). Therefore, except for these three tests a combined kappa for the 91 paired observations was calculated and presented in tables 1 and 2. For the 3 other tests the kappa for the better couple (MdK-PV) was calculated and presented in italics in table 2.

Asking for increase of pain on coughing, sneezing or straining and urinary incontinence resulted in consistent answers. Patient reports of numbness and decreased strength are inconsistent items. Overall, the agreement between the observers on the diagnosis after history taking shows substantial disagreement.

This disagreement is considerably reduced by the physical examination. The neurological investigation (decreased muscle strength, sensory loss, lowered reflexes, SLR, crossed SLR, and Bragard) resulted in consistent findings. Reflex differences were as consistently reported as absent reflexes. For the SLR, the most consistent testing technique defined a positive test result as an increase of typically dermatomal radicular pain. The definitions of a positive SLR as an increase of any pain in the leg or any restriction of SLR below 45° provided less consistent results. The consistency of the crossed straight leg raising benefitted from labelling an angle of crossed straight leg raising below 75 degrees as positive. The investigation of the lumbar spine was unreliable.

DISCUSSION

Was the consistency reported here optimal?

While accurate tests need to be consistent, consistent tests need not have sufficient discriminative value. In this study, the addition of interpretations to the diagnostic strategies to increase diagnostic accuracy may have been at the expense of test consistency. Also, a difference in clinical experience and the potential difference in clinical competence between the neurologist and the neurological residents may have been an additional source of disagreement in this study. Nishikawa⁸⁶ reported, however, that basic clinical skills may not improve after graduation from medical school. It is interesting that the two couples of observers reached similar agreement levels regardless of whether they had discussed and trained the testing technique together or not. This suggests that the agreement reported is not greatly influenced by the fact that the observers took a prior effort to increase their consistency. The exceptions were the Valleix pressure point and Naffziger's sign, more complex and perhaps less obvious parts of the physical examination; training such tests seems to have a relevant effect on the consistency attained.

An advantage of incorporating interpretations into the diagnostic tests and allowing for differences in clinical experience and skill between observers is that the test results

reflect everyday practice (increased external validity). Again for this reason, the investigators had defined the testing technique as a synthesis of the original descriptions in the literature and the daily routine. Variation in the disease status and presentation by the patient is another source of inconsistency. Often the patient complained of increasing pain during the investigation. Likewise, the second investigation might have been more painful than the first. This would partly explain why the second observer attained a higher proportion of positive test results in the pain-related physical examination items. Similarly, the first assessment of the history may have influenced the patient's answers in the second. The history items investigated are but an approximation of the actual history taking routine. The clinician will not ask the whole set of questions in a fixed order but will likely elaborate upon those questions the patient answered positively. Though it is questionable whether this increases the consistency of the signs and symptoms, it may add to the diagnostic accuracy.

What should a consistent history and physical examination include in the patient with pain in the leg?

The following remarks are under the assumption that the tests are diagnostically accurate. Overall, the history is inconsistent; in 1 of 4 patients a second doctor does not seem to agree with the absence or presence of a definitely positive history. However, an increase of pain in the leg on coughing, sneezing or straining, a feeling of coldness in the leg and urinary incontinence are substantially consistent contributions to the history and the diagnosis as a whole. Asking for a feeling of decreased strength without further elaboration should be abandoned as the answers are hardly reproducible.

In our study, the most reliable part of the physical examination is the examination of the muscle strength, either through testing of strength as proposed by the Medical Research Council⁸² or observation of walking on toes or heels. Of these items, the perhaps more complicated assessment of walking on heels and anterior tibial and peronei muscle strength seem to be least reliable. The frequent association of these tests with pain might (partly) explain this. Sensory loss can be demonstrated more reliably for the fine touch than for the pain modality. Though the latter seems more objective, the distinction between normal and abnormal is more difficult (should an occasional mistake in determining sharp and blunt be called abnormal?). Reflex differences are as consistently reported as absent reflexes, but occur more frequently and are therefore more useful. The SLR test has a substantial inter-observer agreement when typically dermatomal pain on SLR was rated. Labelling any pain provoked in the leg on SLR as positive was found to be less consistent. This is the reverse of what McCombe *et al.*⁷⁵ reported. It may be explained by the fact that the assessment of any pain in the leg in patients presenting with pain in the leg is more difficult than the assessment of any pain in the leg during SLR in patients presenting with low back pain, only. Van de Hoogen⁸⁷ found lower consistency values but the time interval between the two assessments was much greater allowing for more biologic variation or actual resolving of the clinical picture. The crossed SLR test was more consistent if a

maximum angle of restricted SLR was specified. Often the disagreement resulted from the fact that the first observer had stopped the straight leg raising at e.g. 80 degrees because of low back pain, while the second observer proceeded lifting the leg to e.g. 90 degrees regardless of the low back pain and then stopped because of pain in the contra lateral leg. The Valleix pressure points were inconsistent signs. A pain irradiating into the leg during Kemp's manoeuvre was inconsistent. As for the SLR, the disagreement might have diminished if only typically dermatomal pain would have been noted. Unfortunately, the type of pain in the leg was not registered for the Kemp sign. Results of investigating the low back are poorly comparable for different clinicians, but -if the intra-observer consistency is good- may be valuable to follow the patient over time. Overall, clinicians disagree on the presence of lumbosacral nerve root involvement in 1 of 5 patients.

How can the consistency findings in table 1 and 2 be interpreted?

The kappa (κ) expresses the proportion agreement observed, accounting for the chance agreement. Landis and Koch⁸⁸ provide qualitative terms for ranges of the κ ; 0.0-0.2 signifies slight consistency, 0.2-0.4 fair, 0.4-0.6 moderate, 0.6-0.8 substantial, and 0.8-1.0 almost perfect consistency.

However, the interpretation of the κ is not as straightforward as mentioned above because of the following four characteristics: (1) κ values are influenced by the number of categories of test results⁸⁹. The latter is illustrated at the bottom of both tables by the fact that a simple reduction of the diagnostic trichotomy (is the radicular syndrome present, doubtful or absent?) to a dichotomy (definitely present or not definitely present) improves the κ . (2) The κ depends on the relative number of positive test results (thus, indirectly the disease prevalence)^{89 90}. When a test is positive in 95% of the patients, the clinicians may have an overall agreement of 90%, yet may totally disagree on the patients with negative test results. In other words, high proportions of positive (or negative test) results result in a high chance agreement and a misleading overall agreement. In another example, the κ for a particular observed agreement, when 50% of the test results are positive, is higher than the κ for the same observed agreement, when almost all or almost no test results are positive⁹¹. Furthermore, the interpretation of the kappa itself also depends on the proportion of positive tests. According to Altman⁸⁹ it would then be misleading to compare kappa values from different studies when the proportions of positive test results substantially differ. (3) In addition, the precision of the κ diminishes (demonstrated by the higher standard deviation in tables 1 and 2) towards extremely low and high proportions of positive results. (4) When the spectrum of disease varies (due to different

---

⁹¹Suppose two clinicians would agree on the positivity of the SLR in 77 of 91 cases as did the two observers in table 2, but because of a different prevalence of disease they would agree on ten positive cases and on 67 negative cases. The chance agreement between the two would rise from 52% to 70% and the κ would decline from 0.68 to 0.50.

presentations, severity of symptoms and/or comorbidity in the various settings^{22 66}, the technical and interpretative difficulties of tests may vary. In other words, the mix of clear-cut and difficult cases varies accross populations⁹¹. With a greater proportion of diagnostically difficult cases the consistency of findings may be expected to decrease^{****}. For example, the assessment of an increase of typical pain on SLR in patients already presenting with pain in the leg and may be more difficult than the assessment of pain in the leg on SLR in patients without pain in the leg. Therefore, the higher likelihood of "nerve root involvement" and, hence, the differing proportions of positive history and physical examination items in our population compared to previous study populations of low back pain patients may lead to different consistency values. In summary, a valid comparison of kappas between studies is limited by differences in test definition, disease prevalence and/or patient spectrum.

****Suppose that all the disagreement occurred in those cases the clinicians labeled as difficult and that they agreed on all easy cases. In our study there would then be 30 easy normal cases, 47 easy and 14 difficult abnormal cases. If a general population of 1000 subjects were to have been investigated this case mix might have been 939 easy normal cases, 47 easy and 14 difficult abnormal cases. For this general population the agreement would be 95% with a κ of 0.73 compared with a 85% agreement and a $\kappa = 0.68$ in ours study. This spectrum effect seems to oppose the mathematical effect in the previous footnote.

Thou cold sciatica, cripple out senators, that their limbs may halt as lamely as their manners

Timon of Athens, William Shakespeare

ABSTRACT

The value of history and physical examination in the diagnosis of sciatica due to disc herniation is uncertain. Previous diagnostic studies have been flawed by referral and verification bias and by an inappropriate gold standard test. Therefore, a diagnostic study was designed to investigate the association between clinical findings and a Magnetic Resonance Imaging (MRI) diagnosis of nerve root compression.

Patients with pain radiating into the leg, and of sufficient severity for bed rest therapy to be considered were investigated. Clinical findings and MRI evidence of nerve root compression was investigated.

274 patients were selected. In 2 by 2 tables, nerve root compression was associated with six history and ten physical examination findings (depressed reflexes, four nerve root tension signs, three signs of decreased muscle strength, and two signs of decreased spinal mobility). When the contributions of other findings were accounted for, six symptoms and two signs (increased finger-floor distance and paresis) significantly predicted nerve root compression. The straight leg raising test was not significantly predictive. The Receiver Operating Characteristic (ROC) showed that the area-under-the-curve increased from 0.80 to 0.83 when the physical examination was added to the history.

The bivariate results agree best with our ideas of valuable signs and symptoms of nerve root compression. However, most tests had lower sensitivity and specificity values than previously reported. Moreover, the multivariate analysis and the ROC showed that most of the physical examination findings, including the straight leg raising test, add little diagnostic information to the history.

INTRODUCTION

About one in every five male and one in three female adults will experience pain radiating into the leg at some time during their life⁹². The pain is called sciatica if it is presumed to originate in lumbosacral nerve roots or the sciatic nerve (although there is no generally accepted definition¹). A major cause of sciatica is nerve root compression. The initial diagnosis of sciatica due to disc herniation is based on history and physical examination^{1 93}.

A systematic review of the value of history and physical examination in radicular syndromes concluded that none of the tests considered had a high sensitivity or specificity⁹⁴. Another systematic review⁹⁵ showed that the literature on these tests should be considered as inconclusive for several reasons. Firstly, no reports on diagnostic value were found for most of the questions in the history and for several tests in the physical examination (as was also pointed out by Andersson and Deyo¹). Secondly, diagnostic test results should be compared with an appropriate gold standard, but previous diagnostic studies referred against a gold standard of disc herniation instead of nerve root compression. Disc herniations need not result in nerve root compression^{96 97}, while nerve root involvement can occur without disc herniation. Thirdly, most studies selected patients on the basis of the same clinical findings that were the goal of the subsequent diagnostic study (work-up or sequential ordering bias^{66 68 67 22}). Finally, patients are likely to be subjected to a series of tests, especially during history-taking and physical examination. Multivariate analysis allows an expression of the independent contribution of the clinical findings to the diagnosis. None of the previous diagnostic studies in this area investigated the independent diagnostic value or applied multivariate analysis.

Because of these limitations in the literature, a cross-sectional diagnostic study was designed to investigate the diagnostic value of history and physical examination in patients presenting to a primary care physician with a new episode of pain radiating into the leg. The reference diagnosis of nerve root compression⁶⁴ was established by Magnetic Resonance (MR) imaging.

PATIENTS AND METHODS

Study population

Between February 1995 and December 1996, 50 general practitioners (GPs) in Maastricht (the Netherlands) and surrounding villages referred patients to the Neurology department of the Maastricht University Hospital if they presented for the first time with the current episode of (1) pain radiating into the leg below the gluteal fold, and (2) pain of sufficient intensity to justify bed rest therapy for 14 days. Patients were excluded for (1) previous spinal surgery, (2) pregnancy, (3) co-morbidity of such a nature that the patients' well-being was primarily determined by it, (4) MR imaging contra-indications (metal-containing prosthesis, pacemaker, etc.).

Figure 1a. T1-weighted sagittal MR image showing a L4/L5 disc herniation with annular rupture.

Figure 1b. MR radiculogram

Figure 1c. T1-weighted MR image showing compression of the left L5 nerve root by a mediolateral disc herniation

Investigations

Within two days of referral by the GPs, all patients selected underwent a standardized history-taking and physical examination. Test definition was based on good clinical practice, standard textbook sections and/or scientific literature and was summarized in a protocol. The straight leg raising test (SLR) was considered positive when typically dermatomal pain was provoked by raising the straight leg. The Bragard sign was considered positive when - at five degrees below the angle of the positive SLR - the pain was provoked by dorsiflexion of the foot. The Kemp sign was considered positive when lateroflexion and rotation of the lumbar spine (to decrease the size of the intervertebral foramen), together with axial pressure, provoked leg pain. Definitions and inter-observer variabilities for these and all other tests have been reported elsewhere in chapter 3 of this thesis.

The gold standard in all patients was MR imaging of the lumbar spine within 24 hours after the clinical examination. It was performed using a 0.5 T system according to the following scanning protocol: (1) sagittal and transverse T1-weighted sequences (TR/TE 400-600/20 msec) with 4 and 5 mm slice thickness, respectively, and (2) a sagittal dual-echo protondensity T2-weighted sequence (fast spin-echo TR/TE 4500/25-150 msec) with 5 mm slice thickness. Additionally, MR radiculography was performed consisting of two heavily T2-weighted fast spin-echo sequences (TR/TE 6000/450 msec) with 4mm slice thickness, oriented parallel to the emerging L5 root, about 20° left- and right-oblique to the coronal plane, reformatted with a maximum intensity projection protocol⁹⁸. A standardized assessment of all MRI studies was done by a neuroradiologist (JW), experienced in spinal MRI.

The assessment focused on the presence or absence of nerve root or nerve root sleeve compression, graded on a four-point scale (definitely absent/probably absent/probably present/definitely present). Criteria used for grading included the presence of protruding annulus or extruded nucleus material or lateral recess narrowing on standard MR images, and flattening and compression of the ventrolateral border of the dural sac and/or emerging nerve root sleeve, with obliteration of the surrounding epidural fat. As regards the MR radiculography images, the criteria included dural impressions, kinking and swelling of the intradural nerve root and cut-off of the root sleeve filling. Figure 1 gives an example of MRI evidence of nerve root compression due to a disc herniation in one of the patients included.

The clinical investigator was unaware of the MRI findings, while the neuroradiologist was unaware of the clinical findings.

Statistical analysis

A bivariate analysis was based on two by two tables relating history and physical examination items to nerve root compression on MRI (with “definitely” and “probably” combined). Sensitivity, specificity, and likelihood ratios for positive and negative test results (LR+ and LR-) were calculated, as well as the diagnostic odds ratio (DOR with $DOR = LR+/LR-$)^{19 99}. Statistical significance of the association was assessed using Pearson’s chi square statistic.

Multiple logistic regression analysis allows the history and physical examination to

be analyzed as a series of tests with nerve root compression on MRI as the dependent variable. The diagnostic odds ratios of the tests are then adjusted for the effect of other test results acquired simultaneously, and express the independent diagnostic value of each individual test. In a backward stepwise logistic regression analysis^{89 100}, the likelihood ratio test was used for each step at a significance level of 0.05. Firstly, a set of general patient characteristics (all variables in table 1) was modeled, leading to model A. Secondly, a set of history items (all variables in table 2) was modeled (model B) including also the significantly predictive variables of model A. Because many of the signs tested similar modalities, these were divided into two groups. The "essential" group included the more generally accepted signs of nerve root compression, while the "alternative" groups contained mostly alternatives to the essential tests, which were either less (inter-observer) consistent, or less well-known. Determining hypalgesia, hypesthesia, and depressed ankle or knee tendon reflexes were "essential". Assessment according to the guidelines of the Medical Research Council was the "essential" method of establishing decreased muscle strength, while abnormal walking on heels or toes, and abnormal knee bending were the "alternative" method. The SLR was the "essential" nerve root tension sign, while the Bragard sign¹⁷, reversed SLR, Valleix pressure points, Kemp sign⁷⁷, and Naffziger sign⁷⁶ were the alternatives. The finger-floor distance was an "essential" spinal sign, while decreased lordosis, antalgic posture, paravertebral hypertonia, vertebral percussion tenderness were "alternative" signs of spinal mobility/tenderness. A final reduced model C was constructed, based on all "essential" signs and the significant predictors of model B. Finally, "alternative" final models were constructed for each of the "alternative" tests by substituting the "alternative" for its corresponding "essential" sign.

The probability of nerve root compression on MR Imaging predicted by the various reduced models was calculated for each patient. Subsequently, to evaluate the predictive power of each model, Receiver Operating Characteristics (ROC) curves were plotted, displaying the sensitivity versus (1-specificity) for each cut-off point of predicted posterior probability^{101 102}. The areas under the ROC curves were calculated. The probability cut-off with the maximum expected gain was determined as the point with the highest sum of sensitivity and specificity¹⁰³. As a reference, the sensitivity and specificity of the observer's summary statement on the presence of nerve root compression (at the bottom of table 2 and 3) was depicted.

Ethics

The procedures of this study were in accordance with the ethical standards of the revised Helsinki declaration of 1983 and were approved by the Maastricht University Hospital ethical committee.

RESULTS

338 patients were referred to the diagnostic study. 35 patients did not meet the inclusion criteria of sufficient pain intensity to justify 14 days of bed rest therapy. 29 patients were excluded because of previous spinal surgery (2 patients), pregnancy (1), metal-containing corpus alienum (1), claustrophobia (6), and unavailability of MRI within 24 hours (19). Baseline characteristics were no different for the groups with and without MR scans. A standardized history, physical examination and assessment of MRI were completed for 274 patients. The kappa value for the intra-observer variability of diagnosing the absence or presence of nerve root compression was 0.83. Table 1 depicts the most important patient characteristics. Root compression was found on 56% of the MRI studies. The pain radiating into the leg was clinically

Table 1. Baseline patient characteristics in their relation to nerve root compression (n=274)

Characteristic	Percentage of patients with characteristic	odds ratio (p-value)
Age (mean 46, SD 12)		
16-40 years	33	
41-50 years	34	1.3 (0.42)
51-81 years	33	2.2 (0.01)
Male gender	49	1.8 (0.01)
Duration of disease (median 19 days)		
14 days or less	46	
15-30 days	22	1.7 (0.11)
more than 30 days	33	0.7 (0.15)
Living alone	18	1.3 (0.45)
High education level	26	1.6 (0.09)
Job type		
no job	40	
cognitive	25	1.1 (0.87)
standing/walking/lifting	23	0.5 (0.03)
physically demanding	12	0.6 (0.22)
Any co-morbidity	39	1.3 (0.31)
Smoking	47	0.7 (0.14)
Sports activities	43	1.1 (0.61)
Exercised back/abdominal muscles	42	0.8 (0.45)
Family history positive for sciatica	33	1.1 (0.67)
Quetelet index (mean 26, SD 5)		

Note- These general patient characteristics were all part of the first "general" set of variables of the multiple logistic regression. SD= standard deviation. The p-value is based on the Chi-Square (χ^2) statistic.

diagnosed as a lumbosacral radicular syndrome in 67% of patients, while other diagnoses included coxarthrosis, spondylolisthesis, regional neoplastic growth, peripheral neuropathy, and pain of tendomyogenous or other origin. The sample included sciatica patients both with and without (various degrees of) neurological deficit. Two patients suffered from urinary incontinence and saddle anaesthesia.

In the bivariate analysis, age, gender and duration of disease were associated with nerve root compression (table 1). Three history items discriminated significantly (i.e. had an odds ratio significantly different from 1.0) with respect to nerve root compression (table 2): typically dermatomal pain - mostly because of a high

Table 2. The diagnostic accuracy of the history (n=274)

Diagnostic test item	Sensitivity (%)	Specificity (%)	LR+	LR-	diagnostic odds ratio	p
typically dermatomal distribution	90	31	1.3	0.3	3.8	<0.001
more pain on coughing, sneezing or straining	50	67	1.5	0.7	2.0	0.004
dermatomal coldness in the leg	31	80	1.6	0.9	1.8	0.03
more pain on sitting	54	43	1.0	1.1	0.9	0.67
less pain on walking or standing	20	79	1.1	0.9	0.9	0.85
less pain on lying down	60	43	1.1	0.9	1.1	0.58
sudden onset	40	58	0.9	1.0	0.9	0.70
paroxysmal pattern	44	62	1.1	0.9	1.3	0.35
worse in the evening/night	31	66	1.1	0.9	0.9	0.54
with known cause	25	70	0.8	1.1	0.8	0.33
subjective muscle weakness	27	62	0.7	1.2	0.6	0.06
subjective sensory loss/ numbness	40	58	0.9	1.0	0.9	0.70
dermatomal paraesthesias	51	41	0.9	1.2	1.2	0.20
disturbed urinary passage	4	97	1.6	1.0	1.0	0.50
urinary incontinence	4	94	0.7	1.0	0.7	0.49
previous back pain episodes	76	29	1.1	0.8	1.3	0.61
previous leg pain episodes	36	57	0.8	1.1	0.7	0.54
positive history	78	48	1.5	0.5	3.1	<0.001

Note-155 patient had positive gold standard and 122 negative gold standard.

LR+= likelihood ratio of positive test result, LR-= likelihood ratio of negative test result

p= p-value of Pearson's chi square, Bold: $p \leq 0.05$

Table 3. The diagnostic accuracy of the physical examination (n=274)

Diagnostic test item	Sensitivity	Specificity	LR+	LR-	diagnostic odds ratio	p
abnormal walking on toes	9	98	5.4	0.9	5.8	0.01
paresis	27	93	4.1	0.8	5.2	<0.001
decreased lordosis	26	90	2.7	0.8	3.3	0.001
abnormal walking on heels	19	93	2.9	0.9	3.3	0.003
finger-floor distance ≥ 25 cm	47	73	1.7	0.7	2.4	0.001
absence of ankle/knee tendon reflex	15	93	2.2	0.9	2.4	0.02
positive straight leg raising	64	57	1.5	0.6	2.3	0.001
positive Bragard	52	62	1.4	0.8	1.8	0.02
positive Kemp's sign	35	76	1.5	0.9	1.7	0.05
positive Naffziger sign	16	93	2.1	0.9	2.4	0.03
painful/reduced hip mobility	4	93	0.5	1.0	0.5	0.21
antalgia	13	90	1.3	1.0	1.3	0.49
vertebral tenderness	4	98	1.6	1.0	1.6	0.49
paravertebral hypertonia	6	89	0.6	1.1	0.5	0.15
sensory loss: hypesthesia	28	66	0.8	1.1	0.8	0.28
hypalgesia	17	84	1.1	1.0	1.1	0.73
positive crossed SLR	13	91	1.4	1.0	1.4	0.36
positive reversed SLR	25	80	1.2	0.9	1.3	0.38
positive Valleix pressure points	11	95	2.1	0.9	2.3	0.09
positive history & examination	81	53	1.7	0.4	4.7	<0.001

Note- LR= likelihood ratio, LR+= of positive test result, LR-= of negative test result, p= p-value of Pearson's chi square, Bold: $p \leq 0.05$, SLR= straight leg raising

sensitivity (90%)-, coldness in the leg -mostly a high specificity (80%)-, and increased pain on coughing, sneezing or straining. Ten physical examination items, shown in bold in table 3, discriminated significantly with respect to root compression on MRI.

The multiple logistic regression analysis produced a model predicting the probability of nerve root compression. Table 4 shows the significant predictors, including two general patient characteristics (age and duration of disease), four symptoms from the history, and two signs from the physical examination. The SLR was not a significant predictor, while an increased finger-floor distance was. An alternative model with "decreased lordosis" substituted for "increased finger-floor distance", showed lordosis had an odds ratio (OR) of 3.9 (1.7-8.9). Paresis was another significant predictor. The alternative analysis showed that strength could also be assessed by having the patient walk on heels (OR=3.2 (1.2-8.7)) or toes (OR=6.2 (1.1-35)) instead of checking individual muscle strength.

Table 4. Predictors of mechanical nerve root involvement (n=274)

Test		Adjusted diagnostic odds ratio	95% confidence limits
History			
Age	41-50 vs 16-40 years	1.8	1.3-2.6
	51-81 vs 16-40 years	2.8	1.9-4.2
Duration of disease	15-30 vs < 15 days	2.2	1.5-3.3
	>30 vs <15 days	0.8	0.6-1.1
Paroxysmal pain		1.8	1.3-2.5
Pain worse in the leg than in the back		4.5	3.3-6.2
Typically dermatomal distribution		3.2	2.2-4.7
Worse on coughing, sneezing or straining		2.0	1.4-2.7
Physical examination			
Finger-floor distance	5-24 cm vs 0-4 cm	1.1	0.7-1.6
	>25 cm vs 0-4 cm	2.8	1.9-4.3
	missing vs 0-4 cm	1.0	0.4-2.1
Paresis		5.2	3.3-11.6
Additional tests			
Decreased lumbosacral lordosis*		3.7	2.4-5.7
intercept		-3.511	

Note- The final reduced model included the variables shown. The adjusted odds ratios are the exponential values of the coefficients in the reduced model. The intercept for this model was -3.511 (standard deviation= 0.62).

The sensitivities and specificities for each possible predicted probability cut-off are depicted in the ROC in figure 2. The area under the curve (AUC) for the model using history findings only was 0.80. It increased to 0.83 when the examination findings were added. Maximum predictive gain of the model occurred at a predicted

probability value of 0.625; the sensitivity and specificity were 72% and 80%, respectively.

Figure 2. Receiver Operating Characteristic

- upper curve: diagnostic value of the model with the variables of table 4 (area-under-the-curve = 0.83)
- ..... lower curve: diagnostic value of the best model based on general patient characteristics and history (area-under-the-curve = 0.80)
- maximum diagnostic gain of model according to Connell & Koepsell
- Δ sensitivity and 1-specificity for the observer's diagnosis after history
- ▲ sensitivity and 1-specificity for the observer's diagnosis after history & physical examination

DISCUSSION

The bivariate results show that 16 patient characteristics and clinical findings are significantly associated with nerve root compression. For most of these 16 tests, these associations have been reported previously. However, the overall impression is that sensitivity and specificity values in this study were lower than those previously reported⁹⁴ or those shown in chapter 2 of this thesis. Different study populations and designs⁶⁷ may explain this.

The participating patients represent a primary care population, primarily selected on the basis of radiating pain into the leg, with a broad disease and co-morbidity spectrum. Both sexes were equally represented in our study, while previous secondary care patient series had average male to female ratios of 2:1. Also, neurological deficit was far less frequent in our patient sample than in previously described series (see chapter 2). This strongly favours our opinion that our primary care based study of patients not selected on any other clinical finding than radiating pain does not suffer from the verification bias¹⁰⁴ and spectrum bias²¹ of previous studies^{67 22}.

MRI was used as the present state-of-the-art procedure to demonstrate nerve root compression^{64 105 106 72}. Sensitivity and specificity values reported for MRI have ranged from 71 to 100% and from 50 to 86%, respectively^{107 108 64}. The limitation of these values is that they reflect the agreement with another gold standard test, which itself may be inferior to MRI. Alternatively, the imperfect sensitivity of MR imaging may be explained by changes in the pathoanatomy of the nerve root with posture and over time¹⁰⁹. The actual specificity is uncertain. Previous studies have reported on disc herniation in asymptomatic subjects on MR imaging^{110 96}, CT¹¹¹ and caudography¹¹². However, none referred to the presence of nerve root compression. The only indication of the imperfect specificity, therefore, seems to be that 12% of our patients had MRI findings of root compression on the asymptomatic side. The additional value of MR radiculography⁹⁸ has as yet not been unequivocally demonstrated.

The only history item that could be compared to the literature was the typically dermatomal pain distribution, which had a 90% sensitivity, compared to a 99% sensitivity in the study by Kerr and colleagues²⁹. Strikingly, subjective muscle weakness and sensory loss actually diminished the likelihood of nerve root compression. These findings are in disagreement with current opinion and the authors' prior beliefs. The selection bias operative in previous studies may have affected our understanding of the true relations of these complaints with nerve root compression. Apparently, subjective weakness or sensory loss has no neurological basis in most patients.

Table 3 shows that ten physical examination items are related to the reference diagnosis. What is also shown is that even before the physical examination, the a priori likelihood of disease has been altered by the history. Rather than considering only one or two tests, one should use history and physical examination as two batteries of tests, and interpret the items in each battery simultaneously. The independent and simultaneous contributions of tests to the diagnostic work-up of the sciatica patient have not been studied previously. Table 4 and the ROC curve in figure

2 reflect that, after the history has been taken, few signs in the physical examination contribute to a more accurate diagnosis. Typically dermatomal distribution of pain, an increase of pain on coughing/sneezing/straining, paroxysmal pain and predominant leg pain are indicators of nerve root compression, which is an agreement with clinical experience. The predictive value of paresis also accords with general beliefs¹. The alternative analysis shows that using the MRC guidelines to check strength may be substituted by having the patient walk on heels or toes. A surprising predictor is an increased finger-floor distance. Apparently, the patient who bends over to touch the floor with his fingers does not only stretch the nerve roots and the sciatic nerve but also delordates the lumbosacral spine. The latter action may accentuate the mechanical effect of the disc herniation on the nerve root. A decreased lumbosacral lordosis is predictive of nerve root compression and seems exchangeable with the finding of an increased finger-floor distance. Part of the restriction in bending may be caused by the decreased lordosis. The increased finger-floor distance should nevertheless be preferred because of its higher inter-observer consistency¹¹³. Another striking aspect is that the SLR and other nerve root tension signs are not significantly predictive of root compression. They may indicate nerve root tension or irritation, but not necessarily nerve root compression on MRI.

The models were based on the investigator's observations, yet predicted nerve root compression better than the investigator did (figure 2). The ROC curves were based on an evaluation of the predictive rule in the same research population from which the rule was derived. Such evaluations lead to an overestimation of the performance of the model. In future research, the model should therefore be tested in a diagnostic study of a different, yet clinically similar, research population.

We conclude that the main component in the diagnosis of sciatica due to disc herniation is the history and that few signs add any further diagnostic information. The patient most likely to have nerve root compression is a 51-81 year-old patient, who 2-4 weeks ago developed a paroxysmal, typically dermatomal pain, predominantly in the leg. This pain worsens on coughing, sneezing or straining. It is accompanied by decreased muscle strength and increased finger-floor distance (or, alternatively, decreased lumbosacral lordosis). Nerve root tension signs such as the SLR may be indicators of root irritation but are no clear indicators of nerve root compression.

5

The patho-anatomy of the clinical findings in the sciatica patient

A Magnetic Resonance Imaging study

ABSTRACT

Anatomical details of nerve root compression may explain the production of the signs and symptoms of sciatica? Anatomical studies have offered many theories without clearly demonstrating the clinical relevance of the observations. Clinico-anatomical series are scarce and are affected to a great extent by selection bias.

Firstly, a schematic drawing of the lumbar anatomy was made based on the literature and *in vitro* anatomical observations. Then, a detailed and standardized clinical and Magnetic Resonance Imaging (MRI) diagnosis was made in primary care patients presenting with pain irradiating into the leg. Clinical and MRI findings were correlated. Finally, the anatomical drawing was compared with the clinical data.

Clinical data showed that the higher the level of symptomatic disc herniations, the more likely the compression will be more laterally situated. Classical findings of sciatica (typically dermatomal pain, pain increase on coughing, sneezing or straining, and positive straight leg raising) were most likely with compression of the nerve root in the axilla and with mediolateral disc herniations.

L3, L4, L5 and S1 nerve roots each tend to be compressed at different sites along the rostral-caudal course of the nerve root. Disc herniations become symptomatic at different sites for each disc level. The drawing produced a priori could well explain these findings. The occurrence of particular clinical findings can be predicted by specific patho-anatomical findings.

INTRODUCTION

Unfortunately, “classical” presentations of disease are rare. Sciatica, one of the most frequent syndromes^{92 114} in neurological practice, may present in many different ways. This complicates the diagnostic process. Why are one patient’s signs and symptoms different from the other’s? Of course, pre-morbid differences in e.g. personality and culture may lead to a different presentation of clinical findings. Differences in the underlying pathophysiology of the disease process may be another major cause of the variance¹¹⁵. Since long, a lumbosacral disc herniation has been regarded as the most prominent patho-anatomical finding implicated in the genesis of nerve root irritation and dysfunction¹¹⁶⁻¹²¹. Can the patho-anatomy be further elucidated to explain the variance in the clinical syndrome of sciatica due to disc herniation?

That the level of disc herniation correlates with involvement of specific nerve roots and thus with specific root syndromes is generally acknowledged^{40 54 43 44}. The mechanisms of clinical sign and symptom production are largely speculative, however. Scarce are the studies that correlated patho-anatomical details with clinical features¹²². These clinico-anatomical studies suffer from selection bias much the same as the majority of the studies on the clinical diagnosis of sciatica reviewed in chapter two.

Most theories about symptom production have been based on anatomical observations¹²³ and biomechanical experiments^{124 125 126}. Especially noteworthy is the study by Falconer and colleagues reporting *in vivo* and *in vitro* observations¹²⁰. They proposed the following anatomical arrangement of extrathecal nerve roots and intervertebral discs: the nerve root sleeve branches off the dural sac at a level clearly caudal to the disc at L3, less caudal to the disc at L4, at the lower disc level at L5 and at the upper disc level at S1. Cailliet’s schematic drawings¹²⁷ also show this anatomical arrangement. Our own anatomical dissections reconfirmed their observations and led to the schematic drawing in figure 6. These dissections also showed that the angle at which the extrathecal nerve root branches off the dural sac is different for each level (as Cailliet already reported). Moreover, the length of extrathecal nerve root between dural sac and dorsal root ganglion was found to increase towards lower disc levels. This seems in keeping with Reulen, Muller and Ebeling’s study of the bony dimensions of the spinal canal¹²⁸. They reported that the pedicles are closer to the disc space towards lower vertebral levels and that the length of the foraminal bony canal increases towards lower root levels. Falconer and colleagues also showed both *in vitro* and *in vivo* that the nerve root moves during a straight leg raising manoeuvre. These biomechanical observations on the SLR have been further elaborated on in experimental work by others^{129 124}. It remains to be shown whether such hypothesized mechanisms are indeed operative *in vivo*.

Previous investigations have addressed the anatomical findings in highly selected patients. The majority of tests commonly performed in the sciatica patient have not been correlated with anatomical details at all. If the focus of the investigation is on why patients exhibit a particular set of clinical findings, the study population should ideally be a sample of the entire population likely to undergo such tests. Therefore, a

cross-sectional sample of primary care patients with pain irradiating into the leg was subjected to clinical and MRI investigations. It was investigated (1) whether the site of compression along the course of the nerve root was comparable for the L3, L4, L5 and S1 nerve roots, (2) whether the site of a symptomatic disc herniation was different for L3L4, L4L5 and L5S1 disc herniations, and (3) whether clinical findings, not regarded as indicative of level of nerve root involvement, (such as the straight leg raising (SLR) test, Kemp sign, Naffziger sign, increased finger-floor distance), were related to specific sites of nerve root compression (along the rostral-caudal course of the nerve root). Differences between patients in the pattern of dermatomal irradiation of the pain, decreased muscle strength, reflex loss and sensory loss can be explained by the involvement of either L3, L4, L5 or S1 nerve roots.

METHODS

The procedures of this study were in accordance with the ethical standards of the institutional ethical committee and with the revised Helsinki declaration of 1983. Between February 1995 and December 1996, 50 general practitioners (GPs) in Maastricht and surrounding villages referred patients to the Neurology department of the Maastricht University Hospital if they presented for the first time with the current episode of (1) pain irradiating into the leg below the gluteal fold, and (2) pain intensity sufficient to justify bed rest for 14 days. Exclusion was based on (1) previous spinal surgery, (2) pregnancy, (3) unavailability for follow-up visits (moving plans, severe co-existing disease), (4) co-morbidity such that their well-being was being determined primarily by this co-morbidity, (6) MRI contra-indications (metal-containing prosthesis, pacemaker, etc.).

Baseline investigations

A standardized history and physical examination was performed in all selected patients within two days of referral by the GPs. The inter-observer variability and diagnostic value of these are reported in chapter three and four. The history reported in this chapter highlights three questions: (H1) Point out the distribution of the pain with one finger? The observer assessed whether the distribution was "typically radicular", (H2) How does the pain in the leg change on coughing, sneezing or straining? (H3) Does the pain increase on sitting? The four physical examination items were: (P1) the straight leg raising test (=SLR) (a positive result was defined as a provocation of irradiating pain into the leg in a dermatomal distribution upon flexion of the extended leg in the hip)^{13 10}, (P2) the Kemp sign: provocation of the pain by hyperextension and lateroflexion of the back by the observer with subsequent axial pressure^{77 78}, (P3) finger-floor distance more than 25 cm= distance between middle fingers of both hands and the floor when the erect patient anteflexes with legs fully extended in the knees, (P4) the assessment of paresis: the extensor hallucis longus, anterior tibial, long and short peronei, gastrocnemius, hamstring, quadriceps, iliopsoas,

and gluteus maximus et minimus muscles were tested adopting the technique and the rating scale described by the Medical Research Council⁸².

The Magnetic Resonance Imaging (MRI) assessment

All patients had MRI studies of the lumbar spine within 24 hours after the clinical examination. It was performed using a 0.5 T system according to the following scanning protocol: (1) sagittal and transverse T1-weighted sequences (TR/TE 400-600/20 msec) with 4 and 5 mm slice thickness, respectively, and (2) a sagittal dual-echo protondensity T2-weighted sequence (fast spin-echo TR/TE 4500/25-150 msec) with 5 mm slice thickness. Additionally, MR radiculography was performed consisting of two heavily T2-weighted fast spin-echo sequences (TR/TE 6000/450 msec) with 4mm slice thickness, oriented parallel to the emerging L5 root, about 20° left- and right-oblique to the coronal plane, reformatted with a maximum intensity projection protocol⁹⁸. A standardized assessment of all MRI studies was done by a neuroradiologist, experienced in spinal MRI.

Figure 1. Site of disc herniation

The assessment focused on the presence or absence of nerve root or nerve root sleeve compression, graded on a four-point scale (definitely absent/probably absent/probably present/definitely present). Criteria used for grading were presence of protruding annulus or extruded nucleus material or lateral recess narrowing on standard MR images, and flattening and compression of the ventrolateral border of the dural sac and/or emerging nerve root sleeve, with obliteration of the surrounding epidural fat. On the MR radiculography images, the criteria were dural impressions, kinking and swelling of the intradural nerve root and cut-off of the root sleeve filling.

Figure 2 Site of root compression

The site of disc herniation was recorded as none, medial, paramedian, mediolateral or foraminal (see figure 1). The site of nerve root compression was recorded as no compression, compression at the dural sac, at the shoulder of nerve root sleeve, at the axilla of the nerve root sleeve, or at the dorsal root ganglion (see figure 2). Ungradable was recorded if there were technical difficulties or large hernias that could not be registered in one particular category. Examples are shown in figure 3.

There were additional assessments with regard to annular rupture (absent/uninterpretable/present), and degeneration of nucleus pulposus (one-level/none or all-levels). The intra-observer variability of the MRI assessment was investigated. The clinical investigator was unaware of the MRI findings and the neuroradiologist of the clinical findings.

Statistical analysis

Kappa was calculated for the intra-observer variability of the MRI assessment. Three correlates of the patho-anatomy were investigated: (1) the level of root compression, (2) the site of root compression and (3) the site of disc herniation. The associations between these patho-anatomical substrates was investigated (table 2a+b). Whether pain intensity in the leg and in the low back on a visual analogue scale could be predicted by site of disc herniation was investigated with multiple linear regression (table 3). The percentage of patients with the sign or symptom for each site of disc herniation, site of root compression or level of root compression was calculated (table 4a+b). To explore which of the patho-anatomical substrates correlated best with presence or absence of a particular clinical finding, logistic regression was used. Categorical variables representing the level of root compression, the site of root compression and the site of disc herniation were simultaneously entered into a model to investigate which classification was most predictive of the clinical finding. Any patho-anatomy associated with the clinical finding in the logistic model at a 0.05 significance level, was marked.

Figure 3a. Sagittal T1W MRI: foraminal disc herniation with root ganglion compression

Figure 3b. Transverse T1W MRI: dorsal root ganglion compression (same patient as in 3a)

Figure 3c. Transverse T1W MRI: Mediolateral disc herniation with root shoulder compression

Figure 3d. Transverse T1W MRI: paramedian herniation with shoulder compression

Figure 3e. MR radiculogram (with MRI of 3d)

Figure 3f. Transverse T1W MRI: Medial disc herniation with dural sac compression

RESULTS

274 patients with pain irradiating into the leg participated in the study. 27 patients had been excluded because MRI studies had not been possible (due to pregnancy, metal-containing corpus alienum, claustrophobia or unavailability of scan time). The baseline characteristics were similar for the groups with and without MR scans. Mean age was 46 years (± 12), 49% were male, median duration of disease was 19 days. Root compression was found on 56% of MR scans. The pain irradiating into the leg was diagnosed as sciatica in 67% of patients. Other diagnoses included spondylolisthesis, cox arthrosis, regional neoplastic growth, peripheral neuropathy, and pain of tendomyogenous or other origin. The intra-observer consistency of the MRI assessment ranged from moderate to very good (kappa's are presented in table 1).

Table 1. Intra-observer variability of MRI assessment

MRI findings	kappa
nerve root compression on Magnetic Resonance Imaging	1.0
nerve root compression on Magnetic Resonance Radiculography	0.83
annular rupture	0.92
disc herniation	0.83
site of root compression	
none - dural sac - shoulder - axilla - dorsal root ganglion	0.60
site of disc herniation	
none - medial - paramedial - mediolateral - foraminal	0.52
Nerve root level	
none - L3 - L4 - L5 - S1	0.81

Table 2a shows that symptomatic disc herniations at the L3L4 level occur most frequently in the intervertebral foramen. At L3L4, compression of L4 is two times more likely than compression of L3, at L4L5, compression of L5 three times more likely than compression of L4 and at L5S1, compression of S1 five times more likely than compression of L5.

Table 2a.   **Distribution of site of disc herniation and nerve root involved per level of disc herniation**

	level of disc herniation (%)		
	L3L4 (n=15)	L4L5 (n=62)	L5S1 (n=106)
Site of disc herniation			
uninterpretable	0	2	4
medial	0	18	6
paramedian	20	15	3
mediolateral	20	42	43
foraminal	60	24	11
Nerve root			
L3	33	4	1
L4	66	22	1
L5	0	72	16
S1	0	2	82

Note- The percentage of disc herniation levels with a particular site of disc herniation and nerve root involvement is shown.
ungradable- specification of site impossible due to technical difficulties or very large disc herniations

Table 2b shows that the site of compression of the nerve root is different for each of the nerve roots. The chance of compression of the dorsal root ganglion at a certain root level decreases from L3 down to S1 (20 times as likely in L3 than in S1). The chance of compression of the axilla of the nerve root is, however, increasingly more likely towards lower levels of root compression.

Table 2b.   **Distribution of site of nerve compression per nerve root involved**

site of compression	Level of nerve root involvement (%)			
	L3 (n=7)	L4 (n=20)	L5 (n=58)	S1 (n=66)
ungradable	0	0	9	9
dural sac	29	20	22	11
shoulder	29	45	38	47
axilla	0	5	26	32
root ganglion	43	30	5	2

Note- The percentage of levels of nerve root involvement with a particular site of root compression is shown.

typically dermatomal distribution and an increase of pain on coughing, sneezing or straining are especially likely with compression of S1. Also, there is a trend for both symptoms to occur more frequently with mediolateral herniations and axilla compression. The most typical site of nerve root compression for pain to increase on sitting is at the shoulder of the root sleeve.

Table 3a. The history versus site of disc herniation and site/level of nerve root compression
Percentage of patients with a positive test for each particular site of involvement

Typically dermatomal distribution of pain					
medial herniation	77	dural sac compression	78	L3	57
paramedian herniation	82	shoulder compression	88	L4	90
mediolateral herniation	92	axilla compression	92	L5	85
foraminal herniation	80	ganglion compression	78	S1	97■
Increase of pain on coughing, sneezing or straining					
medial herniation	40	dural sac compression	41	L3	14
paramedian herniation	53	shoulder compression	48	L4	40
mediolateral herniation	56	axilla compression	55	L5	48
foraminal herniation	40	ganglion compression	39	S1	58■
Increase of pain on sitting					
medial herniation	59	dural sac compression	33	L3	57
paramedian herniation	50 ■	shoulder compression	63 ■	L4	40
mediolateral herniation	57	axilla compression	51	L5	40
foraminal herniation	36	ganglion compression	39	S1	70

Note- The percentage of patients with the particular site of herniation, compression or root involvement that exhibited a positive test result is shown (77% of patients with medial herniations had a typically dermatomal distribution of pain). It should be noted that in patients with radiating pain into the leg without nerve root compression the three symptoms were positive in 73, 25 and 58%, respectively.

Table 3b shows the relation of MR patho-anatomy of nerve root and disc to examination findings. Straight leg raising is especially likely with compression in the axilla of the nerve root sleeve, irrespective of level (see figure 4). The same holds for another nerve root tension sign, the Kemp sign. A finger-floor distance of more than 25 cm in patients with irradiating pain into the leg is significantly more often present when the site of the compressing disc herniation is mediolateral (figure 5). Finally, compression of the L4 root most often leads to a paresis. Not shown in the table is that 29% of L3 and 35% of L4 compressions had unilateral absence of the knee tendon reflex, while 5% of L5 and 14% of S1 compressions had unilateral absence of the ankle tendon reflex.

Figure 4. Distribution of site of root compression per test

Table 3b. Physical examination versus site of herniation and site/level of root compression. Percentage of patients with a positive test for each particular site of involvement

Straight leg raising test					
medial herniation	53	dural sac compression	48	L3	57
paramedian herniation	58	shoulder compression	56	L4	90
mediolateral herniation	64	axilla compression	80 ■	L5	85
foraminal herniation	44	ganglion compression	56	S1	97
Kemp sign					
medial herniation	31	dural sac compression	22	L3	14
paramedian herniation	29	shoulder compression	30	L4	40
mediolateral herniation	38	axilla compression	45 ■	L5	48
foraminal herniation	20	ganglion compression	33	S1	58
Finger-floor distance 25 cm or more					
medial herniation	29	dural sac compression	37	L3	33
paramedian herniation	46	shoulder compression	44	L4	37
mediolateral herniation	60 ■	axilla compression	43	L5	39
foraminal herniation	18	ganglion compression	27	S1	57
Paresis					
medial herniation	14	dural sac compression	33	L3	14
paramedian herniation	18	shoulder compression	25	L4	55
mediolateral herniation	18	axilla compression	22	L5	28
foraminal herniation	44	ganglion compression	18	S1	20

Note- It should be noted that in patients without nerve root compression on MRI the four signs were positive in 50, 25, 35 and 12% of the cases.

Figure 5. Distribution of site of disc herniation per test

DISCUSSION

In the present study MRI investigations were assessed and correlated with clinical findings. It was shown that L3, L4, L5 and S1 nerve roots each tend to be compressed at different sites along the rostral-caudal course of the nerve root. It was also shown that disc herniations become symptomatic at different sites for each disc level. Finally, it was shown that clinical findings are related to specific patho-anatomical aspects.

The MRI and clinical investigations were more elaborate than described here. For the purpose of clarity, we decided to limit our investigation to the findings and anatomical aspects correlated presently. MRI may be argued not to be a gold standard of spinal anatomy. It is, however, the state of the art radiological imaging procedure of the spine and can be applied in all patients (unlike operation) and be correlated with clinical findings (unlike *in vitro* anatomical studies).

Endoscopical observations of the intrathecal nerve roots¹³⁰ and related anatomical work by one of the authors, and the observations by Falconer *et al.*, Reulen *et al.* and Cailliet had prompted us to draw a schematical model of the thecal sac, nerve roots, vertebrae and intervertebral discs (figure 6). This figure indeed seems to provide an explanation for the clinical data presented. It shows that L3 roots are more likely to be compressed towards the dorsal root ganglion and by a caudad L3L4 disc herniation, because the root sleeve exits below the L2L3 disc level. On the contrary, L5 roots are more likely to be compressed by the cephalad L4L5 disc herniation, because the sleeve exits directly posterior to the disc level. It also shows that S1 roots are more likely to be compressed in the axilla, because of an even more cephalad exiting point of the sleeve. Reulen *et al.* have shown that at each disc level, the inferior facet joint is situated higher relative to the disc (parallel to the higher exiting point of the root sleeve).

Figure 6.
Schematic drawing of vertebrae,
discs, dural sac and intra- and
extrathecal nerve roots

It seems that the typical symptoms of nerve root involvement (typically dermatomal pain distribution, pain increase on coughing, sneezing or straining, positive SLR and Kemp test) are associated most with compression of the nerve root in the axilla. It is imaginable that a compression in the axilla leads to a greater fixation of dural sac and nerve root (sleeve) than a compression of the shoulder. This may interfere more profoundly not only with nerve root dynamics, but also with the nutritional support of the extrathecal part of the nerve root^{131 132}.

Sitting down generally extends the lumbar spine. Parke and Watanabe¹³¹ have shown that the intrathecal cauda equina roots become tense on extension of the lumbar spine. In view of this, the nerve root would be especially vulnerable to compression at the shoulder. The clinical data indeed supported this.

Overall, mediolateral disc herniations were most likely to provoke the signs and symptoms of sciatica. Interestingly, an increased finger-floor distance was the most pronounced finding with this particular type of disc herniation. In another study of patients with pain irradiating into the leg¹³³, this sign was shown to be one of the strongest predictors of nerve root compression.

The author of this thesis gratefully acknowledges E. Beuls and H. van Mameren for the many inspiring sessions dealing with the anatomy of the spine.

The conservative treatment of sciatica

Not for their elevated thoughts will their books be searched through, but for some casual sentence, which allows conclusions
Bertolt Brecht

6 Systematic review of conservative treatment for sciatica due to disc herniation

---

ABSTRACT

Most patient with sciatica due to disc herniation are managed conservatively at first. Because the natural course is generally favourable, the additional value of many conservative treatment options, including bed rest, may be questioned. A systematic review of conservative treatment of sciatica is lacking. Therefore a systematic review was carried out.

A standardized literature search for randomized controlled trials (RCTs) was done. Trial methodology and outcome were registered independently by three observers. If possible, trial results were statistically pooled (= meta-analysis).

18 RCTs were finally selected. The methodological quality of these trials varied considerably. Pooled odds ratios were calculated for several treatment types. Neither traction, exercise therapy, manipulation nor drug therapy were unequivocally shown to be beneficial. Epidural steroids seem superior to placebo. No RCT of bed rest was found.

Clinical heterogeneity and methodological quality aspects had important effects on review results. Small sample sizes in most studies may partly explain the absence of significant effects in many studies. From the review results the natural course in secondary care patients was estimated. Overall, 51% of patients with sciatica were shown to have a favourable natural course over a period of 10 days. The authors recommend well-designed trials investigating the effectiveness of epidural steroids in acute sciatica, spinal manipulation and bed rest.

INTRODUCTION

Sciatica is a common complaint in both primary and secondary health care ^{4 238}. The lifetime cumulative incidence for low back complaints is 80%. Approximately 10% of these low back pain episodes will be accompanied by sciatica. The latter may be defined as pain irradiating into the leg supposedly due to lumbar nerve root involvement. Disc herniation is regarded as the most frequent cause.

After a diagnosis of nerve root involvement (and often disc herniation) has been established, the majority of patients are managed conservatively. The reported success rates of conservative therapy range from 35 to 97% in various case series^{134 135} and trials. Bed rest is regarded as the mainstay of conservative therapy^{135 114 136}. However, few recommendations for bed rest seem to be evidence-based. The trials by Deyo and colleagues and Malmivaara and colleagues have provided evidence against the usefulness of bed rest in acute low back pain^{137 138}. While bed rest as a treatment for low-back pain has been reviewed^{139 140}, it has not been reviewed as a treatment for sciatica.

The purpose of the study reported here was to systematically review the value of the conservative therapies, particularly bed rest, for the treatment of sciatica due to disc herniation. Trial results were statistically pooled, if the study characteristics permitted this. Because many studies have reported a spontaneous improvement of sciatic complaints ^{141 111 142}, we focused on randomized controlled trials (RCTs) since these control for this favourable natural course. Meanwhile, the placebo groups were thought to give an impression of the natural course of sciatica.

Finally, the effect of major study characteristics on review results was studied with special attention for the effect of clinical heterogeneity, as discussed by Thompson¹⁴³. In particular, it was hypothesized that the outcome might differ between studies of acute and studies of chronic episodes of sciatica or between studies of subjects with a definite radicular origin of their signs and symptoms and studies of subjects with a less well differentiated radicular syndrome.

METHODS

The literature was searched using the optimal search strategy for randomized controlled trials, published in the BMJ¹⁴⁴, combined with a search strategy for the disease entity (employing the terms "intervertebral dis* displacement", "dis? near herniat*", "sciatic*", "hernia nuclei pulposi", "nerve root involvement/entrapment", "dis? near rupture*", "dis? near slipped", "dis? near protrus*", "dis? near entrus", "dis? near sequestrat*", "nerve entrapment syndrome"). Included were phase I-IV studies according to Sackett¹⁴⁵ with subjects with sciatica as a well demarcated main or subgroup of the study were included. Excluded were studies not dealing with lumbar disc herniation, diagnostic studies, descriptive or review studies, phase V studies (case series/case report) according to Sackett, studies dealing exclusively with invasive procedures (chemonucleolysis, percutaneous nucleotomy, micro-/discectomy, fusion) and studies with less than 20 patients in the smallest group.

Three readers independently read the final selection of trials. The methodological assessment criteria had been used and published in several other reviews^{146 147 148 139}. Several criteria were slightly adapted for specific use in trials dealing with sciatica. Some other criteria for aspects of the quality of the report were found to be missing because of their indirect but important methodological implications. Therefore, several criteria were added based on the items 8, 13, 14, 15, 20, 22, 23, 27, 28, 30, 31, and 32 of the checklist proposed by The Standards of Reporting Trials Group¹⁴⁹. Studies were considered to provide good clinical differentiation of radicular pathology if both information was given on (1) typical radicular pain distribution, (2) positive nerve tension signs and (3) neurological complaints and at least two of these three were used as inclusion criteria. Studies were regarded as studies of acute disease if all patients had a duration of disease shorter than 3 months.

Pooling was considered possible if more than one placebo-controlled trial addressed a particular therapy. The random effects method described by Dersimonian and Laird²⁵ was used for statistical pooling of the odds ratios for improvement. Improvement was defined as improvement in pain, in the global measure of improvement or return to work. The odds ratios are presented on a logarithmical scale (figure 1) because point estimates and the width of the 95% confidence interval are then more realistically presented¹⁵⁰. Overall odds ratios for several study characteristics were presented in figure 2. Linear regression of the log (odds) on these characteristics weighted by the estimated variance¹⁵¹ did not reveal statistical significance for the trends observed. A sensitivity analysis was done for the epidural steroid intervention with calculation of odds ratios if only studies using comparable groups, blinding of outcome assessors and intention-to-treat analysis were included (see epidural steroid (Q) in figure 1). The exclusion criterion "less than 20 patients in the smallest group" most affected the estimation of epidural steroid effectiveness as 8 of 11 studies were excluded¹⁵²⁻¹⁵⁹. Therefore, the pooled odds ratio for epidural steroid treatment was additionally determined for the total of 11 studies to investigate the effect of this exclusion criterion (see epidural steroids (11) in figure 1).

To provide an impression of the natural course the percentage of patients in the placebo group with partial or complete improvement was plotted against the time interval (figure 3). The outcome of placebo in the four groups of the Coxhead study¹⁶⁰ was averaged for the estimation of the natural course, but because of the unknown effect of interactions the Coxhead study was exempted from pooling.

RESULTS

Automated searches of Medline (Silverplatter version/SPIRS) for the period 1966-March 1995, and of Embase/Excerpta Medica for the period 1984-1995 resulted in the retrieval of 1841 and 1310 studies, respectively. A hand search based on references in the literature and references by experts in the fields resulted in 5 additional studies. Nineteen studies were finally selected (references 160-178). As one study was discussed in two papers, eighteen were read by three readers (MdK, PS, PV).

The 18 trial reports described 24 RCTs. The methodological scores for each of these 24 RCTs in the 18 papers are depicted in table 1. 22% of the 24 RCTs had a poor description of the randomisation procedure, 44% a poor or unknown extent of homogeneity of the research population, and 78% a small sample size. Patients, outcome assessors and data entry/analysis were not blinded in 30, 30 and 94% of studies, respectively. 60% had a limited effect measurement, 40% had no intention-to-treat-analysis, and 70% had no variability measures for the data reported. In none of the trial reports was an evaluation of the efficacy of the blinding procedures discussed.

Table 2 shows the essential characteristics of the 15 placebo-controlled RCTs. The 9 trials using a non-placebo control group are described in table 3. Overall, two RCTs^{161 163} investigated primary care and 22 RCTs secondary care populations. Eleven RCTs provided "good differentiation from non-radicular pathology". Nine studies included patients with sciatica of less than 3 months duration only. Seven studies met the three major requirements of adequate comparability of groups, observer blinding and intention-to-treat analysis. The pooled odds ratios for these seven studies was lower than for the remaining lower quality studies (non-significant). The same trend of lower pooled odds ratios could be observed for acute versus non-acute disease, and for studies with better versus studies with poorer clinical differentiation.

The effect measure in most studies was a patient derived global measure of improvement (gmi) or perceived effect. On five occasions^{171 177 169 168 176} was the experimental intervention shown to have a statistically significant effect. One study with good methodological quality ¹⁷¹ showed a significant effect. Of the 13 studies showing non-significant differences, ten had sample sizes smaller than 50 patients per group.

Table 1. Methodological scores per study

Study	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	SS	=sum score
Weber, 1993	2	3	1	0	5	4	8	6	3	3	9	0	9	4	2	1	8	1	1	0	2	72	
Hofferberth	1	1	5	0	7	0	8	6	0	3	9	3	9	2	2	0	8	1	1	3	2	71	
Snoek	2	3	1	0	7	0	8	6	3	3	9	0	9	3	2	1	8	1	1	0	2	69	
Goldie	2	3	0	1	6	0	8	6	3	0	9	3	9	1	1	0	8	1	0	0	2	63	
Berry	2	0	1	0	3	4	8	6	3	3	7	0	9	4	1	1	8	2	1	0	0	63	
Gibson	1	2	0	0	5	0	8	6	3	3	9	0	9	2	1	0	8	2	1	0	2	62	
Dilke	2	3	1	0	0	0	8	6	3	3	9	3	9	2	1	0	8	1	1	0	2	62	
Weber, 1984	1	1	1	0	0	0	8	6	3	3	9	0	9	2	0	1	8	0	0	0	2	57	
Mathews, 1	1	2	1	0	6	0	8	6	3	3	9	0	9	2	2	1	0	1	0	1	0	56	epidural steroid intervention
Ljunggren, 1	2	2	1	1	7	0	8	3	3	3	0	0	7	4	0	1	8	0	1	0	2	53	manual traction vs. exercises (1992)
Porsman	2	0	3	0	6	0	8	6	0	3	9	0	9	3	1	0	0	0	0	0	2	52	
Weber, 1980	1	1	0	0	6	0	8	6	3	3	7	0	7	1	0	1	0	1	1	0	2	48	
Coxhead	1	1	1	0	4	11	5	0	0	3	0	0	0	4	2	1	8	2	1	0	2	46	
Coxhead	1	1	1	0	4	11	5	0	0	3	0	0	0	4	2	1	8	2	1	0	2	46	
Tesio	1	0	0	0	6	0	8	3	3	3	0	0	9	6	2	1	0	1	1	0	2	46	manipulation intervention
Mathews, 2	1	1	1	0	6	4	8	0	3	3	0	0	9	2	2	1	0	1	1	0	2	45	
Weber, 1983	2	1	5	0	7	4	5	0	0	3	0	0	0	3	2	1	8	1	1	0	1	44	
Coxhead	1	1	1	0	4	11	5	0	0	0	0	0	0	4	2	1	8	2	1	0	2	43	
Coxhead	1	1	1	0	4	11	5	0	0	0	0	0	0	4	2	1	8	2	1	0	2	43	
Borns	1	1	1	0	6	0	8	3	0	0	9	0	0	1	1	0	0	2	1	0	2	36	
Ljunggren, 2	1	4	1	0	7	0	8	3	0	3	0	0	0	5	1	1	0	1	1	0	0	36	autotractor vs manual traction (1984)
Mathews	1	0	1	0	4	4	8	0	3	0	0	0	7	1	1	0	0	1	1	0	2	34	traction intervention
Larsson	2	4	1	0	6	0	8	0	3	0	0	0	0	1	1	1	0	1	1	0	1	30	

Study population: A homogeneity (2), B comparability of prognostic variables (4), C randomization adequate (5), D intended sample size justified (1), E number/flow of participants (4), drop-outs < 20% (1), < 10% (2), loss to follow-up (3), F sample size > 50 in the smallest group (4), > 100 in the smallest group (7)

Interventions: G experimental + reference intervention described (8), H control group (6), I co-interventions avoided/ equally divided (3), J compliance adequate (3)

Table 2. Description of studies with placebo group

	Q	P	D	R	A	T	Experimental group	Placebo group
Drugs								
Weber, 1993, (72) VAS pain/return to work	Y	Y	N	N	Y	30	Piroxicam 40 mg/d for 2 days, 20 mg/d for 12 d	Placebo tablet, same dosage
	52/120	vs		41/94				
Goldie (63) improvement in pain	Y	Y	N	N	Y	14	50 times 25 mg indomethacin thrice-daily	Placebo tablet, same dosage
	15/25	vs		18/25				
Weber, 1983 (48) improvement in pain	N	Y	N	N	Y	5	Phenylbutazone, 1200 mg/d for 3 d, 600 mg/d. for 2 d.	Placebo tablet, same dosage
	14/27	vs		8/30				
Traction								
Weber, 1984 (57) GMI	Y	Y	Y	Y	N	7	1. Intermittent Tru-trac, 1/3 of weight, 20min./d. 5-7 d.	7 Kp traction simulation
	19/37	vs		16/35				
							2. Spina-Trac of 40-70 Kp, 20min/d. for 5-7 d.	fictitious traction
	5/21	vs		5/23				
Mathews, 3 (34) GMI	N	N	N	N	Y	4	> 45 kg traction force each weekday for 7 days	infra red heat thrice-weekly
	40/77	vs		27/54				
Larsson (30) GMI	N	N	Y	N	Y	21	Auto-traction ad mod. Lind + corset	corset, no simulation of traction
	20/41	vs		8/41				
Epidural steroids								
Snoek (69) GMI	Y	Y	Y	Y	Y	2	2ml methylprednisolone 80 mg	epidural 2 ml saline
	18/27	vs		10/24				
Dilke (62) pain/return to work	Y	Y	Y	N	N	30	10 ml methylprednisolone	1 ml saline in intraspinal lig.
	16/35	vs		4/36				
Mathews, 1 (56) GMI	N	Y	Y	N	Y	30	2 ml methylprednisolone acetate 80 mg 1-3 times	2 ml lignocaine in sacral hiatus
	14/21	vs		18/32				
Intramuscular dexamethasone								
Hofferberth (71) GMI	Y	Y	N	N	N	11	schedule per day : (24-24-24-24-24-12-12-12-8-4	neutral ampulla (?) injection
	27/38	vs		29/53				
Persman (52) GMI: complete	N	Y	Y	N	N	9	4mg/ml intramuscularly for 7 d (16-8-8-4-4-4-4)	0,9 % saline injection
	13/25	vs		14/24				
Other therapy								
Gibson (62) GMI: complete	Y	Y	N	N	N	7	4 times a day 2 tablets of Chymoral	4dd2 placebo tablets
	7/46	vs		10/47				
Mathews, 2 (54) GMI	N	N	Y	Y	Y	14	Manipulation each day if necessary	Infra red heat thrice-weekly
	98/123	vs		56/84				
Berry (63) VAS	N	Y	N	N	Y	7	tizanidine 12 mg/days for 7 days	Placebo tablets, same dosage

Legend to table 2

Studies are mentioned by the first author's name and the methodological score in brackets. The second line of the first column shows the outcome measure used (GMI= global measure of improvement, VAS= visual analogue scale).

The 2nd column shows the scores (Y=yes or N=no) for the Q/P/D/R/A/T categories (Q= meeting major methodological prerequisites (comparability of groups, blinded outcome assessment, and intention-to-treat analysis), P= patient blinding, D= good differentiation of radicular from non-radicular pathology, R=MRI/CT/myelography diagnosis, A= acute disease. T= timing of outcome measurement in days).

The 3rd and 4th column feature a description of the intervention with numbers. The numbers in bold represent the number of successfully treated patients per group divided by the total number of patients per group (there are no numbers for the Berry report as this study featured Visual Analogue Scale scores instead of number of patients improved).

Table 3. Description of studies without placebo group

		Q	P	D	R	A	T	Intervention 1	Intervention 2
Ljunggren, 1	(53)	N	N	Y	Y	N	7	Manual traction, 5-7 days, 300 N	Isometric exercises, 5-7 d
50 patients, GMI		no significant difference							
Tesio	(46)	N	N	N	N	N	14	Autotraction ad mod. Natchev	anual traction 35% of weight
44 patients, GMI		in favour of 1							
Weber, 1983	(44)	N	N	Y	Y	N	365	Discectomy	6 weeks of physical therapy
126 surgical candidates, GMI		in favour of 1							
Coxhead	(46)/(43)	N	N	N	N	N	28	Manipulation (Maitland)	factorial design
resp. 322, 292, 292 and 292 patients, GMI		no significant differences						Manual traction Tru-Trac Exercises Corset	
Borns	(36)	N	Y	N	Y	N	4	Tiaprofenic acid 400mg intramuscularly for 4 days	Ketoprofen 200 mg intramusc. for 2 d.
40 patients, GMI		no significant differences							
Ljunggren, 2	(36)	N	N	Y	Y	N	14	Autotraction ad mod. Lind	Manual traction 1/3 to 1/1 of weight
49 patients, GMI		no significant differences							

Note- The table 2 legend holds for table 3 with the exception of the text in bold. In table 3 the text in bold qualitatively summarizes between-group differences in outcome.

The Coxhead study used a factorial design with each intervention being compared with no or 1-3 of the other interventions. Because compliance with exercises and corset was not described, the experiments involving these 2 therapies had a methodological sum score of 43 instead of 46.

Pooling was possible for (1) non-steroidal anti-inflammatory drugs (NSAIDs), (2) traction therapy, (3) intramuscular steroids, and (4) epidural steroids. The results are presented in Figure 1. The pooled odds ratio for the 3 RCTs investigating the effect of epidural steroids versus placebo was 3.0 (95% confidence interval 1.3-6.8).

Figure 1. odds ratios per study (with placebo group) and pooled odds ratios

Note- Tests of homogeneity; NSAID: chi-sq.=4.33, 2df, p>0.1, Traction: chi-sq.=4.82, 3df, p>0.1, Steroids i.m.: chi-sq.=1.74, 1df, p>0.1, Epidural steroids: chi-sq.=2.68, 2df.

In a sensitivity analysis, the odds ratio was 4.2 (1.8-9.9) if only the higher quality trials were pooled and 2.3 (1.2-4.4) if the 8 RCTs that were excluded because of sample size below 20 per group would have also been included in the pooling. No significant effect was demonstrated for NSAID's, intramuscular steroids or traction. In other randomized trials, several modes of traction were compared with equivocal results: while autotraction was reported to lead to greater improvement in complaints than wearing a corset or manual traction after respectively one and two weeks, two other trials showed poorer results for autotraction than for manual traction.

Chymoral, tizanidine and manipulation treatment were addressed by one single placebo-controlled RCT. Only the latter therapy was shown¹⁶⁹ to be superior to placebo (local heat) after two weeks. The study with a factorial design¹⁶⁰ showed a small positive trend. The methodological quality of these twotrials was better. A negative but non-significant trend was reported for chymoral.

In the placebo and heat groups 51% (±20) of patients had partially or completely recovered after 10 days (±9). The outcome was positively correlated with the time interval until effect measurement (see figure 2). It did not depend on the methodological score. The outcome of placebo/heat intervention did not differ significantly between patient groups with acute sciatica and groups with chronic sciatica.

No RCT could be found studying the use of bed rest. 7 studies specified using bed

rest as a co-intervention for an average of 7 days. Placebo interventions actually involved strict bed rest in 4 of the studies. In 2 studies bed rest was not given, while the other 9 did not specify the role of bed rest as a co-intervention.

Figure 2. The recovery over time

Note- The straight line is a linear transformation of the observed recovery rate (SPSS curve estimation, $p,0.05$).

DISCUSSION

There seems to be insufficient evidence supporting the effectiveness of most of the conservative therapies for sciatica. Only a few therapy types were the subject of multiple placebo-controlled trials. Of these types only epidural steroids proved to be effective. The other therapy types were the topic of isolated RCT's that generally had considerable methodological flaws. Of the latter only manipulation had a positive effect. Bed rest has not been studied in a RCT.

The results of any review are determined to a fair extent by the literature search, the selection criteria and the scale used¹⁷⁹. Our search was an elaboration of the optimally sensitive search strategy reported by Dickersin et al¹⁴⁴. Medline and Embase do not include all physical therapy journals, however. Particularly in this field some relevant studies may not have been selected. As reported elsewhere¹⁸⁰ a publication bias and the risk of false positive results may affect systematic reviews. Both risks are more prominent in smaller studies¹⁸¹. Also, the chance of significant prognostic imbalances

between intervention groups is greater. Thus, studies with less than 20 patients in the smallest group were excluded. The impact of this exclusion criterion on the effect estimation for epidural steroids was shown to be small.

In many studies the details of the clinical picture and the randomization procedure were not provided. Information on the number of patients with typically radicular pain, positive nerve tension signs or neurological complaints is required for assessing study homogeneity¹⁴³. Such information also helps the reader to assess the degree of likelihood that the patients investigated indeed had radicular pathology and not diseases that lead to sciatica-like pain without affecting disc or nerve root (e.g. tendomyogenous or hip afflictions). Moreover, if patient populations have different clinical pictures the natural course and response to conservative therapies may differ. A clear delineation of a radicular syndrome is complicated by the varied terminology for disorders of the lumbar intervertebral disc and nerve root (for example, see the search strategy under methodology). If the clinical syndrome of subjects in the research populations is not adequately described, a trial in "disc herniation" patients may not be assumed to be equivalent to a trial in "sciatica" patients. Figure 2 serves as an example of the effect that differences in patient population's clinical pictures may have on RCT outcome. Pooling of all placebo-controlled trials showed an overall effect of conservative therapy. When only the studies with good differentiation of radicular from non-radicular pathology were pooled, the pooled odds ratio tended to be lower. This suggests an overestimation of therapy effectiveness in populations of patients selected without firm diagnostic criteria for sciatica. Figure 2 also shows a trend for RCTs with good comparability of intervention groups, blinded outcome assessment and intention-to treat analysis to yield lower estimates of effectiveness (albeit non-significant in a weighted linear regression analysis). It underlines the importance of interpreting trial results in the context of study methodology.

With regard to the methodological flaws, the following remarks are made. Firstly, the absence of blinding of patient or observer in 30% of studies does not reflect the fact that for some interventions patient blinding is impossible. It also does not reflect that for some outcome measures absence of observer blinding introduces less bias than for other measures. Strikingly, a blinded data entry and analysis was lacking in all studies but one. Secondly, various trials report the differences between experimental and reference intervention to be non-significant. Insufficient sample size may explain this result in 77% of the studies. In all these cases the sample size was smaller than 100. Larger trials are definitely required for positive statements on treatment efficacy. For example, according to Pocock a total of 196 patients would be required¹⁸¹ for establishing a difference of 20% with statistical significance (assuming $\pi_1=60\%$, $\pi_2=40\%$, $\alpha=0,05$ and $1-\beta=0,80$). Finally, some methodological shortcomings may result from incomplete reporting and are not necessarily shortcomings of the trials themselves. Limitations set forth by the publishing journal may contribute to such incomplete reporting.

Non-significant tests of homogeneity do not exclude heterogeneity and are therefore an inadequate justification for a fixed effects approach. The pooled odds ratios were therefore determined by the more robust random effects method of

Dersimonian and Laird²⁵. The result were not substantially different from the pooled Mantel-Haenszel or Peto's odds ratio.

Figure3. Pooled odds ratios for subgroups

According to our review, improvement in sciatica patients is greater for epidural steroids than for placebo treatment. No serious side effects were reported for a total of 437 included patients, while 26 patients complained of transient headache or transient increase of sciatic pain. The 1997 RCT of 80 mg/ 10ml. epidural methylprednisolone vs a 2 ml saline solution by Carette *et al.*¹⁸² was beyond the time frame of this review's literature search. Nevertheless, it deserves special attention here, because its sample size was much larger than that of the epidural steroid RCTs selected (if the RCT by Carette *et al.*¹⁸² had been included with the other 3 RCTs in our study, the pooled odds ratio for the effect of epidural steroids would have been 2.2 (95% CI 1.0-4.7)). In the trial by Carette *et al.*, pain in the leg after 6 weeks had been somewhat less in the methylprednisolone group and there had been no significant difference in primary outcome or number of patients improved. On average the patients had had sciatica for 13 weeks. Thus, the absent or small effect may apply particularly to chronic cases of sciatica. In contrast, the pooled odds ratio for previous RCTs of epidural steroids was significant in our study. Strikingly, two of the three RCTs^{170 169} in our review exclusively investigated patients with a disease duration of less than 3 months. Hence, epidural steroids may be more effective in acute cases of sciatica. It should be remarked, however, that a publication bias¹⁸⁰ may be a cause of overestimating odds ratios in any meta-analysis. Future research is recommended to study this as well as other research goals such as the investigation of the additional benefit of fluoroscopic guidance for steroid application¹⁸³ and of the value of short interval repeated injections¹⁸⁴. The pooled results for traction and NSAIDs suggest they do not benefit

the sciatica patient. Naturally, effectiveness of particular types of traction or NSAID therapy cannot be excluded but none of the types seems clearly superior to the other. Manipulation may have beneficial effects on radicular complaints. The study supporting this, however, had significant shortcomings (e.g. no intention-to-treat analysis, no control for attention and expectations of the patients, no blinding of patients).

If placebo treatment is regarded as a reflection of the natural course, half of the patients showed an improvement after one to two weeks, while partial or complete improvement may occur in more than 70% in the long run. This has also been described in other studies¹³⁵. Follow-up radiological studies have reported that a decrease in disc herniation size to occur in 60-76% of cases^{185 111} over a period of 0.5-2 years, but spontaneous regression of herniated disc only partly explains the favourable natural course. Recurrences are said to occur in 44% of cases over an 8 year interval¹³⁵, while Weber reports a 24% recurrence rate over 4 years for the conservatively treated group. It should be acknowledged that bed rest is often an unspecified contributor to most “placebo” treatments and as such may alter the natural course to an unknown extent.

Notwithstanding the absence of a significant effect of bed rest for acute low back pain^{138 137}, bed rest is one of the mainstays of the conservative treatment of lumbar disc herniation with sciatica. We found no randomized controlled trial specifically studying its efficacy, however.

We recommend that a report on a RCT studying sciatica discusses the clinical syndrome and the differentiation from non-radicular pathology in detail to allow better assessment of heterogeneity both between intervention groups and between studies. The blinding of patient, caregiver, outcome assessor and data analyst and the subsequent evaluation of its efficacy should also receive more attention. Epidural steroids seem to be beneficial in sciatica patients and warrant further investigation. Spinal manipulation needs to be further investigated. There is no evidence that traction, non-steroidal anti-inflammatory drugs, intramuscular steroids, tinnazidine or chymoral are superior to placebo. It seems appropriate to study the use of bed rest in a randomized controlled trial.

ABSTRACT

Bed rest is widely advocated for sciatica and other signs and symptoms of a lumbosacral radicular syndrome, but its effectiveness has not been established.

To study the effectiveness of bed rest in patients with a lumbosacral radicular syndrome of sufficient severity to justify a treatment with 14 days of bed rest, a trial was conducted in a primary care setting in Maastricht and surrounding villages. 183 subjects were randomly assigned to either bed rest or watchful waiting for a duration of 14 days. Primary outcome measures were global measures of improvement over 2 and 12 weeks. Secondary measures were a variety of functional and pain scales (at 2, 3 and 12 weeks), work absenteeism and surgical intervention. Both the observer's assessment of the outcome and the data entry and analysis were blinded.

No significantly higher rate of improvement was found in the 92 patients in the bed rest group than in the 91 in the watchful waiting group neither at two nor at twelve weeks. Visual Analogue Scales of pain intensity and bothersomeness of complaints, McGill Pain Questionnaires, and functional status questionnaires revealed no significant differences between the two groups. Work absenteeism and surgical intervention rates were similar for both groups.

Bed rest does not provide a better chance of recovery for patients with sciatica and other signs and symptoms of a lumbosacral radicular syndrome. Therefore, such patients may rest in bed, but may mobilize when possible.

Keywords: bed rest - sciatica - lumbosacral nerve root involvement - RCT

INTRODUCTION

Pain irradiating into the leg (sciatica) and signs and other symptoms of lumbosacral nerve root involvement^{18 76} indicate a lumbosacral radicular syndrome (LSRS). This syndrome is often due to a disc herniation¹¹⁶. A variety of conservative strategies may add to the benign natural course of the syndrome. Among these therapies, bed rest is generally regarded as a mainstay. Similarly, bed rest used to be regarded as a valuable treatment for low back pain up to a decade ago, although there was little objective data¹⁸⁶ to support this view. Since 1986, evidence of the inefficacy of bed rest in low back pain patients has accumulated^{137 138}. Though bed rest is widely used to treat symptoms and signs of lumbosacral nerve root involvement¹¹⁴, a randomized controlled trial (RCT) to justify this has been lacking¹⁸⁷. In fact, the value of bed rest seems so undisputed that it has been an integral part of many control treatments in trials of other conservative therapies. Yet, bed rest may be detrimental to the patient's mental and physical condition¹⁸⁸⁻¹⁹⁶ and it may be associated with work absenteeism and -in the case of hospitalisation- with substantial medical costs. Also, the optimum duration of bed rest is unknown. An inquiry by Van Leeuwen and Perquin¹¹⁴ and an inquiry by the authors among 96 general practitioners in the Maastricht region revealed that the duration most frequently advocated was 14 days.

Therefore, a randomized, controlled, observer-blinded trial was designed to evaluate the efficacy of bed rest for a duration of 14 days for the treatment of a lumbosacral radicular syndrome. A control group termed the "watchful waiting" group was designed to resemble the natural course. It was hypothesized that the success rate of bed rest would be at least 20% higher than the rate in the control group (both before and after adjustment for possible confounders and effect modifiers defined *a priori*).

METHODS

The study protocol

The study was approved by the institutional review board of the Maastricht University Hospital. The intended sample size was 196 (assuming two-sided testing, $\alpha=0.05$, $1-\beta=0.80$, success rates 50 and 70%¹⁸¹). For the sample size calculation, success was defined as any improvement over 2 weeks. Other primary outcome measures were the patient's perception of reasonable improvement after 12 weeks, of strong improvement after 2 weeks and the blinded observer's perception of any improvement after 2 weeks, strong improvement after 2 weeks, and reasonable improvement after 12 weeks. By way of secondary outcome measures, the rate of further specialist evaluation (indicated by the blinded observer) and surgical intervention were compared. Additionally, Visual Analogue Scales (VAS) of the bothersomeness of the major complaint¹⁹⁷ and of the separate pain intensities in the leg and low back, averaged for the past week, were compared. The scales ranged from 0 to 100 mm. Differences in the qualitative description of pain were investigated by the

McGill Pain Questionnaire. Patients choose words that best fit their pain resulting in a total Pain Rating Index¹⁹⁸. Furthermore, a modified Roland Disability Scale (Roland DS)^{199 200} and the revised Oswestry Low Back Pain Questionnaire (Oswestry LBPQ)²⁰¹ were used to evaluate the back-specific functional status²⁰². Patients came in for follow-up visits at 2, 12 and 52 weeks after the time of inclusion. They were requested to fill out questionnaires after 3 and 26 weeks.

Patient selection

Between February 1995 and December 1996, 50 general practitioners (GPs) in Maastricht and surrounding villages asked patients to participate in the study. Patients with a first time presentation to the GP for the current complaints were referred to the Neurology Department of the Maastricht University Hospital if they had (1) pain radiating into the leg below the gluteal fold, and (2) pain intensity sufficient to justify bed rest for 14 days as a therapy. Patients were excluded on the basis of (1) previous spinal surgery, (2) pregnancy, (3) penitentiary confinement, (4) unavailability for follow-up visits (moving plans, severe disease), and (5) co-morbidity of such a nature that their well-being was primarily determined by this co-morbidity. GPs were asked to register the eligible patients not referred. All patients had baseline examinations within two days of referral. In addition to history and physical examination, magnetic resonance (MR) imaging of the lumbar spine was performed consisting of T1-weighted sagittal and transverse cuts, sagittal proton density/T2-weighted cuts and coronal-oblique MR myelographic images. The assessment of the MR images focussed on nerve root compression. The researchers then reascertained the selection criteria and applied two further criteria. Patients were required to have lumbar nerve root involvement (with at least two of the following: typically radicular pain distribution, pain in the leg increasing with coughing, sneezing, or straining, decreased muscle strength, sensory loss, reflex loss, positive nerve tension test). They were excluded if there was an indication for direct surgical intervention, due to either (a) intractable pain, (b) a rapid progression of a paresis of short duration or (c) cauda equina syndrome.

The intervention assignment

A randomization schedule, balanced by blocks of ten, was made using a computer-generated random number table. A research nurse allocated patients to the interventions according to their case number, and instructed them on the nature of the intervention. She gave a written instruction and an intervention-specific diary. The instruction was standardized to minimize instruction and attention related differences in expectation for the two groups.

The interventions

Patients allocated to the bed rest treatment group were instructed to stay in the supine or the lateral recumbent position with one pillow under the head for a duration of 14 days. Toilet visits and showering were permitted. Patients in the "watchful mobilizing" group were instructed to mobilize whenever this was possible, without

clearly straining the back or provoking the pain. Bed rest was not prohibited. Patients could call on the researchers around the clock in case of intolerable pain or worsening paresis. The neurological status and the feasibility of another week of watchful waiting or bed rest was assessed in the patient's home after one week by a research assistant.

Compliance

Compliance with the allocated intervention was recorded daily in an anonymous patient diary. Pain and functional status were also registered in the diary.

Co-interventions

A pain medication algorithm was applied (paracetamol 1000 mg q8h was standard, supplemented by codeine 10-40 mg q4h or substituted by naproxen 500 mg q8h when necessary). Temazepam 10 mg a day was to be taken in case of insomnia. Other interventions for the radicular complaints were advised against.

Blinding of outcome assessment, caregiving and data entry and analysis

Before and after intervention allocation, and before each follow-up visit, patients were instructed to conceal the nature of the intervention for the blinded outcome assessor. The caregiver was blinded in all cases but a few in whom the outcome and nature of the trial intervention were thought to be helpful in planning further management. The group code meanings were unmasked after data entry, analysis, and the compilation of tables 2-5. Both patient, caregiver and outcome assessor were blinded for the MRI results until surgical intervention was indicated.

Data control, entry and analysis^{89, 181}

P-values were calculated for differences in baseline characteristics as a means of quantifying the differences observed. They should not be interpreted as the usual chance statistics, because -with proper randomization- all differences in baseline findings can only be due to chance.

The outcome was analyzed according to the "intention to treat" principle. An initial simple analysis compared the primary outcome of both groups using a chi-square test. Changes in VAS and Questionnaire scores were compared using Mann-Whitney tests. Work absenteeism and time to surgical intervention were compared using Kaplan-Meier curves and logrank tests⁸⁹. Subsequently, the data were analysed with adjustment for covariables. The primary outcome was used as a dependent variable in a logistic regression model¹⁰⁰. The model included the following possible confounding covariates: age (16-40, 41-50 and 51-81 years), gender, a VAS of observer's opinion of disease severity, a VAS score for the bothersomeness of the major complaint and the four covariates defined a priori - (1) nerve root compression on MR Imaging, (2) paresis, (3) duration of disease and (4) previous episodes of pain in the leg. Interaction terms for the latter four covariates and the intervention group were added to investigate effect modification. Similarly, confounding and effect modification of the secondary outcome measures were investigated using linear⁸⁹ and

Cox regression²⁰³ (only for the duration of work absenteeism).

RESULTS

Trial profile (table1).

Of the 338 patients referred by the GPs, 227 were considered eligible for the trial. 44 of these patients declined participation. They had slightly less severe disease (as

Table 1. Trial profile
338 patients referred

97	not meeting inclusion criteria
	5 - no pain radiating below gluteal fold 30- insufficient pain in the leg 62- no radicular syndrome tendomyogenous (29), hip pathology (7), spinal canal stenosis (2), mono-neuropathy (2), psychogenous (3), peripheral neuropathy (1), unclear (18)
58	meeting exclusion criteria
	1 - pregnancy 2 - previous spinal surgery 1 - serious co-morbidity 44-declining participation unwillingness to comply with bed rest (18), previous bed rest (15), unwillingness to miss work (5), improving already (3), bed rest increases pain (3) 6 - short term surgical intervention indicated paresis < 24 hours old (5), intractable pain (1) 4 - requiring further diagnostic work-up cauda equina syndrome (2), malignancy (2)

183 patients selected & randomized

0 drop-outs

2 loss to follow-up at 2 weeks

unwillingness to return (2); both reported improvement

1 bed rest

1 mobilisation

14 loss to follow-up at 3 months

unwillingness to return (10), moved (2), serious unrelated illness (2), penitentiary confinement (1)

7 mobilisation

7 bed rest

scored on the VAS by the observer), but had similar characteristics elsewhere. 183 patients were eventually selected and randomized; none dropped out before completion of the interventions. One patient in each intervention group did not return after two weeks; both reported improvement. After three months, 7 patients in each group were lost to follow-up. Their baseline characteristics and success rates after 2 weeks were similar to those of the 169 remaining patients. 12 patients in the bed rest and 10 in the control group did not undergo MR Imaging. Their baseline characteristics were similar to those that were scanned.

Outcome

Neither after two weeks nor after three months was there any significant difference in primary outcome between the bed rest group and the watchful waiting group.

Table 2. Comparability of patient groups

Characteristic	Bed rest N = 92	Watchful waiting N = 91
Age	44 (12)	48 (12)
Sex (% male : % female)	51:49	61:39
Duration of pain (median in days)	16	15
Any previous episode of pain in the leg (%)	34	39
Satisfaction score	4.2 (1.3)	4.5 (1.4)
Visual Analogue Scale (mm)		
Major complaint	85 (16)	87 (16)
pain in the leg †	62 (22)	68 (21)
pain in the back	49 (30)	45 (33)
McGill Pain Questionnaire PRI-T	19.1 (10)	19.7 (11)
Revised Roland Disability Scale	5.5 (3.9)	5.2 (3.8)
Oswestry Low Back Pain Questionnaire	27.3 (9.5)	29.0 (7.9)
Paresis (%)	14	23
Sensory loss (%)	41	43
Reflex differences -		
ankle tendon (%)	20	21
knee tendon (%)	3	9
Definite clinical diagnosis	84 %	88 %
Nerve root compression on MR Imaging	53 %	66 %

Note- Average values with standard deviations in brackets are indicated in the 2nd and 3rd column unless indicated otherwise in the first column, either by the median, a fraction of patients in relation to the total number of patients per group or the male to female ratio for sex.

11 patients did not undergo MRI scanning. There were 3 missing values for Roland DS and Oswestry LBPO and 1 for the VAS score of the major complaint.

PRI-T: Total Pain Rating Index

MR: Magnetic Resonance

mm: millimeters

† = p-value for difference between groups < 0.05

table 3). The rates of reporting for "any" and "strong" improvement were similar. These results were confirmed by the logistic regression analysis with adjustment for baseline differences in age, gender, duration of disease, nerve root compression on MR Imaging, paresis distribution, previous episodes of sciatica, VAS of the observer's opinion, and VAS of the bothersomeness of the major complaint.

Table 3. Primary outcome

Outcome Measure	Bed rest %	Watchful waiting %	Difference % (95%CI)	Adjusted odds ratio (95% CI)
after 2 weeks				
Effect perceived by patient				
Any improvement	70	65	5 (-9,18)	1.2 (0.6-2.3)
Strong improvement	37	35	2 (-12, 16)	1.1 (0.6-2.0)
Effect perceived by observer				
Any improvement	73	65	8 (-5,21)	1.4 (0.6-2.6)
Strong improvement	17	21	-4 (-15,8)	0.7 (0.3-2.0)
after 12 weeks				
Effect perceived by patient				
Improvement	87	87	0 (-10,10)	1.0 (0.4-2.9)
Effect perceived by observer				
Improvement	86	89	-3 (-13,7)	0.6 (0.2-1.7)

Note- The 2nd and 3rd columns present fractions of successfully treated patients in relation to the total number of patients in the groups. The 4th column shows the unadjusted difference with the limits of the

There was no significant difference in any of the secondary outcome measures after two and three months (table 4). Of 21 patients with muscle paresis in the control group, 2 had recovered after 2 weeks, and an additional 12 after 3 months. Of the 13 patients with a paresis in the bed rest group, 4 had recovered after 2 weeks and an additional 4 after 3 months. Furthermore, 4 patients had developed a paresis during watchful waiting compared to 2 patients during bed rest.

There were no significant differences in three of the four subgroups defined a priori: the duration of disease, paresis and nerve root compression on MR Imaging (in other words the interaction terms in the model were not significant). Neither the reported improvement, nor the VAS scores or functional status were differentially affected. The fourth covariate, previous episodes of sciatica, showed a higher rate of "any improvement" in the bed rest group than in the control group after two weeks (table 5). The significant effect modification of this subgroup was not confirmed when the rate of strong improvement or any of the secondary outcome measures after two

Table 4. Secondary outcome measures

Outcome parameter	Mean Bed rest (SD)		Mean watchful waiting (SD)		p-value for change	Adjusted Difference	95% CI
after two weeks							
Satisfaction score	5.6	(1.7)	5.8	(1.5)	0.41	-0.1	(-0.6,0.3)
Visual Analogue Scale (mm)							
Major complaint	45	(34)	49	(33)	0.74	2.9	(-7.0,12.9)
Pain in the leg	36	(28)	44	(27)	0.60	2.3	(-5.4,10.0)
Pain in the back	30	(27)	30	(28)	0.16	1.4	(-6.1,8.8)
McGill Pain Questionnaire	13.3	(10)	14.8	(11)	0.30	1.3	(-1.7, 4.3)
after three weeks							
Revised Roland DS	8.2	(6.2)	9.2	(6.3)	0.07	-1.6	(-3.7,0.4)
Oswestry LBPQ	21.5	(11)	20.1	(11)	0.34	-2.7	(-1.0,6.4)
after three months							
Satisfaction score	7.3	(1.6)	7.5	(1.2)	0.59	-0.1	(-0.6, 0.3)
Visual Analogue Scale (mm)							
Major complaint	20	(31)	18	(28)	0.36	-1.4	(-10.2,7.5)
Pain in the leg	16	(26)	14	(24)	0.06	-1.6	(-9.2,6.0)
Pain in the back	19	(25)	22	(27)	0.51	5.3	(-2.2,12.8)
McGill Pain Questionnaire	7.5	(9)	6.8	(8)	0.74	-0.6	(-3.3, 2.1)
Revised Roland DS	15.2	(7)	15.7	(7)	0.32	-0.5	(-2.6,1.6)
Oswestry LBPQ	11.2	(10)	10.5	(11)	0.29	-0.4	(-3.6,2.9)
Requiring further follow-up (%)	40		40				
Requiring discectomy (%)	17		19		0.93		
Days of work absenteeism (median)	65		66		0.84		
Resuming work during follow-up (%)	77		84		0.91		

Note- Columns 2 and 3 show the mean values for the outcome parameters with the standard deviation in brackets. Column 4 shows the p-value (2-sided) for the Mann Whitney statistic for comparison of the mean change in outcome parameters. Column 5 shows the difference in outcome, adjusted for prognostic factors. The bottom part depicts event rates for surgery and resuming work, with p-values for the comparison of survival curves.

weeks were compared. At three months, no interaction was found between previous episodes of sciatica and the intervention type was found at three months.

Table 5. Subgroup analysis

Outcome Measure after 2 weeks	Bed rest %	Watchful waiting %	Difference % (95 % CI)	Adjusted Odds ratio (95% CI)
Effect perceived by patient				
Any improvement				
-no previous sciatica	64	73	-9 (-26,7)	0.6 (0.2-1.3)
-previous sciatica	81	51	29 (8,51)	4.5 (1.4-15.1)
Strong improvement				
-no previous sciatica	41	39	2 (-16,19)	1.1 (0.5-2.4)
-previous sciatica	29	29	0 (-22,22)	1.0 (0.4-3.9)
Effect perceived by observer				
Any improvement				
-no previous sciatica	69	66	3 (-14,20)	1.1 (0.4-2.4)
-previous sciatica	81	63	18 (-4,39)	2.6 (0.8-9.3)
Strong improvement				
-no previous sciatica	20	23	-3 (-18,11)	0.7 (0.3-1.9)
-previous sciatica	13	17	-4 (-21,13)	0.6 (0.1-2.5)

Trial methodology

Patients in the bed rest group compared well with those in the watchful mobilizing group (table 2), apart from slightly less pain in the leg on the Visual Analogue Scale .

The blinded outcome assessor reported unsuccessful blinding in 10 of 183 patients and was correct in 9 of these 10. The outcome assessor became aware of the MRI findings before the first outcome assessment in 3 patients.

The compliance assessment in the patient diaries showed that patients had had a daily average of 21.4 ($\pm$ 3.8, median 22.3) hours of bed rest in the bed rest group versus 10.0 ($\pm$ 3.8, median 8.9) in the control group. The reasons for not complying on a particular day were usually "resolution of complaints" (5 and 4%), "incompatibility with work" (20 and 10%) and "it didn't work" (12 and 9%).

The consumption of analgesics was similar in both groups ($p=0.37$, MWU). Temazepam use lasted, on average, 2 days longer in the control group ($p=0.002$, MWU). Patients in the control group applied heat more frequently than the BR patients ($p=0.0004$, MWU). Other co-interventions were equally distributed with homeopathy being used most frequently (by 10% of patients in both groups).

DISCUSSION

Our results show that the advice to stay in bed for most of the day does not lead to a higher eventual success rate at two or twelve months follow-up visits. Neither the pain

level in leg or low back, nor the degree of functional impairment, the need for surgery or the work absenteeism were significantly lower in the bed rest group. The absence of a statistically significant effect cannot be explained by a lack of power, as the 95 % confidence interval for the effect excludes any clinically relevant treatment effect (defined a priori). Our design contrasted the most frequently used bed rest strategy (according to an inquiry among GPs) with a period of watchful mobilizing. The pragmatic interventions in our study were chosen as realistic and practical reflections of general and neurological practice at the expense of some contrast between the interventions. Theoretically, the compliance with the interventions is irrelevant in intention-to-treat-analysis. Nevertheless, it is informative to know that a substantial contrast remained, with the bed rest patients being out of bed about 2½ hours a day, and the watchful mobilizing patients being out of bed for an average of 14 hours a day.

Bed rest may provide some pain relief as suggested from Greek antiquity onwards⁸. However, it provides no cure for lumbar nerve root involvement and does not alter the natural course. This distinction is shown in figure 1. During the 14 days, pain intensity in the leg may be lower in the bed rest group (though baseline levels were also lower), but the change over time is identical, in other words, the slopes of the lines are equal. Also, as soon as the intervention period ends, the difference between the effect of bed rest and mobilisation disappears.

Figure 1. Visual analogue scale score of pain intensity in the leg over time

Bed rest is prescribed to patients for various presentations of lumbosacral nerve root involvement. Therefore, the trial design was such that the findings might be applied to any patient in a general population with a first time presentation of the current episode of nerve root involvement. However, we do not know whether the primary care physicians did indeed refer all eligible patients for this study instead of applying the usual criteria for referral to the neurological outpatient clinic. Though GPs were asked to register eligible non-referred patients, this registration was not very reliable. It is possible that a higher proportion of patients with a more serious and clear presentation of nerve root involvement have ended up in the intended random sample of a primary care population. In any case, the patients referred but declining participation had baseline characteristics very similar to those participating.

Subgroup analyses are controversial, but, when interpreted with caution, they may provide valuable non-experimental information. It was hypothesized that patients with MRI evidence of nerve root compression would have a different reaction to bed rest compared to those without such MRI evidence. Yet, no significant differences were found and, if anything, the findings suggest that MRI evidence of nerve root compression predicted a trend in favour of watchful waiting. Subgroups were also based on previous episodes of leg pain because they were thought to reflect past experiences of patients and to have influenced patient expectation for the current episode. If patients cannot be blinded and the outcome measure is highly subjective (even if relevant and practical), the outcome reported can differ between such subgroups. The results show that the percentage of patients reporting any improvement after two weeks was significantly higher in the bed rest group if patients had had previous episodes of leg pain. If "previous episodes of leg pain" partly reflects expectation bias as was hypothesized, outcome measures less prone to expectation bias were hypothesized to show less interaction of previous episodes of leg pain with the intervention. The blinded observer's opinion (even if subjected to the patient's unmasked report), as well as defining as a success "strong" rather than "any improvement", and all secondary outcome measures in particular were regarded as such measures. Indeed, the subgroup differences for "previous episodes" subgroups were not significant for all these measures. Probably, patients with previous pain episodes are more familiar with bed rest as it is the standard treatment for sciatica and may expect less favourable results of the watchful waiting strategy. Patients with and without previous episodes of pain in the leg had similar symptom severity and functional status. In view of these considerations, it seems recommendable to incorporate more specific denominators of patient expectation into regression models in future trials.

Our trial results are in keeping with previous low back pain RCTs¹³⁸⁻¹³⁷, which showed that prolonged bed rest is not indicated for low back pain. Though there were no obvious adverse effects of bed rest in our trial, some may occur. Any substantial degree of bed rest may diminish the patient's physical condition¹⁸⁸⁻¹⁹⁶. Moreover, the patient may be strongly conditioned against any activity in the case of pain. Generally, such conditioning is thought to influence the tendency towards chronicity of pain and prolonged work absenteeism^{204, 205}.

Our observer-blinded trial of 14 days of bed rest versus 14 days of watchful waiting in 183 patients has shown that a bed rest regimen does not cure a higher proportion of patients than a watchful waiting regimen. In view of the potentially detrimental effects of bed rest on the patient's physical and psychosocial condition, it is recommended not to subject patients with a lumbosacral radicular syndrome to 14 days of bed rest. Instead, one should let nature run its mainly benign course. The patient may rest in bed but may mobilize when possible. The patients should be instructed and watched regarding the development of paresis or cauda equina syndrome.

Life is full of trials
on somebody's T shirt on a sunny day in spring

ABSTRACT

The prognostic value of the clinical findings elicited in an initially presenting sciatica patient is unknown. A prospective study was designed to investigate whether history and physical examination findings could predict outcome.

A primary care sample of sciatica patients were examined and outcome was registered as improvement reported by the patient after two weeks and after three months. Furthermore, the prediction of Roland disability scale scores, and visual analogue scale scores for pain in the leg and low back at three months was investigated.

In 183 patients, a variety of clinical findings were shown to independently predict outcome. The most consistent predictors of a favourable outcome were more pain on sitting, previous sporting activities, and increased finger-floor distance. The most consistent predictors of an unfavourable outcome were a positive straight leg raising test, a higher Quetelet index, and numbness/subjective sensory loss.

The predictors in this study can indicate the prognosis in sciatica patients at an early stage. This prognosis may help to fine-tune treatment decisions and improve patient selection in trials of conservative therapy strategies.

INTRODUCTION

Neurologists frequently see patients with sciatica in their neurological outpatient clinic¹¹⁴. General practitioners (GPs) with an average practice encounter these patients as often as 6 times per year⁶. The majority will have a favourable natural course^{161 206 207}. A systematic review¹⁸⁷ and subsequent conservative therapy trials^{206 182} have shown that no conservative strategy is able to improve the natural course. Probably, the favourable natural course of sciatica leaves little room for benefit from additional treatment. We wondered whether it might be possible to predict possible unfavourable outcome at an early stage. This might allow conservative therapies not effective for the entire group to be aimed at subgroups likely to have an unfavourable outcome or lead to a preference for early surgery as opposed to prolonged conservative strategies. Such unfavourable outcome may take several forms, such as the patient reporting no improvement, persisting or worsening sciatica, persisting, worsening or ensuing low back pain, persisting functional disability or prolonged work absenteeism.

Little is known about factors that can predict the clinical course in patients presenting for the first time with a new episode of radicular pain radiating into the leg. Nykvist *et al.*²⁰⁸ investigated how clinical findings one year after hospitalisation for sciatica predicted the eventual outcome of rehabilitation. Several authors have reported on factors predictive of the outcome of surgery. Sedentary work (as opposed to physically active work), absence of motor or sensory deficits and pre-operative autotractor seem to be correlated with excellent outcome (defined as the patient's subjective statement)²⁰⁹, and abnormal hysteria and hypochondriasis subscales of the Minnesota Multiphasic Personality Inventory^{210 211} seem to be related to poor outcome of surgery. However, these findings cannot be assumed to apply to patients with sciatica when they first enter the neurologist's or GP's office.

Because the prognostic value of findings elicited from a patient initially presenting with a radicular syndrome is unknown, a prospective study was designed to investigate whether history and physical examination could predict short and long term outcome.

PATIENTS AND METHODS

Study population

In the Netherlands, patients are registered in a particular GP's practice and have their own GP. Referral by the GP is required before a specialist at the hospital can be consulted. An unselected primary care spectrum can be thus be obtained by recruitment of patient directly at the GP's practice. Between February 1995 and December 1996, 50 general practitioners (GPs) in Maastricht and surrounding villages invited patients to participate in a study on sciatica. Patients were referred to the Neurology department of the Maastricht University Hospital if they presented for the first time with the current episode of (1) pain radiating into the leg below the gluteal fold, and (2) a pain intensity sufficient to justify bed rest therapy for 14 days. Exclusion was based on (1) previous spinal surgery, (2) pregnancy, (3) unavailability

plans to move house), and (4) co-morbidity of such a nature that the patients' well-being was primarily determined by it. This study was performed concurrently with a randomized controlled trial of bed rest and a diagnostic study of the value of history and physical examination in patients with pain in the leg. The researchers reviewed the selection criteria and applied two further selection criteria for the trial and the prognostic study. Patients had to have signs and symptoms of sciatica** (at least two of the following signs and symptoms: typically radicular pain distribution, more pain in the leg on coughing/sneezing/straining, decreased muscle strength, sensory loss, reflex loss, positive nerve root irritation sign). They were excluded if there was an indication for immediate surgical intervention, being either (a) intractable pain, (b) rapidly progressive paresis or severe paresis of limited duration or (c) cauda equina syndrome.

Baseline examinations

All subjects selected were examined within two days of referral by the GPs. This examination included a standardized history and physical examination, the technique and inter-observer variability of which have been reported elsewhere²⁰⁶.

Three classes of independent variables were investigated for their prognostic value:

I Patient chart/demographic items: age, gender, educational level, living alone, job type, previous episodes of pain in the leg, previous episodes of low back pain, family history, co-morbidity, smoking, sporting activities, exercised back/abdominal muscles, Quetelet index (weight (kg) divided by squared length (m²)). An additional variable was included indicating membership of bed rest or watchful waiting group in the trial.

II History items: more pain in the leg than in the back, sudden onset of pain, cause of pain known, pain worse in the evening/night, paroxysmal pain, pain already improving, observer's opinion of complaint severity on a visual analogue scale (VAS), typically dermatomal pain distribution, more pain on coughing/sneezing/straining, more pain on sitting, less pain on lying down, less pain on standing/walking, subjective muscle weakness, numbness/subjective sensory loss in the leg, cold sensations in the leg in a dermatomal distribution, paraesthesias in the leg, disturbed urinary passage, urinary incontinence, worrying about health.

III Physical examination items: decreased lordosis, finger-floor distance, paresis, hypesthesia, hypalgesia, ankle tendon reflex difference, knee tendon reflex difference, straight leg raising (SLR) test, reversed SLR, Valleix pressure points (tenderness along the sciatic nerve⁷⁶), Kemp sign (pain on lateroflexion and axial pressure of lumbar spine⁷⁷), Naffziger sign⁷⁶.

Outcome parameters

It was investigated how the parameters of the baseline examination could predict the major improvement over the first two week period as reported by the patient.

** The term sciatica as it is used here refers to a syndrome of lumbosacral nerve root involvement. Although the latter term seems the more accurate, the term sciatica is more widely used.

Furthermore, the following five aspects of outcome at three months follow-up were investigated: substantial (= a combination of the categories reasonable and major) improvement as reported by the patient, work absenteeism, the VAS score for pain in the leg, VAS score for low back pain, and modified Roland Disability Scale score^{200 199 202}. Patients were seen by the outcome assessor at baseline and after 2 and 12 weeks.

Statistical analysis

Firstly, a bivariate analysis was performed (the results of which are only shown for the parameters significant in the multivariate analysis). Dichotomous baseline parameters were compared to dichotomous outcomes by the chi-square for independent proportions. Continuous outcome measures were compared for the two classes by logistic regression analysis. A Pearson correlation coefficient was calculated for the associations between continuous variables. Multiple logistic regression analysis¹⁰⁰ was used to predict major improvement after two weeks, and substantial (= "reasonable" + "major") improvement and work absenteeism after three months. Set I was modelled first, after which set II was modelled, incorporating the predictive variables from the model of set I. The predictors from set I and II were then used with the variables of set III. Presented are all the variables of the final reduced models. Multiple linear regression analysis was used to predict the VAS scores and the modified RDS, with sets 2 and 3 being consecutively added to the model of set 1. The non-parametric Mann Whitney U test was used to analyse differences in baseline characteristics between patient subgroups⁸⁹.

RESULTS

183 patients were selected. Of 338 patients referred by the GPs, 97 did not have sciatica of sufficient severity to require treatment with 14 days of bed rest and 58 were excluded on the basis of our exclusion criteria. The eligible patients declining participation (n=44) had slightly less severe disease (as scored on the VAS by the observer) and a somewhat lower score on the affective dimension of the McGill Pain Questionnaire¹⁹⁸. Two patients did not return after two weeks; both reported improvement. After three months, a total of 14 patients had been lost to follow-up; 9 were unwilling to return, 2 had moved, 2 had incurred serious unrelated illness and 1 had been imprisoned. Their baseline characteristics and success rate after two weeks were similar to the 169 remaining patients. The baseline characteristics of the 183 patients selected are shown in table 1.

The top half of table 2 shows the signs and symptoms that predicted the probability of major improvement over the first two week period. Each variable presented was an independent variable in the logistic model. Factors indicating favourable prognosis were those with an odds ratio greater than 1.0, while factors indicating poor prognosis were those with an odds ratio below 1.0. The odds ratios could also be used to calculate the probability of short term recovery in a particular patient. For example, a patient with pain which is worst at night and which increases upon coughing, sneezing

or straining, with a Quetelet index of 27 and with a finger-floor distance of 30 centimetres could be calculated to have an 8% probability of short term recovery. The bottom half of table 2 shows the variables predicting the likelihood of substantial improvement over three months as reported by the patient.

Table 3 shows that some findings specifically predicted either pain in the leg, low back pain or functional disability at three months. For example, higher pain intensities in the leg were predicted by having co-existent illnesses, a physically demanding job, the absence of more pain on sitting, a higher Quetelet index, and/or the absence of paresis.

Some findings predicted more than one outcome measure. This is illustrated by the SLR test, which predicted pain levels at three months both in the leg and low back. An increased finger-floor distance indicated a favourable prognosis on five outcome measures both in the short and the long term. The Quetelet index was the only other finding to indicate both short and long term (leg pain) outcome. Patients active in sports reported less functional restriction after three months and less back pain.

Table 1. Summary of patient characteristics and outcome (n=183)

	Mean/ Distribution %	(SD)
Patient characteristics		
Age	46	(12)
Male gender	56%	
Duration of pain (median in days)	16	
Any previous episode of pain in the leg	37%	
Any previous episode of low back pain	73%	
Paresis	18%	
Sensory loss	42%	
Reflex differences - ankle tendon	20%	
knee tendon	7%	
Nerve root compression on MR Imaging	59%	
Employed	61%	
Outcome measures		
after 2 weeks -		
major improvement over 2 weeks	67%	
after 3 months -		
substantial improvement over 3 months	83%	
VAS pain in the leg	15	(28)
VAS pain in the back	20	(26)
Roland disability scale	16	(10)
Work absenteeism	32%	

Note- SD= standard deviation

Second column: all values followed by “%” represent percentages of patients with the particular characteristic. Duration of disease is described by the median in days. All other values are means and are followed by the standard deviation in the third column.

Table 2. Predictors of improvement perceived by patient (2 final reduced logistic models)

Predictors of improvement	%	crude odds ratio	p-value	adjusted odds ratio	(95%CI)
Major over two weeks					
Pain worse in the evening/night	53	0.4	0.005	0.4	(0.2-0.9)
Pain already improving	79	2.1	0.03	2.7	(1.2-6.0)
Pain on pressure	64	0.8	0.003	0.3	(0.1-0.5)
Quetelet index	§	0.96	0.28	0.87	(0.78-0.97)
Finger-floor distance					
5-24 cm	73	2.1	0.09	2.3	(0.8-6.6)
25 cm or more	68	1.6	0.24	4.0	(1.4-12)
Hypesthesia	73	1.5	0.22	2.4	(1.1-5.1)
Substantial over 12 weeks					
Duration of disease					
14-30 days	96	1.7	0.55	1.5	(0.3-8.4)
more than 30 days	69	0.2	<0.001	0.1	(0.04-0.4)
Pain worsening on sitting	94	3.6	0.008	5.0	(1.5-17)
Numbness/subjective sensory loss	82	0.4	0.05	0.2	(0.1-0.7)
Finger-floor distance					
5-24 cm	88	3.0	0.03	3.4	(0.9-12)
25 cm or more	95	7.9	<0.001	5.1	(1.3-21)

Note- Complete reduced models are presented. The intercepts were 2.29 (-0.73, 5.32) and 1.82 (0.51, 3.12) for the major and substantial improvement models, respectively.
95%CI= upper and lower limit of the 95 percent confidence interval.
First column: the reference category for duration of disease is 0-13 days and for finger-floor distance 0-4 cm. Second column: percentages of patients with the findings who reported major improvement after two weeks or substantial improvement after three months. Third column: odds ratios for the bivariate association between outcome and predictor. Fourth column: p-values for the bivariate associations. Fifth column: adjusted odds ratios in the final reduced multiple logistic regression model. Sixth column: 95% confidence intervals for the adjusted odds ratio.
§- no percentage shown because the Quetelet index was a continuous measure

Thirty-two percent of patients who had jobs (i.e. 112 patients) had not returned to work after three months. The likelihood of work absenteeism could be predicted by the variables in table 4. According to the model, a 55-year-old employed patient with a positive SLR test and more pain in the leg than the back but without any of the other findings had a likelihood of 90% of still being absent from work after three months. Without positive SLR, the likelihood was 68% and if the patient's age was 30, it would be 40%.

Table 3. Predictors of leg pain, back pain and functional disability after three months: final reduced linear regression models

Predictors for each measure	Sign absent	Sign present	p-value	Coefficient	(95%CI)
VAS of pain in the leg					
Co-morbidity	11	21	0.02	9	(2, 16)
Physically very demanding work	13	24	0.06	10	(0.3, 20)
Duration of disease				6	(2, 10)
14-30 days	11	10	0.76		
more than 30 days	11	27	0.001		
Pain worse on sitting	20	10	0.02	-9	(-16, -2)
Quetelet index	§	§	0.02	0.96	(0.2-1.7)
Finger-floor distance				-8	(-12, -3)
4-24 cm	28	14	0.02		
25 cm or more	28	10	0.001		
Paresis	16	9	0.16	-11	(-20, -2)
Positive straight leg raising test	6	19	0.003	15	(7, 22)
VAS of low back pain					
VAS of low back pain at baseline				0.15	(0.04, 0.3)
Duration of disease				6	(2, 10)
14-30 days	17	15	0.55		
more than 30 days	17	31	0.007		
Sporting activities	24	16	0.06	-8	(-15, -1)
Urinary incontinence	16	43	0.03	22	(4, 41)
Finger-floor distance				-7	(-11, -2)
5-24 cm	32	21	0.10		
25 cm or more	32	15	0.004		
Depressed knee tendon reflex	21	10	0.12	-12	(-25, -0.4)
Positive straight leg raising test	12	24	0.006	14	(7, 22)
Roland Disability Scale					
Sports activities	15	17	0.13	3.1	(0.3-5.9)
Numbness/subjective sensory loss	17	14	0.07	-3.3	(-6, -0.5)
Finger-floor distance				2.5	(0.7, 4.3)
5-24 cm	14	14	0.89		
25 cm or more	14	18	0.004		
Depressed knee tendon reflex	15	23	0.001	8.6	(4, 13)

Note- Complete reduced models are presented. The intercepts were -13 (-35, 9), -17 (-40, 5) and 12 (9, 15), respectively.

VAS=visual analogue scale. §- the association of the two continuous variables was calculated as a Pearson correlation coefficient; $r=0.2$.

For columns 1 and 4 see legend to table 2. Columns 2 and 3 show the mean scores for the group without and the group with the finding, respectively. Column 5 shows the multiple linear regression coefficients in the final reduced models for the variables and column 6 the 95%CI. The coefficient represents the difference in scores on the VAS and Roland Disability Scale (RDS) for patients with versus those without the finding. A negative value for the VAS implies a lower pain intensity if the finding is present. A negative value for the RDS indicates more functional disability.

Table 4. Predictors of work absenteeism at three months (in the final reduced model)

Predictors of work absenteeism	%	odds ratio	p-value	adjusted odds ratio	95%CI
Age					
41-50	29	0.8	0.51	1.5	(0.6, 4.2)
51-81	77	5.0	0.02	13.4	(1.9, 96)
Paroxysmal pain	21	0.4	0.05	0.4	(0.1, 1.0)
More leg than back pain	39	3.3	0.01	3.8	(1.2, 11.9)
Positive straight leg raising test	37	3.6	0.02	4.1	(1.0, 16.5)
Positive Valleix pressure points	69	6.2	0.001	7.0	(1.6, 30)

Note- See legend to table 2; the second column shows the percentages of patient with the finding that have not returned to work after three months.

DISCUSSION

Our study shows that a number of signs and symptoms in the history and physical examination predict different types of outcome in the patients examined. Several methodological comments need to be made, however.

All patients had signs and symptoms indicating lumbosacral nerve root involvement at entry, but a definite gold standard for sciatica was lacking. Neurophysiological investigations are specific but not sensitive enough²¹². Radiological investigations or perioperative observations are relatively accurate in demonstrating nerve root compression^{64 213}, but non-compressive etiologies of root involvement are not identified. Therefore, the accuracy of the diagnosis of nerve root involvement depends on how the diagnostic entity itself is defined. Our definition closely follows that of the Dutch consensus form on the radicular syndrome²²¹. According to this definition, all patients selected had sciatica at presentation. Selection bias could have occurred if the eligible non-selected patients had different characteristics than those selected. Physicians were explicitly instructed not to exclude patients on the basis of history and physical examination findings. Nonetheless, some eligible patients may not have been referred. If they had less severe disease, the percentage of patients with a poor outcome predicted to have a poor outcome may have been overestimated, while the percentage of patients with a favourable outcome predicted to have a favourable outcome may have been underestimated²¹⁴. The eligible patients that were referred but nonetheless refused participation did not show different baseline characteristics.

Statistically significant differences in the VAS scores and the Roland Disability Scale need not imply clinical relevance. For example, patients with a typically dermatomal distribution have a VAS score that is an average of 10 mm. higher - on a

dermatomal distribution have a VAS score that is an average of 10 mm. higher - on a scale from 0 to 100, and with adjustment for other prognostic variables - than patients without such pain distribution. The clinical implications of this difference are unclear. The implications of globally perceived improvement and work absenteeism are, by contrast, apparent.

There may be a serious pitfall in appraising the relation between the test and the outcome. The course represented is not the natural course of the disease. Firstly, the randomly assigned trial interventions (bed rest or a watchful waiting regimen) may have differentially affected the course. The regression analysis corrected for this differential effect. Secondly, patients may have received intercurrent therapies, e.g. physical therapy or surgery) in later stages. If these therapies are related to particular outcomes, tests providing the indications for these therapies invariably predict the outcome. Three of the significant predictors in the models were associated significantly with eventual surgery; the Valleix pressure points, more pain in the leg than in the back and a physically demanding job. Therefore, it could be that the Valleix pressure points and more pain in the leg than in the back actually predict eventual surgery and indirectly prolonged absenteeism due to the surgical procedure. Indeed, the eventual surgery and work absenteeism at three months were found to be associated. A physically demanding job is related to a higher pain intensity in the leg at three months. Maybe, this is because patients with a physically demanding job are less likely to be operated upon eventually. Patients that were eventually operated upon tended to have a lower pain intensity in the leg.

Because the baseline characteristics of the patients lost to follow-up were not clearly different from those who completed the study, little bias will have been due to the 8 percent loss to follow-up.

Some pre-morbid patient characteristics seem to predict the outcome in patients with sciatica. Strikingly, greater physical fitness seems to be associated with a favourable outcome. Being active in sports predicts less work absenteeism and less low back pain, while higher relative weight (Quetelet) predicts a decreased likelihood of short term recovery and more long term leg pain. A physically very demanding job, however, predicts higher leg pain levels after three months.

Several symptoms of the history predict outcome. Two symptoms stand out: a poor outcome is more likely when the patient complains of numbness/subjective sensory loss, and a favourable outcome is more likely with pain that worsens on sitting.

The straight leg raising test is an important predictor of unfavourable prognosis. Roland *et al.*²¹⁵ reported the same for low back pain patients. In our study, leg pain and low back pain at three months tend to be higher in patients who had a positive SLR. The odds ratio for work absenteeism was 4.1. In a related diagnostic study¹³³, the odds ratio of the SLR for nerve root compression was 2.0. It is remarkable that the SLR test is a more accurate sign of future work absenteeism than of nerve root compression. The value of the Valleix pressure points is illustrated by the increased likelihood of work absenteeism when the Valleix test is positive in addition to a positive SLR; in a 45-year-old patient with paroxysmal pain predominantly in the leg, the likelihood would increase from 26 to 71%.

The most universal predictor was increased finger-floor distance. It predicted a more favourable outcome both in the short and in the long term (all outcome modalities except work absenteeism). This finding has been shown to be one of the most accurate signs of sciatica due to disc herniation¹³³.

Most conservative interventions are ineffective for the general group of patients with sciatica¹⁸⁷. They may, however, have an effect in a particular subgroup, e.g. groups with a particular prognosis. The predictors of this study indicate the prognosis in sciatica patients at an early stage, which may help to fine-tune treatment decisions. For example, the decision to prescribe a more extensive physical therapy program might take into account the likelihood of a future unfavourable outcome. In addition, it may improve patient selection for trials of conservative therapy strategies.

INTRODUCTION

While several papers support the efficacy of lumbosacral discectomy in patients with sciatica due to disc herniation^{216 70 176}, it is indicated in a subgroup of patients with sciatica, only. The first prerequisite for operation seems to be the presence of nerve root compression supported by both the clinical picture and radiological images. The nineteenth century brought the description of the clinical picture of sciatica^{10 13} and Mixter and Barr¹¹⁶ described the relation of sciatica to disc herniation. Since then, the clinical picture of sciatica due to disc herniation has been described in various expert opinion articles^{18 1}, cross-sectional diagnostic studies¹³³, and systematic reviews^{94 95}. Many papers have addressed the radiological signs of nerve root compression^{72 217 64 218 105 96 54}.

The problem related to this first requirement for operation is how high the likelihood of nerve root compression should be to warrant operation²¹⁹. The optimal likelihood depends on the situation and the patients' and doctors' preferences. Moreover, the fact that even patients with near certain nerve root compression are not automatically operated, points towards additional requirements. Such a (largely subjective) requirement is that the intensity or bothersomeness of complaints should be sufficient for both patient and the treating physician to consider surgery.

The final and maybe most troublesome requirement is that the complaints of nerve root compression should have existed for a sufficient time period before surgery is considered. In view of the favourable natural course of sciatica due to disc herniation^{161 187 207 220 206}, and the possible complications of surgery, most physicians will probably prefer to await the spontaneous outcome. The Dutch consensus on the lumbosacral radicular syndrome has advised a period of 6-8 weeks before ordering additional radiological procedures and considering eventual operation²²¹. The optimal duration of waiting is unknown, however. Is there a way of predicting an eventual poor outcome of conservative therapy with the subsequent need for surgery, when the patient initially presents?

To answer this question, a prospective study was designed. Patients with signs and symptoms of sciatica were investigated with regard to findings that might predict the future need for operation. This prediction was evaluated by logistic regression at two stages: (1) how could future surgery be predicted at the first consultation (model A)?, and (2) how could it be predicted at a second consultation after two weeks taking into account the disease course over two weeks (model B)? The benefit of these predictions might be in expediting MRI studies and eventual surgery. We assumed that a model used to expediate MRI studies should have a specificity exceeding 95%; no more than 5% of the patients treated conservatively would then be predicted to eventually require surgery. If the model were used to expediate surgery, a cut-off probability should be chosen with a 100% specificity. In this way, none of the patients with nerve root compression on MRI and with a favourable outcome of conservative treatment would have been predicted as eventually requiring surgery. Based on these assumptions, the value of the predictive models was estimated. These estimations would indicate if it would be worthwhile to validate the models in additional research.

PATIENTS AND METHODS

Study population

Between February 1995 and December 1996, 50 general practitioners (GPs) in Maastricht and surrounding villages asked patients to participate in a study on sciatica. Patients were referred to the Neurology department of the Maastricht University Hospital if they presented for the first time with the current episode of (1) pain radiating into the leg below the gluteal fold, and (2) pain intensity sufficient to justify bed rest therapy for 14 days. Exclusion was based on (1) previous spinal surgery, (2) pregnancy, (3) penitentiary confinement, (4) unavailability for follow-up visits (moving plans, severe disease), and (5) co-morbidity of such a nature that the patients' well-being was primarily determined by it. This study was done concurrently with a randomized controlled trial of bed rest²⁰⁶ and a prognostic and diagnostic study³³ of the value of history and physical examination in patients with pain radiating into the leg. The researchers reascertained the selection criteria and applied two further selection criteria for the trial and the prognostic study. Patients had to have signs and symptoms of lumbar nerve root involvement (at least two of the following signs and symptoms: typically radicular pain distribution, pain in the leg increasing with coughing/sneezing/straining, decreased muscle strength, sensory loss, reflex loss, positive nerve tension test). They were excluded if there was an indication for direct surgical intervention, being either (a) intractable pain, (b) progressive, severe decreased muscle strength of limited duration or (c) cauda equina syndrome.

Baseline examinations

All patients selected were examined within two days of referral by the GPs and again two weeks later. The technique and inter-observer variability of history and physical examination are reported elsewhere¹¹³. At stage I, two classes of independent variables were investigated for their predictive value:

I History/questionnaire items: age, gender, high education level, living alone, mentally demanding job type, previous episodes of pain in the leg, previous episodes of low back pain, family history, co-morbidity, smoking, sports activities, exercised back/abdominal muscles, more pain in leg than back, sudden onset of pain, cause of pain known, pain worse in the evening/night, paroxysmal pain, pain already improving, typically dermatomal pain distribution, more pain on coughing, sneezing or straining, more pain on sitting, numbness in the leg, paraesthesias in the leg, cold feeling in the leg, worrying about health, intensity of pain in the leg scored by the patient on a visual analogue scale (VAS), complaints considered severe by the physician, Pain Rating Index of McGill Pain Questionnaire¹⁹⁸ for sensory and affective dimensions (PRI-S and PRI-A), questionnaire items difficulty putting on socks/stockings, and pain while dressing.

II Physical examination items: finger-floor distance, paresis, hypesthesia, hypalgesia, ankle tendon reflex difference, knee tendon reflex difference, straight leg raising (=SLR, the test was positive if there was typically dermatomal pain upon straight leg raising at an angle less than 45°), reversed SLR, crossed straight leg

raising, Kemp sign, Naffziger sign.

To take the time course into account, a second stage analysis incorporated results of the two week follow-up visit (instead of baseline values) for typically dermatomal pain distribution, pain worse on sitting, SLR, and crossed SLR. Also, for pain intensity in the leg the change of VAS over two weeks was assessed instead of the VAS at baseline. The PRI-A of the McGill Pain Questionnaire after two weeks was investigated instead of the PRI-A at baseline.

Statistical analysis

Backward stepwise logistic regression analysis¹⁰⁰ with the likelihood ratio test was used to predict surgical intervention in the following 6 months. At the first stage, set I was modeled. Then, set II was modeled also incorporating the predictive variables from the model of set II. The resulting reduced models were based on data collected at first consultation and predicted the likelihood of surgery in the next 6 months. The probability of surgery for each patient can be estimated using the formula below table 2. At any given probability cut-off, a number of patients will be below the cut-off and a number above. These numbers can be compared in a two by two table with the numbers actually operated. In this way, sensitivity and specificity values can be calculated for any probability cut-off value. The (1- specificity) versus sensitivity values for the various probability cut-offs of both models are depicted in the Receiver Operating Characteristic (ROC)^{101 102} in figure 1. The area-under-the-curve was estimated non-parametrically. This area reflects the overall predictive properties of each model. At the second stage, variables investigated at the two weeks consultation were also taken into account. The predictive properties of this model were also investigated using a ROC curve (figure 1). The probabilities at which the specificity would be 95% and 100% (as motivated in the introduction) were determined. To illustrate the possible consequences of applying the models using these probabilities, an example was worked out in which patients would indeed receive the predicted surgery at an early stage (see figure 2 and table 4).

RESULTS

29 of 212 eligible patients declined participation. They had slightly less severe disease (as scored on the VAS by the observer) and a somewhat lower score on the affective dimension of the McGill Pain Questionnaire¹⁹⁸. 183 patients were eventually included. Two patients did not return after two weeks; both reported improvement. After three months a total of 14 patients were lost to follow-up. Their baseline characteristics and success rate after 2 weeks were similar to the 169 remaining patients. 22 patients did not receive MR scanning because of a metal-containing corpus alienum, claustrophobia, or unavailability of MRI within 24 hours (a criterium

Table 1. Baseline patient characteristics (n=183)

Patient characteristics	Mean/ Distribution %	(SD)
Age	46	(12)
Male gender	56%	
Duration of pain (median in days)	16	
14 days or less	48%	
15-30 days	26%	
more than 30 days	26%	
Any previous episode of pain in the leg	37%	
Any previous episode of low back pain	73%	
Visual Analogue Scale (mm)		
Major complaint	85	(17)
pain in the leg	65	(21)
low back pain	47	(32)
McGill Pain Questionnaire		
PRI-A	3.7	(3.1)
PRI-S	10.0	(5.8)
Paresis	18%	
Sensory loss	42%	
Reflex differences -		
ankle tendon	20%	
knee tendon	7%	
Definite clinical diagnosis	85%	
Nerve root compression on MR Imaging	59%	
Co-morbidity	34%	
Sports activities	44%	
Employed	61%	
Receiving worker's compensation	8%	
Requiring surgery in following 6 months (n=30)	16%	

for the overall diagnostic study). Their baseline characteristics were not different from those that underwent MRI studies. The characteristics of the 183 patients selected are shown in table 1. Thirty patients eventually had surgery.

Overall, the eventual need for surgical intervention could be accurately predicted. Table 2 shows 6 baseline variables that independently predicted eventual surgery: a mentally demanding job, sudden onset of pain, more pain on coughing, sneezing or straining, difficulty putting on socks/stockings, VAS of pain intensity in the leg, and the reversed straight leg raising test. Together with the intercept they formed the reduced logistic model A.

Table 3 shows model B based on both baseline data and the data after two weeks as well as the change over time. The predictors were: duration of disease, previous episodes of low back pain or pain in the leg, a positive family history for low back pain or pain in the leg, a mentally demanding job, paraesthesias, a pain worsening on sitting after 2 weeks, the pain intensity in the leg on a VAS, PRI-A of McGill Pain Questionnaire, and the crossed SLR.

Table 2.   **Reduced model A: the variables at first consultation predicting surgery**
(multiple logistic regression analysis with eventual surgery as dependent variable)

Baseline predictors of surgery	crude odds ratio	95%CI	adjusted odds ratio	95%CI
History/questionnaire				
Mentally demanding job	3.8	1.7-8.6	4.0	1.5-10.3
Sudden onset	0.28	0.1-0.7	0.22	0.07-0.68
More pain on coughing/sneezing/straining	5.3	2.0-13.6	4.3	1.5-12.4
Difficulty putting on socks/stockings	1.6	0.7-1.3	2.3	0.86-6.0
VAS of pain intensity in the leg	1.03	1.0-1.05	1.033	1.004-1.062
Physical examination				
Reversed straight leg raising	2.5	1.1-5.8	3.2	1.2-8.9
exponential value of the intercept	3.6*10 ⁻³	2.9*10 ⁻⁴ -4.4*10 ⁻²		

Note- 95%CI: limits of the 95% confidence interval
VAS: Visual analogue scale
The probability of surgery in the next 6 months can be calculated as:
Probability (X)= 1 / (1 + e ^{-Y})
Y= α + ln(4)*X1 + ln(0.22)*X2 + ln(4.3)*X3 + ln(2.3)*X4 + ln(1.033)*X5 + ln(3.2)*X6
α=intercept and X1 through X6 are the six variables in the order of the table.

The predictive properties are further summarized by the Receiver Operating Characteristic (ROC) curves in figure 1. Overall, model B provided a better prediction of surgery than model A. The area-under-the-curve was 0.86 for model A and 0.93 for B. Figure 1 also shows that model A might be preferred , when a very sensitive (few false-negatives, upper right corner of ROC) prediction is called for, and model B, when a high specificity (few false positives, lower left corner) is the purpose.

DISCUSSION

Our study shows that it is possible to predict eventual surgery in sciatica patients. Already at first consultation, several characteristics of patients with sciatica predict the eventual need for surgery. A mentally demanding job (as opposed to a mentally non-demanding job or no job), a gradual onset of pain, an increase of pain on coughing, sneezing or straining and difficulty putting on socks/stockings are the four history items predicting surgery. Of the physical examination findings, solely a reversed straight leg raising test (SLR) predicts surgery. Of these findings, only an increase of pain on pressure and the reversed SLR are generally considered to be findings of nerve

Table 3.   **Reduced model B: the variables at second consultation predicting surgery**
 (multiple logistic regression analysis with eventual surgery as dependent variable)

Baseline predictors of surgery	crude odds ratio	95%CI	adjusted odds ratio	95% CI
History/questionnaire				
Duration of disease at presentation				
15-30 days	3.4	1.3-8.7	5.8	2.7-12.5
more than 30 days	1.8	0.7-5.1	1.2	0.6-2.7
Previous episodes of pain in the leg	1.0	0.4-2.3	0.17	0.08-0.3
Previous episodes of low back pain	2.8	0.9-2.1	7.3	2.8-19.2
Positive family history for				
pain in the back	3.3	0.9-12.4	10.1	3.2-31.2
pain in the leg	1.0	0.4-0.9	4.1	1.9-8.7
Mentally demanding job	3.8	1.7-8.6	4.4	2.3-8.2
Paraesthesias	0.5	0.2-1.2	0.2	0.08-0.36
Pain worsening on sitting after 2 weeks	4.4	0.6-2.3	24.3	9.6-62
Change over 2 weeks in leg pain VAS	1.05	1.02-1.07	1.058	1.040-1.076
PRI-A of McGill Pain Questionnaire	1.3	1.1-1.4	1.35	1.22-1.49
Physical examination				
Crossed straight leg raising	2.1	0.8-5.7	4.2	1.8-10.2
exponential of the intercept	2.3×10^{-3}	5.3×10^{-4} - 1.1×10^{-2}		

root compression.

The model based on first and second consultation allowed an even better prediction of surgery; the area-under-the-curve was 0.93 for model B versus 0.86 for model A. Strikingly, only three of the predictor in the model are generally regarded as indicators of nerve root compression predicted surgery; paraesthesias, pain increasing on sitting and a crossed SLR. Another remarkable observation is that model B contains other baseline variables than model A. Some variables in model A may predict eventual surgery partly because they predict poor outcome after two weeks. In a model with outcome measures after two weeks, such variables will lose part of their predictive value. Reversely, the duration of disease contributes to the model only if the outcome after two weeks is accounted for. Meanwhile, it reflects that for most patients about 4-6 weeks of complaints will have passed by before surgery is considered. Interestingly, patients with previous episodes of back pain are more likely to undergo surgery, while patients with previous episodes of leg pain are less likely. The latter group's experience with previous sciatica episodes that have resolved without surgery may decrease their inclination towards surgery. On the contrary, patients that have had low back pain before, may be more disconcerted by the new experience of radiating pain

into the leg. Patients in mentally demanding professions may be more likely to undergo surgery because of a more assertive attitude towards surgeons, or because surgeons may be more willing to operate patients working in physically demanding jobs.

Figure 1. Receiver Operating Characteristic for the prediction of eventual surgery by model A + B

In this respect, it is also striking that patients with a higher score on the affective dimension of the McGill pain questionnaire are more likely to undergo surgery. Finally, the change over time of pain intensity in the leg is a strong predictor. It seems logical that pain level after two weeks in comparison to the baseline levels are related to the likelihood of surgery.

After having stated that surgery can be predicted, an important question is how good the predictive properties of the models are. According to the ROC curves and their corresponding areas-under-the-curve, the overall predictive value is high. The following paragraph will illustrate the potential clinical relevance of the model.

Probability cut-offs can be chosen to suit the particular demand. If patients are very likely to eventually undergo operation, it might be desirable not to wait until this indeed happens, but to operate at an earlier time. If the latter decision were based on the model, the probability cut-off would have to have a very high specificity. Every

false positive prediction would imply a patient operated who would otherwise not have been operated. A specificity of 100% was possible for model A, but with a sensitivity of 7%, only. Hypothetically, model A might then reliably expediate surgery, but only in 7% of the patients that eventually received surgery. For model B, the probability cut-off with a 100% specificity was 0.60. The sensitivity of this probability value was 57%. This means that surgery might be expediated in 17 of 30 patients eventually operated (see figure 2). This seems a relevant gain.

In table 4, the estimated gain in waiting time is calculated. Per 30 patients operated, 100 weeks of waiting time could be gained if the models were applied. With an employment rate of 60%, estimated indirect costs of 500 dollars per week related to waiting, and an annual operation rate of 20000 in the US, a gain of 3.3 weeks per patient would amount to saving 20 million dollars of indirect costs per year in the US alone.

There are several reasons why the models cannot be directly applied, however. Firstly, determining the diagnostic properties of a prediction rule in the population from which it was derived will likely overestimate the predictive value. Such a prediction rule should first be validated in another study population. Secondly, the prediction rule was formulated for a population directly derived from primary care. It remains to be shown whether the characteristics of the prediction rule can be extrapolated to more selected secondary care populations²¹⁴. Finally, the prediction

Figure 2. An example of model B application

Table 4. Time elapsed between consultation at two weeks and eventual operation

	Usual indications for surgery	Usual indications for surgery + Model B
Total number operated n=30		
Mean time (weeks):	7.2	3.9
Median (weeks):	6.5	1
Sum (weeks):	217	117
Total number predicted as eventual surgical candidates n=17		
Mean time (weeks):	5	1
Median (weeks):	6.9	1
Sum (weeks):	117	17

Note- The following are the underlying assumptions:

At a probability cut-off of 0.60, Model B predicts eventual surgery in 17 patients. There are no false positive predictions (indeed, all patients were operated). If the positive prediction had resulted in operation in week 3 instead of further awaiting the natural course, the total waiting time for these 17 would have been 17 instead of 117 weeks.

rule is dependent on the indications for surgery currently applied in our hospital. If other indications for surgery are used, the predictive rule may no longer hold. The models only show that given a set of indications (namely those in our hospital) eventual surgery in sciatica can be predicted. Importantly, most of the predictors do not seem to be obvious indications for surgery. The nature of the predictors in the model and particularly the striking lack of classic indicators of nerve root compression suggest that many non-medical factors influence the decision to operate. It is not unlikely that the latter is a universal situation.

General discussion

In chapter one, the research goals for this thesis were described. The following eight chapters featured the background for each goal, the methods of the research conducted, the results and a discussion of the research. The present chapter offers a short synthesis of the material of the previous chapters and a general discussion. In the first section the history and physical examination in patients suspected of sciatica is highlighted and in the second section the conservative treatment of sciatica.

The history and physical examination in the patient suspected of sciatica

The essential question of the first part of this thesis was what the value of history and physical examination is in the diagnosis of sciatica due to disc herniation? The research goal was aimed at sciatica due to disc herniation in particular, and not at sciatica *per se*. The following paragraph implicitly motivates this choice.

When and why should a set of signs and symptoms be labelled sciatica?

Sciatica is the clinical syndrome of lumbosacral nerve root involvement and not a diagnosis in the true sense. Hence, the gold standard for such a syndrome may be any fair agreement between physicians as to which signs and symptoms are supposed to indicate nerve root involvement. The Dutch consensus statement regarding sciatica is such an agreement²²¹. Most definitions of the sciatic syndrome include references to a dermatomal distribution of the leg pain, an increase of the leg pain on coughing, sneezing or straining, paraesthesias in the leg, urinary incontinence, positive nerve root tension signs (SLR, Bragard, Kemp, Naffziger, etc.), antalgic posture, and neurological deficits such as absence of or difference between ankle or knee tendon reflex, decreased muscle strength and/or sensory loss.

Does the physician's conclusion as to whether sciatica is present or not have any consequences? The absence of sciatica may be a comfort to the patient. Low back pain syndromes without nerve root involvement have been reported to have a better prognosis than those with root involvement²²². Not finding sciatica may also be the

incentive to search out another diagnosis. There is no evidence that an initial positive diagnosis of sciatica should have any consequences. The results of the therapeutic review and the trial seem to suggest there should be no definite therapeutic consequences. The question whether particular causes of sciatica should be ruled out at an early stage cannot be completely resolved. Diabetes²²³ and neuroborreliosis are examples of non-mechanical causes of sciatica²²⁴. The prevalence of such causes in a population with irradiating pain into the leg is probably much lower than that of sciatica due to disc herniation. Also, the diabetes patient with sciatica may still have a disc herniation as the most likely cause. Therefore, it seems that the ruling out of non mechanical causes should be reserved for atypical cases, for cases not responding to conservative therapy, or for cases with a high a priori likelihood of non mechanical causes (e.g. cases with a history of tick bite and erythema migrans).

When a patient with pain irradiating into the leg presents, the mind, both of the doctor and the patient, seems to be occupied most by the question whether a disc herniation may explain the complaints? The diagnosis of sciatica due to disc herniation does seem to have consequences. Eventually, radiological verification of a disc herniation may be required and in a minority surgical intervention. Ideally, radiological verification should only be undertaken when surgery is considered^{106 221}. The history and physical examination are the first means of establishing whether there is indeed sciatica due to a disc herniation. The first part of this thesis essentially revolved around the question:

How consistent and accurate is the diagnosis of sciatica due to disc herniation based on the history and physical examination?

The systematic review of the diagnostic literature in chapter two showed that previous studies of history and physical examination in sciatica mostly were reports on hospitalized patients for whom clinical findings and a radiological or peroperative diagnosis was available. Without exception the radiological or peroperative diagnosis was one of disc herniation instead of nerve root compression. The latter is the relevant finding. There is ample evidence that disc herniation can occur without nerve root involvement^{96 225 110 112 141}. Also, all studies suffered from a fair deal of verification bias. This bias occurred in two stages; in all studies patients had already been referred to the setting in which the investigation took place, to some extent on the basis of positive clinical findings (also called referral bias). Then, having arrived in the secondary care setting, most studies verified the diagnosis by radiological investigation or operation preferentially in patients with positive findings. As a result the positive clinical findings had a higher chance of being referred against a gold standard than negative clinical findings. Concurrently, the spectrum of disease complaints of the patients finally investigated tended to be more severe than that of the patient initially presenting with the complaints and in whom history and physical examination are the first means of attaining a diagnosis. Strikingly, the straight leg raising (SLR) test and the crossed SLR were predominantly investigated, while most of the questions that are regularly asked of the patient with irradiating pain into the leg had not been investigated at all. To minimize the bias in the literature search, we tried

to standardize the searches as much as possible (for the therapeutic review we used the search strategy published by Dickersin *et al.*). Reviewers bias (i.o.w. being influenced by one's own preoccupations regarding particular therapies, authors or journals, when assessing the methodological quality of the papers) may have affected the results. Still, blinding of papers was not used. We were unaware of a report on the effectiveness of different types of blinding procedures, especially when the reviewers are very familiar with the original papers. Meanwhile, it is a laborious task. The perpetual question with the methodological assessment in systematic reviews deals with the criteria applied. Many varieties of criteria ckecklists exist. Some intend to judge study quality, others paper quality and still others judge something of both. Globally, the reader should be convinced that the major methodological requirements of studies were checked without any important omissions. We feel confident the checklist we used meets this standard. Of course, it would be desirable to have a checklist that is broadly accepted. Fortunately, an admirable effort to develop such a criteria list for reviewing clinical trials has meanwhile been made by Ariane Verhaege.

The cross-sectional study was designed to minimize the above-mentioned verification bias. The selection of patients for the study was not based on the presence of any clinical finding apart from a sufficiently severe irradiating pain into the leg. It cannot be excluded, however, that the general practitioners applied some degree of additional selection before referring the patient. All eligible patients were supposed to undergo clinical and Magnetic Resonance Imaging (MRI) investigation. However, in 22 patients, MRI scans could not be made as a result of contra-indications or unavailability of MRI. Their baseline characteristics were comparable to those of the 274 selected patients.

An important aspect of any diagnostic study is the choice of gold standard. It seems that MRI is the preferred procedure for the anatomical diagnosis of nerve root compression. In my view, it is questionable whether operation allows a more complete observation of the local anatomy and, more importantly, it is unfeasible in patients without clinical findings suggesting sciatica. MRI may lead both to false positive and false negative diagnoses of nerve root compression, however. Unfortunately, a previous investigation of MRI scans in asymptomatic subjects has addressed the presence of disc herniation and not nerve root compression⁹⁶. In our study, 12% of the patients had nerve root compression on MRI on the asymptomatic side. This suggests that false positives may indeed occur with regard to nerve root compression. Still, it can only be speculated how frequently nerve root compression on MRI on the side of complaints is not the cause of the complaints. False negatives are probably less likely, but also have less drastic consequences (where false positive MRIs may lead to unnecessary operations). The great variance of clinical findings over time seems to suggest that the involvement of the nerve root by a disc herniation is a dynamic situation. These dynamics may explain why some of the veritable nerve root compressions may be missed by MRI. Also, the non mechanical effects of the herniated disc material on the nerve root may lead to nerve root pathology concurrently with an apparent but irrelevant mechanical effect. It is as yet unclear how this inflammatory component should be managed. An important final remark is that

from a pragmatic viewpoint the fallacies of MRI as a gold standard are of secondary importance. Currently, clinical findings of sciatica due to disc herniation of sufficient severity and duration to warrant surgery are an important requirement for surgery in addition to MRI evidence of nerve root compression^{226 221}. Therefore, the pragmatic requirement for the clinical findings is to diagnose nerve root compression as much in accordance with MRI as possible.

Literature described the value of tests for nerve root compression in terms of their direct relation to the gold standard. When multiple tests are performed, as is generally the case in history and physical examination, test results are correlated, however. It is incorrect to consider them as a set of tests, for which the individual diagnostic value expressed for example as likelihood ratios, is additive. It seems appropriate to consider all components of the history as a battery of tests, of which the information is processed simultaneously (in parallel). Then, the physical examination constitutes another battery of tests, interpreted in parallel. Most commonly, the physical examination battery follows the history battery (serial). These assumptions are reflected in the statistical analysis of the diagnostic study.

The overall results of the cross-sectional study indicate that the history provides most of the diagnostic information on nerve root compression. The Receiver Operating Characteristic (chapter 4, figure 2) shows that the physical examination adds very little to the prediction of nerve root compression on MRI. The likelihood of having nerve root compression increases significantly with age, a paroxysmal onset, with pain more in the leg than in the back, with a typically dermatomal distribution of the pain, and with more pain in the leg on coughing, sneezing or straining. A paresis, a finger-floor distance of more than 25 cm, or alternatively, a decreased lumbosacral lordosis are additional features of nerve root compression. Apparently, the SLR is significantly correlated with nerve root compression, but adds less to the diagnostic information already acquired than for example an increased finger-floor distance. Maybe, much of the information in the SLR is already conveyed by a typically dermatomal distribution of pain or a pain increase on coughing, sneezing, or straining. The SLR technique seemingly most widely used was adopted in this research: a provocation of typically dermatomal pain on raising the straight leg was labelled as "positive". Additionally, the potential value of alternative SLR testing techniques was explored. It was found that incorporating an angle restricted below 40 degrees as a condition for positivity would increase the discriminative power of the SLR, also in a multivariate analysis.

The humbling observation in chapter three was that even the standardized history in our investigation had many inconsistent components. And though physical examination findings led to a more consistent diagnosis, there was considerable disagreement between the two observers, even with regard to the revered neurological examination. Some inconsistencies were especially striking. There was little agreement on what constitutes a typically dermatomal distribution. This makes the discriminative properties of this symptom even more telling. Subjective sensory loss and a subjective decrease of muscle strength are not only poorly reproducible, they also indicate (if anything) absence of nerve root compression. This is the reverse of what is generally presumed. Fortunately, the overall conclusion of both observers

ABSTRACT

The duration of complaints is an important factor in the decision to operate upon the sciatica patient, apart from the presence of nerve root compression -both clinically and radiologically-, and the bothersomeness of complaints. Should a given time have elapsed before considering a discectomy, or can patients with a very high likelihood of eventually undergoing surgery be recognized at an early time?

In a prospective study, 183 sciatica patients sampled from a primary care population underwent a standardized history, physical examination and Magnetic Resonance Imaging. Surgery carried out in the next 6 months was recorded. Models were constructed to predict whether patients eventually received surgery.

Two models were constructed. Reduced model A was based on baseline findings, only, and contained 6 variables. Model B incorporated change over time as well and contained 10 variables. The area-under-the-curve (in a Receiver Operating Characteristic) for these models was 0.86 and 0.92, respectively. It was shown that at a probability cut-off of 0.60, model B predicted eventual surgery with a sensitivity of 57% and a specificity of 100%.

Given the requirement of a high specificity, eventual operation could be adequately predicted in a sample of 183 sciatica patients. The application of the model in sciatica patients might lead to earlier operation in a subset of patients resulting in a reduction of duration of illness and associated indirect costs. An important prerequisite for future application would be the validation of the prediction rule in another population.

appeared fairly consistent.

How then to perform an accurate, consistent and efficient clinical investigation of the patient with irradiating pain into the leg. Overall, it might be kept in mind that with increasing age patients are more likely to suffer from nerve root compression. A higher likelihood of having nerve root compression on MRI is furthermore indicated by having paroxysmal pain, pain that worsens on coughing, sneezing and straining and pain that irradiates according to a typically radicular distribution. Because of the inconsistency of finding a typically dermatomal distribution of the pain, it may be preferred to describe the site of pain in anatomical landmarks instead of dermatomes. An essential physical examination would include having the patient attempt to touch the floor with the fingers of both hands with the legs fully extended, investigating his muscle strength and performing the straight leg raising test paying particular attention to a restriction of the angle below 40 degrees.

Is there a distinct relation between the clinical picture on the one hand and the level of root involvement and the location of disc herniation and nerve root compression along the course of the root on the other hand?

The study reported in chapter 5 was done to explore whether the standardized investigation of clinical and MRI findings in an unselected patient population might elucidate the anatomy underlying a number of clinical findings. Of course, the design permitted no conclusions regarding causal relations. It was complicated by the correlation between site of disc herniation, site of nerve root compression and level of nerve root compression. The *a priori* hypothesis was that particular sites of disc herniation and root compression would be positively related to several clinical findings. This hypothesis was, however, not further specified. Hence, the purpose of the study was explicitly explorative. The results confirmed that the probability of exhibiting a paresis or reflex loss is dependent on whether the L3, L4, L5, or S1 nerve root is afflicted. It also supported the *in vitro* anatomical observations of differing length and angulation of the nerve root sleeve and differing distances to the intervertebral disc per nerve root level. Moreover, it was shown that -irrespective of root level- a positive Lasègue (=SLR) occurs relatively more frequently with compression in the pit of the nerve root sleeve and an increased finger-floor distance with mediolateral symptomatic disc herniations.

The conservative treatment of sciatica

The most important question addressed in the second part of this thesis was: **How should sciatica be treated initially?**

Several investigators have made attempts to gather empirical evidence in support of particular therapies. In chapter 6 these attempts are summarized. Reports on

randomized placebo-controlled trials (RCTs) (in english, french, german or dutch and with a minimum of 20 patients per treatment group) of treatment with non-steroidal anti-inflammatory drugs, traction, epidural steroids, systemic steroids, manipulation, chymoral and tizanidine were found. Three epidural steroid RCTs were selected and had a positive effect after pooling. Manipulation was the only other treatment to show a positive effect (in one placebo-controlled trial, only). Bed rest had not been the subject of a RCT.

The major difficulty in performing the meta-analysis was the great clinical heterogeneity of the RCTs and the small sample sizes. Most trials had such small sample sizes that the clinically relevant differences might have reached statistical significance if the sample size had been larger. The other difficulty with small sample sizes is that small studies are more likely to have insufficient comparability of the intervention groups. This relatively great heterogeneity leads to additional noise when statistically summarizing qualitative outcome measures. The effect of this “noise” is not accounted for by so called tests-of-homogeneity as these only address statistical “noise” (and may well be false-negative!). For this reason, it was decided to exclude studies with less than 20 patients per intervention group. This most affected the number of RCTs on epidural steroids; 6 were excluded because of too small sample sizes. In chapter 6, figure 1 shows that this exclusion did not affect the conclusions regarding epidural steroids significantly. The pooled data were not powerful enough to allow definite statements on the effect of several study characteristics on the effect differences between the groups. Trends were detected both for trials studying acute disease and for trials offering a good differentiation of radicular from non-radicular pathology to reveal smaller therapy effects. Considering the benefit of epidural steroids in our data and the only very small benefit reported meanwhile for a much larger trial, it cannot be definitely excluded that epidural steroids are beneficial in sciatica patients. A new RCT, particularly in acute sciatica patients, seems required. As the positive statement on manual therapy is derived from only one trial with questionable methodology, this statement requires further support before being generally accepted.

The data on the natural course of sciatica were also summarized in a graphical presentation of the course. The procedure was basically very different from the other pooling procedures. The latter procedures pooled empirical data of RCTs and even though bias may be operative in the selection and subsequent handling of the data, the meta-analysis profits from the empirical design of the separate trials. On the contrary, the representation of the natural course was essentially a summary of separate cohorts (a meta-analysis not of trials but of the placebo cohorts in the trials). Although the bias in this graphical summary may be considerable, it is probably a robust estimate because both recovery rate and bias are averaged.

Does a period of 14 days of bed rest improve the spontaneous course of sciatica?

In a randomized controlled trial, a period of bed rest was compared with a regimen of watchful waiting, both for a duration of 14 days. There was no significant effect in favour of bed rest for any of the outcome measures.

The study population was not very homogeneous. The exclusion criteria pregnancy, previous spinal surgery and running Worker's Compensation claims served to exclude small subsets of patients that might have a disease course different from the average sciatica patient. Essentially, all other variants of sciatica, for which bed rest might be considered as a therapy, were included. As a result the trial may have lost some homogeneity, but definitely gained in external validity (generalizability of the results). No prognostic factors that differentially affect the response to bed rest or watchful waiting have been reported in the literature. Therefore, we decided there were no arguments to exclude other subtypes of sciatica. For the same reason, we did not prestratify on particular baseline characteristics. The differences in baseline characteristics observed between the groups were not of great significance. This is supported by the multivariate analysis. None of the outcome measures were significantly affected when any of the somewhat unequally divided characteristics were incorporated as confounders. It was hypothesized a priori that the response to bed rest might be different for patients with acute versus chronic complaints, for patients with a first time episode versus a recurrent episode, for patients with versus those without nerve root compression on MRI, and finally for patients with a paresis versus patients without a paresis. The hypotheses were tested by incorporating the variables as effect modifiers in the multivariate logistic regression analysis. In conclusion, the bed rest advice is not effective regardless of the duration of complaints, regardless of presence of nerve root compression on MRI, and regardless of the presence of paresis. Only patients that have had a previous episode seem to fare better on bed rest (but only in their personal subjective statement of the outcome). As discussed in chapter seven, this effect is probably more a result of the patient's expectation and previous experiences than a true effect of bed rest. The sample size was somewhat lower than originally intended, because after an inclusion period of 23 months the research ran out of time. The 95% confidence intervals around the effect estimates show that, in spite of this, no clinically relevant difference in primary outcome has been unjustly denied. The drop-out rate as well as the loss to follow-up at three months were acceptable.

The intention-to-treat analysis of the trial led to the conclusion that the advice to stay in bed for a period of 14 days does not produce a better outcome than watchful waiting. In our view, the compliance with both intervention groups was high. In other words, a lack of contrast is not a likely cause of the absence of a significant difference. Even if the compliance would have been poor, the resulting loss of contrast might be interpreted as inherent to the bed rest strategy. As long as an adequate effort is made to optimize the compliance, a lack of compliance with bed rest need not signify a methodological shortcoming.

So far, no therapy has been unequivocally shown to improve the outcome in patient with sciatica. Awaiting the eventual outcome of the natural course as a strategy seems both to be practical and to avoid prescribing a multitude of therapies of unestablished virtue. Nevertheless, relying solely on "watching and waiting" will leave both physician and patient incompletely satisfied. Of course, some hope might be vested in

the next trial of epidural steroids or a new trial of the efficacy of manual therapy. Also, new therapies e.g. addressing the inflammatory component of sciatica, might be developed. An alternative to focussing on the treatment is to focuss on the patients treated. Many treatment options have been shown not to work for low back pain in its aspecific definition. A refinement of classification schemes for low back pain might reveal therapy benefit for specific subgroups. Similarly, when the overall impression is that no therapy is efficacious for the entire group of sciatica patients, subgroups of patients might still benefit from particular therapies. Classification of sciatica into subgroups based on etiology is troubled by the inaccuracy of the etiologic diagnosis (with a lack of gold standard) and/or a lack of specific therapeutic options. Inflammation of the nerve root is the most telling example of an etiology that is probably of great relevance in the development of sciatica as shown in experimental studies^{227 228 229 230}. Maybe, enhancement of lumbosacral nerve roots on MRI studies with contrast administration provides a means of making inflammation more likely^{231 232 233}, but as yet inflammation cannot be accurately diagnosed. Consequently, specific treatments (new options or existing ones such as epidural steroids) cannot be implemented exclusively in patients with sciatica due to inflammation. Epidural fibrosis²³⁴, and sciatica nerve root cysts²³⁵, are examples of sciatica etiologies that can be diagnosed accurately, but treatment options are lacking or of unproven benefit. Only sciatica due to disc herniation seems to be an unequivocal subgroup: the diagnosis is relatively accurate and specific treatment options, especially surgical ones, exist (the value of conservative treatment is much the same as for the overall group of sciatica patients). Yet, surgical therapy seems to be indicated only for a subgroup within the subgroup of sciatica due to disc herniation, namely those that have had a sufficient severity of pain for a sufficient time period. Spinal stenosis is also recognized as an important subgroup²³⁶ but the body of evidence supporting diagnosis and treatment is much smaller than for nerve root compression²³⁷. Alternatively, classification might be based on expected outcome. Patients expected to have a favourable spontaneous recovery might receive no treatment or treatment solely directed at expediting recovery or preventing recurrences. More extensive treatment strategies might be reserved for patients expected to have an unfavourable natural course. The question central to these presumptions is:

Can the outcome of sciatica be predicted at an early stage?

In chapter 8 it is shown that a favourable outcome in sciatica patients can be predicted by an increased finger-floor distance, having more pain on sitting and being active in sports. Consistent predictors of an unfavourable outcome are a positive straight leg raising test, a higher Quetelet index and subjective numbness or sensory loss. Several types of outcome can be calculated by using the models presented. In the text the example was given of the patient older than 51 years with more pain in the leg than in the back and a positive straight leg raising test having a likelihood of not returning to work within three months of more than 90%. The benefit of rehabilitation programs aimed at return to work might be investigated specifically in this group or at least any trial investigating such programs should define such patients as a subgroup.

There is an important pitfall in appraising the relation between the clinical finding

and the outcome. The course represented is not the natural course of the disease. Half of the patients were randomly assigned to keep in bed for two weeks, while the other half were assigned to a regimen allowing mobilisation. Some received physical therapy or underwent surgery at a later stage. If these latter therapies are related to particular outcomes, tests that form the indications for these therapies invariably predict the outcome.

As mentioned before, surgery is a treatment option in a subgroup of patients with sciatica due to disc herniation of sufficient severity and duration. However, **can the patient with a very high likelihood of eventually requiring surgery be predicted at an early stage?**

Chapter nine suggests the answer is yes. The application of the regression model in sciatica patients might lead to earlier operation in a subset of patients resulting in a reduction of duration of illness and associated indirect costs. This hypothesized improvement in patient management should be empirically tested, but not until the prediction rule has been validated in a new research population. An important additional prerequisite is that the indications for surgery are similar in the next population investigated as these surgical indications are more or less implicitly incorporated into the present model.

Whenever scientific research is concluded, the hope is that some of the questions raised at the start of the research are resolved. The previous section attempted to fulfil some of this hope. Meanwhile, the inavoidable observation is that many new questions have arisen. I will mention only some of the most obvious questions, all of which might be objectives of future research:

We have shown that the straight leg raising test (SLR) is not a very powerful predictor of nerve root compression, but the addition of the requirement of a restriction of the angle below 40 degrees improved diagnostic gain. Many variations of the testing procedure exist, however. Such variations may also affect the diagnostic accuracy of the SLR. It will be worthwhile to investigate how alterations in testing technique affect the diagnostic accuracy.

In chapter three and four the inter-observer variability and the accuracy of history and physical examination were reported on. It might be interesting to compare not only the test result between the observers but also the diagnostic accuracy between the observers to discover sources of inaccuracy.

Diagnosing the inflammatory component of sciatica is a great challenge. A possible indicator of nerve root inflammation may be nerve root enhancement on MRI with Gadolinium. Relatively few studies have addressed the topic, and they have dealt with small and highly selected patient populations. The diagnostic properties of MRI with regard to nerve root enhancement deserve a more rigorous investigation.

The systematic review has shown that epidural steroids are effective in sciatica, especially when patients have a disease duration of less than three months. A more recent, methodologically adequate trial of epidural steroids demonstrated only a modest effect of steroids. Yet, the patients' disease duration in the latter trial was three months on average. Another trial of epidural steroids particularly in acute sciatica

patients seems called for.

The efficacy of manipulation in the treatment of sciatica has been investigated in one trial, only. Unfortunately this RCT had several methodological flaws. It is advisory to study the efficacy in another trial.

The preventive effects of exercise therapy particularly in sciatica have not been verified in an empirical study. Nevertheless, it is prescribed frequently both by general practitioners and neurologists. It seems to be yet another candidate for a RCT.

The model proposed to predict surgery in chapter nine requires further validation in a new research population. If the prediction rule is shown to be valid, the effects of the rule can be empirically investigated.

While it shows that the management of sciatica continues to raise new questions, there seems little in the way of evidence based medicine to support the management of cervical nerve root compression. Meanwhile, other challenges beckon in the field of neurology besides the painful afflictions of nerve roots. The medical doctor participating in scientific research is in the luxury position of having much of his research objectives offered to him by his patients. Medical science starts in the effort to find an answer to the questions the patients raise, and medical science evolves in the effort to critically evaluate the answers found. I hope to become a better doctor along the way.

Sciatica is not the hottest topic in neurology, yet, any improvement in the management of the syndrome is bound to profoundly change the neurological outpatient clinic. The impression may have existed that all we needed to know was already known. I hope to have at least shown that this impression is incorrect.

R eferences

---

1. Andersson GBJ, Deyo RA. History and physical examination in patients with herniated lumbar discs. *Spine*. 1996;21:10S-18S.
2. Papageorgiou AC, Croft PR, Ferry S, Jayson MIV, Silman AJ. Estimating the prevalence of low back pain in the general population. *Spine*. 1995;20:1889-1894.
3. Frymoyer JW. Lumbar disk disease: epidemiology. *Instr Course Lect*. 1992;41:217-23.
4. Kelsey JL, White Ad. Epidemiology and impact of low-back pain. *Spine*. 1980;5:133-42.
5. Esser CJ, Weel Cv. Het beloop van uitstralende rugpijn. *Huisarts en wetenschap*. 1985;28:335-337.
6. Barker ME. Pain in the back and leg: a general practice survey. *Rheumatol Rehabilitation*. 1977;16:37-45.
7. Zuidema H. National statistics in the Netherlands. *Ergonomics*. 1985;28:3-7.
8. Robinson JS. Sciatica and the lumbar disk syndrome: a historic perspective. *South Med J*. 1983;76:232-8.
9. Viets HR. Domenico Cotugno: his description of the cerebrospinal fluid, with a translation of part of his *De Ischiade Nervosa Commentarius* (1764) and a bibliography of his important works. *Bull Inst Hist Med*. 1935;3:701-738.
10. Lasegue. Considerations sur la sciatique. *Arch Generales de Medecine*. 1864;2:558-580.
11. Beurmann D. Note sur un signe peu connu de la sciatique.- Recherches experimentales. *Arch Phys Norm Path*. 1884;3:375-380.
12. Hall GW. Neurologic signs and their discoverers. *JAMA*. 1930;95:703-707.
13. Forst JJ. Contribution to the clinical study of sciatica. Translation of: "Contribution a l'etude clinique de la sciatique. These pour le doctoral en medecine (1881). *Arch Neurol*. 1969;21:220-221.
14. Bazzi F. Storia di un sintoma: il Lasegue. *Rivista di storia della medicina*. 1964;8:159-171.
15. Wilkins RB. Lasegue's sign. *Arch Neurol*. 1969;21:219-21.
16. Fajerstajn J. Ueber das gekreuzte Ischiasphaenomen. *Wiener Klin Wochenschr*. 1901;14:41-47.
17. Woodhall B, Hayes GJ. The well-leg-raising test of Fajersztajn in the diagnosis of ruptured lumbar intervertebral disc. *J Bone Joint Surg*. 1950;33A:786-792.
18. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *Jama*. 1992;268:760-765.
19. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology. A basic science for clinical medicine. . 2nd ed: Little, Brown and Company; 1991:439.
20. Arroll B, Schechter MI, Sheps SB. The assessment of diagnostic tests: a comparison of medical literature in 1982 and 1985. *J Gen Intern Med*. 1985;3:443-447.
21. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med*. 1978;299:926-930.
22. Knottnerus JA, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol*. 1992;45:1143-1154.
23. Irwig L, Tosteson ANA, Gatsonis C, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Int Med*. 1994;120:667-676.
24. Begg CB. Biases in the assessment of diagnostic tests. *Statistics in Medicine*. 1987;6:411-423.
25. Dersimonian R, Laird NM. Meta-analysis in clinical trials. *Controlled clinical trials*. 1986;7:177-188.
26. Laird NM, Mosteller F. Some statistical methods for combining experimental results. *Int J Technol Assess Health Care*. 1990;6:5-30.
27. Knutsson B. Comparative value of electromyographic, myelographic and clinical-neurological examinations in diagnosis of lumbar root compression syndrome. *Acta Orthop Scand Suppl*. 1961;49:1-134.
28. Uden A, Landin LA. Pain drawing and myelography in sciatic pain. *Clin Orthop*. 1987;216:124-30.
29. Kerr RSC, Cadoux-Hudson TA, Adams CBT. The value of accurate clinical assessment in the surgical management of the lumbar disc protrusion. *J Neurol Neurosurg Psych*. 1988;51:169-173.
30. Haldeman S, Shouka M, Robboy S. Computed tomography, electrodiagnostic and clinical findings in chronic workers' compensation patients with back and leg pain. *Spine*. 1988;13:345-50.
31. Hudgins WR. The crossed straight leg raising test: a diagnostic sign of herniated disc. *J Occup Med*. 1979;21:407-8.
32. Kosteljanetz M, Espersen JO, Halaburt H, Miletic T. Predictive value of clinical and surgical findings in

REFERENCES

- 1979;21:407-8.
32. Kosteljanetz M, Espersen JO, Halaburt H, Miletic T. Predictive value of clinical and surgical findings in patients with lumbago-sciatica a prospective study (part I). *acta neurochir.* 1984;73:67-76.
33. Spangfort E. Lasegue's sign in patients with lumbar disc herniation. *Acta Orthop Scand.* 1971;42:459.
34. Kosteljanetz M, Bang F, Schmidt-Olsen S. The clinical significance of straight-leg raising (Lasegue's sign) in the diagnosis of prolapsed lumbar disc. *Spine.* 1988;13:393-5.
35. Hirsch C, Nachemson A. The reliability of lumbar disk surgery. *Clin Orthop.* 1963;29:189-195.
36. Aejmelaeus R, Hiltunen H, Harkonen M, Silfverhuth M, Vaha-Tahlo T, Tunturi T. Myelographic versus clinical diagnostics in lumbar disc disease. *Arch Orthop Trauma Surg.* 1984;103:18-25.
37. Mann Nd, Brown MD, Enger I. Expert performance in low-back disorder recognition using patient pain drawings. *J Spinal Disord.* 1992;5:254-9.
38. Hudgins WR. Computer-aided diagnosis of lumbar disc herniation. *Spine.* 1983;8:604-615.
39. Jensen OH. The medial hamstring reflex in the level-diagnosis of a lumbar disc herniation. *Clin Rheumatol.* 1987;6:570-4.
40. Jensen OH. The level-diagnosis of a lower lumbar disc herniation: the value of sensibility and motor testing. *Clin Rheumatol.* 1987;6:564-9.
41. Morris EW, Di-Paola M, Vallance R, Waddell G. Diagnosis and decision making in lumbar disc prolapse and nerve entrapment. *Spine.* 1986;11:436-9.
42. Hakelius A, Hindmarsh J. The significance of neurological signs and myelographic findings in the diagnosis of lumbar root compression. *Acta Orthop Scand.* 1972;43:239-46.
43. Hakelius A, Hindmarsh J. The comparative reliability of preoperative diagnostic methods in lumbar disc surgery. *Acta Orthop Scand.* 1972;43:234-8.
44. Kortelainen P, Puranen J, Koivisto E, Lahde S. Symptoms and signs of sciatica and their relation to the localization of the lumbar disc herniation. *Spine.* 1985;10:88-92.
45. Meylemans L, Vancraeynest T, Bruyninckx F, Rosselle N. [A comparative study of EMG and CAT scan in the lumbo-ischial syndrome. II: Pain in the lumbo-ischial syndrome and the diagnostic value of clinical examination, EMG and CAT scan]. *Acta Belg Med Phys.* 1988;11:35-42.
46. Ljunggren AE. Descriptions of pain and other sensory modalities in patients with lumbago-sciatica and herniated intervertebral discs. Interview administration of an adapted McGill Pain Questionnaire. *Pain.* 1983;16:265-76.
47. Gurdjian. Herniated lumbar intervertebral discs. 1960.
48. Weise MD, Garfin SR, Gelberman RH, Katz MM, Thorne RP. Lower-extremity sensibility testing in patients with herniated lumbar intervertebral discs. *J Bone Joint Surg Am.* 1985;67:1219-24.
49. Wei CP, Cheng WC, Chang CN, Lee ST, Lui TN, Wang ADJ. Upper lumbar disc herniation. *Chang Gung Med J.* 1989;12:193-199.
50. Portnoy HD, Ahmad M. Value of the neurological examination, electromyography and myelography in herniated lumbar disc. *Mich Med.* 1972;71:429-34.
51. Mann Nd, Brown MD, Enger I. Statistical diagnosis of lumbar spine disorders using computerized patient pain drawings. *Comput Biol Med.* 1991;21:383-97.
52. Postacchini F, Cinotti G, Gumina S. The knee flexion test: a new test for lumbosacral root tension. *J Bone Joint Surg Br.* 1993;75:834-5.
53. Damme van W, Hessels G, Verhelst M, Van-Laer L, Van-Es I. Relative efficacy of clinical examination, electromyography, plain film radiography, myelography and lumbar phlebography in the diagnosis of low back pain and sciatica. *Neuroradiology.* 1979;18:109-18.
54. Jonsson B, Stromqvist B. Symptoms and signs in degeneration of the lumbar spine. A prospective, consecutive study of 300 operated patients. *J Bone Joint Surg Br.* 1993;75:381-5.
55. Aho AJ, Auranen A, Pesonen K. Analysis of cauda equina symptoms in patients with lumbar disc prolapse. *Acta Chir Scand.* 1969;135:413-420.
56. Christodoulides AN. Ipsilateral sciatica on femoral nerve stretch test is pathognomonic of an L4/5 disc protrusion. *J Bone Joint Surg Br.* 1989;71:88-9.
57. Albert TJ, Balderston RA, Heller JG, et al. Upper lumbar disc herniations. *J Spinal Disord.* 1993;6:351-9.
58. Aronson HA, Dunsmore RH. Herniated upper lumbar discs. *J Bone Joint Surg.* 1963;45A:311-317.
59. Borovich B, Zaaroor M, Gruskiewicz J. The syndrome of the central L-3-herniated disc with special emphasis on motor involvement. *Acta Neurochir Wien.* 1984;70:115-25.
60. Katznelson A, Nerubay J, Lev-El A. Gluteal skyline (G.S.L.). A search for an objective sign in the diagnosis of disc lesions of the lower lumbar spine. *Spine.* 1982;7:74-5.

62. Kostuik JP, Harrington I, Alexander D, Rand W, Evans D. Cauda equina syndrome and lumbar disc herniation. *J Bone Joint Surgery*. 1986;68-A:386-391.
63. Wiberg G. The absence of the Achilles tendon reflex and motor weakness in determining the level of disc protrusion. *Acta Orthop Belg*. 1969;35:617-9.
64. Jackson RP, Cain J, Jr., Jacobs RR, Cooper BR, McManus GE. The neuroradiographic diagnosis of lumbar herniated nucleus pulposus: II. A comparison of computed tomography (CT), myelography, CT-myelography, and magnetic resonance imaging. *Spine*. 1989;14:1362-7.
65. 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. *Neuroradiology*. 1993;186:731-738.
66. Feinstein AR. Clinical Epidemiology: the Architecture of Clinical Research. Philadelphia: Saunders; 1985.
67. Knottnerus JA. The effects of disease verification and referral on the relationship between symptoms and disease. *Medical Decision Making*. 1987;7:139-148.
68. Lachs MS, Nachamkin I, Edelstein PH, Goldman J, Feinstein AR. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. *Ann Intern Med*. 1992;117:135-140.
69. Knottnerus JA. Diagnostic prediction rules: principles, requirements, and pitfalls. *Primary Care*. 1995;22:341-363.
70. Hoffman RM, Wheeler KJ, Deyo RA. Surgery for herniated discs: a literature synthesis. *J Gen Int Med*. 1993;8:487-496.
71. Deyo RA. Nonsurgical care of low back pain. *Neurosurg Clin N Am*. 1991;2:851-862.
72. Herzog RJ. The radiologic assessment for a lumbar disc herniation. *Spine*. 1996;21:19S-38S.
73. Nelson MA, Allen P, Clamp SE, de-Dombal FT. Reliability and reproducibility of clinical findings in low-back pain. *Spine*. 1979;4:97-101.
74. Waddell G, Main CJ, Morris EW, et al. Normality and reliability in the clinical assessment of backache. *British Medical Journal*. 1982;284:1519-1523.
75. McCombe PF, Fairbank JCT, Cockersole BC, Pynsent PB. Reproducibility of physical signs in low-back pain. *Spine*. 1989;14:908-918.
76. Haerer AF. Chapter 47: Lesions causing low back pain with lumbosacral radiation. *DeJong's The Neurological Examination*. 5th ed. Philadelphia: J.B. Lippincott; 1992:596-602.
77. Kemp A. Een nieuw symptoom bij prolaps van de tussenwervelschijf. *NTV/G*. 1950;94.II:1750-1755.
78. Gathier JC. Radicular disorders due to lumbar discopathy. In: Vinken PJ, Bruyn GW, eds. *Handbook of Clinical Neurology: Tumours of the Spine and Spinal Cord, part II*. Amsterdam/Oxford: North Holland Publ.Comp.; 1976:573-604.
79. Stern WE. Howard Christian Naffziger. 1884-1961. *Surg Neurol*. 1975;4:407-409.
80. Gerhardt JJ, Rippstein JR. Gelenk und Bewegung. Goettingen: Hans Huber Verlag; 1992.
81. Karbowski K. *Schweiz Med Wochenschr*. 1984;114:27-28.
82. Medical Research Council Aids to the examination of the peripheral nervous system. *Memorandum*. London: Her majesty's stationery office; 1975:62.
83. Hijdra A, Koudstaal PJ, Roos RAC. Neurologie. . 1st ed. Utrecht: Bunge; 1994.
84. Bronisch FW. Die Reflexe und ihre Untersuchung in Klinik und Praxis. . 4 ed. Stuttgart: Georg Thieme Verlag; 1973:100.
85. Dyck P. The femoral nerve traction test with lumbar disc protrusions. *Surg Neurol*. 1976;3:163-166.
86. Nishikawa J, Sackett DL. Do fundamental clinical skills improve beyond graduation from medical school? *Clin. Res*. 1989;37:525A.
87. Hoogen HJMM vd. Low back pain in general practice: diagnosis and prognosis. *Ph. D. Thesis. General Practice/EMGO*. Wageningen: Free University Amsterdam; 1996:121.
88. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159.
89. Altman DG. Practical statistics for medical research. 1st edition ed. London: Chapman & Hall; 1996.
90. Walter SD, Irwig LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. *J Clin Epidemiol*. 1988;41:923-937.
91. Knottnerus JA, Volovics A. Medische statistiek en onderzoek in de huisartsgeneeskunde. In: Genootschap NH, ed. Utrecht: NHG; 1990:136.
92. Biering-Sorensen F. A one-year prospective study of low back trouble in a general population. *Danish Medical Bulletin*. 1984;31:362-375.
93. Deyo RA. Early diagnostic evaluation of low back pain. *J Gen Internal Medicine*. 1986;1:328-338.

REFERENCES

94. Hoogen HJMM vd, Koes BW, Eijck JTMv, Bouter LM. On the accuracy of history, physical examination and erythrocyte sedimentation rate in diagnosing low-back pain in general practice. *Spine*. 1995;20.
95. Vroomen PCAJ, de Krom MCTFM, Knottnerus JA. The diagnostic value of history and physical examination in patients with sciatica due to disc herniation; a systematic review. (*submitted*). 1998.
96. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*. 1994;331:69-73.
97. Andersson GB, Weinstein JN. Disc herniation (editorial). *Spine*. 1996;21:1S.
98. Hofman PAM, Wilmink JT. Optimising the image of the intradural nerve root: the value of MR radiculography. *Neuroradiology*. 1996;38:654-657.
99. Knottnerus JA, Volovics A. Medische statistiek en onderzoek in de huisartsgeneeskunde. *NHG-Publikaties*. Utrecht: Nederlands Huisartsen Genootschap; 1990.
100. Kleinbaum DG. Logistic regression: a self-learning text. In: Dietz K, K. G, Krickeberg K, Singer B, eds. *Statistics in the Health Sciences*. 1st ed. New York: Springer; 1994:282.
101. Metz CE. Basic principles of ROC analysis. *Seminars in Nuclear medicine*. 1978;8:283-298.
102. Hanley JA. Receiver operating characteristic (ROC) methodology. *Crit Rev Diagn Imaging*. 1989;29:307-335.
103. Connell FA, Koepsell TD. Measures of gain in certainty from a diagnostic test. *Am J Epidem*. 1985;121:744-753.
104. Bernclot Moens HJ, Hirschberg AJ, Claessens AA. Data-source effects on the sensitivities and specificities of clinical features in the diagnosis of rheumatoid arthritis: the relevance of multiple sources of knowledge for a decision-support system. *Med Decis Making*. 1992;12:250-258.
105. Ellenberger C. MR imaging of the low back syndrome. *Neurology*. 1994;44:594-600.
106. Deyo RA. Magnetic resonance imaging of the lumbar spine; terrific test or tar baby? *N Engl J Med*. 1994;331:115-116.
107. Bischoff RJ, Rodriguez RP, Gupta K, Righi A, Dalton JE, Whitecloud TS. A comparison of computed tomography-myelography, magnetic resonance imaging, and myelography in the diagnosis of herniated nucleus pulposus and spinal stenosis. *J Spinal Disord*. 1993;6:289-95.
108. Szypryt ET, Twining P, Wilde GP, Mulholland RC, Worthington BS. Diagnosis of lumbar disc protrusion. *J Bone Joint Surgery*. 1988;70-B:717-722.
109. Bozzao A, Gallucci M, Masciocchi C, Aprile I, Barile A, Passariello R. Lumbar disk herniation: MR imaging assessment of natural history in patients treated without surgery. *Radiology*. 1992;185:135-41.
110. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*. 1990;72:403-8.
111. Didry C, Lopez P, Baixas P, Simon L. [Lumbar disk herniation with no surgical treatment. Clinical and x-ray computed tomographic follow-up study]. *Presse Med*. 1991;20:299-302.
112. Hitselberger WE, Witten RM. Abnormal myelograms in asymptomatic patients. *J Neurosurg*. 1968;28:204-6.
113. Vroomen PCAJ, de Krom MCTFM, Knottnerus JA. The consistency of the history and the physical examination in the patient suspected of sciatica. (*submitted*). 1998.
114. Leeuwen van RB, Perquin WVM. De bedrustkuur bij het lumbale radiculair syndroom. *Medisch Contact*. 1986;50:1618-1620.
115. Garfin SR, Rydevik B, Lind B, Massie J. Spinal nerve root compression. *Spine*. 1995;20:1810-20.
116. Mixter WJ, Barr JS. Rupture of the intervertebral disc with involvement of the spinal canal. *N Engl J Med*. 1934;211:210-215.
117. Munro D. The diagnosis of posterior herniation of the lumbar intervertebral disks. *N Engl J Med*. 1945;232:149-160.
118. Lindblom K, Rexed B. Spinal nerve injury in dorsolateral protrusions of the lumbar disks. *J Neurosurg*. 1948;5:413-432.
119. Kirstein L. An after-examination of operated and non-operated cases with clinical symptoms of herniated disc. *Acta medica scand*. 1945;120:93-106.
120. Falconer MA, McGeorge M, Begg AC. Observations on the cause and mechanism of symptom-production in sciatica and low-back pain. *J Neurol Neurosurg Psych*. 1948;11:13-26.
121. Briggs H, Krause J. The intervertebral foraminotomy for relief of sciatic pain. *J Bone Joint Surg*. 1945;27B:475-478.
122. Hardy RW, Davis CH. Extradural spinal cord and nerve root compression from benign lesions of the lumbar area. In: Youmans, ed. *Textbook of Neurological surgery*. 3rd ed. Philadelphia: WB Saunders; 1990.

123. Spencer DL, Irwin GS, Miller JAA. Anatomy and significance of fixation of lumbosacral nerve roots in sciatica. *Spine*. 1983;8:672-679.
124. 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.
125. Spencer DL, Miller JAA, Bertolini JE. The effect of intervertebral disc space narrowing on the contact force between the nerve root and a simulated disc protrusion. *Spine*. 1984;9:422-426.
126. Olmarker K, Rydevik B, Holm S. Edema formation in spinal nerve roots induced by experimental graded compression: an experimental study on the pig cauda equina with special references to differences in effects between rapid and slow onset compression. *Spine*. 1989;14:569-573.
127. Cailliet R. Low back pain syndrome. 2nd ed. Philadelphia: F.A. Davis Company; 1968.
128. Reulen H-J, Müller A, Ebeling U. Microsurgical anatomy of the lateral approach to extraforaminal lumbar disc herniations. *Neurosurgery*. 1996;39:345-351.
129. Smith SA, Massie JB, Chesnut R, Garfin SR. Straight leg raising - anatomical effects on the spinal nerve root without and with fusion. *Spine*. 1993;18:992-999.
130. Beuls EAM, Mameren Hv, Vroomen PCAJ. Caudoscopic experiences and a new pathoanatomic concept for sciatica. *Minimally invasive neurosurgery*. 1996;39:1-32.
131. Parke WW, Watanabe R. The intrinsic vasculature of lumbosacral spinal nerve roots. *Spine*. 1985;10:508-515.
132. Rydevik B, Brown MD, Lundborg G. Pathoanatomy and pathophysiology of nerve root compression. *Spine*. 1984;9:7-13.
133. Vroomen PCAJ, Krom MCTFMd, Wilmink JT, Kester ADM, Knottnerus JA. The diagnostic value of the history and physical examination in patients suspected of sciatica- (submitted). 1998.
134. Saal JA, Saal JS. Nonoperative treatment of herniated lumbar intervertebral disc with radiculopathy. An outcome study. *Spine*. 1989;14:431-7.
135. Pearce J, Moll JMH. Conservative treatment and natural history of acute lumbar disc lesions. *J Neurol Neurosurg Psychiatr*. 1967;30:13-17.
136. Bell GR, Rothman RH. The conservative treatment of sciatica. *Spine*. 1984;9:54-56.
137. Malmivaara A, Haekinen U, Aro T, et al. The treatment of acute low back pain- bed rest, exercises, or ordinary activity? *New Engl J Med*. 1995;332:351-355.
138. Deyo RA, Diehl AK, Rosenthal M. How many days of bed rest for acute low back pain? *N Engl J Med*. 1986;315:1064-70.
139. Koes BW, Hoogen HMMvd. Efficacy of bed rest and orthoses of low-back pain. *Eur J Phys Med Rehab*. 1994;4:86-93.
140. Waddell G, Feder G, Lewis M. Systematic reviews of bed rest and advice to stay active for acute low back pain. *Br J Gen Practice*. 1997;47:647-652.
141. Schultz U, Koehler D, Kuemeyer M, Staebler-Lehr A. Zum Spontanverlauf des Discusvorfalls beim lumbalen Wurzelkompressionssyndrom. *Nervenarzt*. 1988;59:661-668.
142. Teplick JG, Haskin ME. Spontaneous regression of herniated nucleus pulposus. *AJR*. 1984;145:371-375.
143. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ*. 1994;309:1351-55.
144. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309:1286-91.
145. Sackett DL. Bias in analytical research. *J Chronic Dis*. 1979;32:51-68.
146. Koes BW, Bouter LM, Beckermann H, Heijden GJMGvd, Knipschild PG. Physiotherapy exercises and back pain: a blinded review. *Br Med J*. 1991a;302:1572-1576.
147. Koes BW, Assendelft WJJ, Heijden GJMGvd, Bouter LM, Knipschild PG. Spinal manipulation and mobilization for back and neck pain: a blinded review. *Br Med J*. 1991b;303:1298-1303.
148. Koes BW, Tulder-van MW, Windt-van-der DWM, Bouter LM. The efficacy of back schools: a review of randomized controlled trials. *J Clin Epidemiol*. 1994;47:851-862.
149. Group Consort. A proposal for structured reporting of randomized controlled trials. *JAMA*. 1994;272:1926-1931.
150. Galbraith RF. A note on the graphical presentation of estimated odds ratios from several clinical trials. *Statist Med*. 1988;7:889-894.
151. Brand R, Kragt H. The importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials. *Statist Med*. 1992;11:2077-2082.
152. Bush K, Hillier S. A controlled study of caudal epidural injections of triamcinolone plus procaine for the

REFERENCES

- management of intractable sciatica. *Spine*. 1991;16:572-5.
153. Cuckler JM, Bernini PA, Wiesel SW, Booth R, Jr., Rothman RH, Pickens GT. The use of epidural steroids in the treatment of lumbar radicular pain. A prospective, randomized, double-blind study. *J Bone Joint Surg Am*. 1985;67:63-6.
154. Haimovic IC, Beresford HR. Dexamethasone is not superior to placebo for treating lumbosacral radicular pain. *Neurology*. 1986;36:1593-1594.
155. Hedeboe J, Buhl M, Ramsing P. Effects of using dexamethasone and placebo in the treatment of prolapsed lumbar disc. *Acta Neurol Scand*. 1982;65:6-10.
156. Kleenerman L, Greenwood R, Davenport HT, White DC, Peskett S. Lumbar epidural injections in the treatment of sciatica. *British Journal of Rheumatology*. 1984;23:35-38.
157. Ridley MG, Kingsley GH, Gibson T, Grahame R. Outpatient lumbar epidural corticosteroid injection in the management of sciatica. *Br J Rheumatol*. 1988;27:295-9.
158. Rogers P, Nash T, Schiller D, Norman J. Epidural steroids for sciatica. *The pain clinic*. 1992;5:67-72.
159. Zahaar MSe. The value of caudal epidural steroids in the treatment of lumbar neural compression syndromes. *J Neurol Orthop Med Surg*. 1991;12:181-184.
160. Coxhead CE, Meade TW, Inskip H, North WRS. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet*. 1981;1065-1068.
161. 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;18:1433-8.
162. Hofferberth B, Gottschaldt M, Grass H, Buttner K. [The usefulness of dexamethasonephosphate in the conservative treatment of lumbar pain--a double-blind study (author's transl)]. *Arch Psychiatr Nervenkr*. 1982;231:359-67.
163. Berry H, Hutchinson DR. A multicentre placebo-controlled study in general practice to evaluate the efficacy and safety of tizanidine in acute low-back pain. *J. International Medical Research*. 1988;16:75-82.
164. Gibson T, Dilke TFW, Grahame R. Chymoral in the treatment of lumbar disc prolapse. *Rheumatology and Rehabilitation*. 1975;14:186-191.
165. Goldie I. A clinical trial with indomethacin (indomec(R)) in low back pain and sciatica. *Acta Orthopaed Scand*. 1968;39:117-128.
166. Weber H, Aasand G. The effect of phenylbutazone on patients with acute lumbago-sciatica. A double blind trial. *J. Oslo City Hospitals*. 1980;30:69-73.
167. Borms T. Comparison of injectable formulations of tieprofenic acid and ketoprofen in acute lumbar sciatica. *Drugs*. 1988;35:85-87.
168. Tesio L, Merlo A. Autotrraction versus passive traction: an open controlled study in lumbar disc herniation. *Arch Phys Med Rehab*. 1993;74:871-876.
169. Mathews JA, Mills SB, Jenkins M, al e. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J of Rheumat*. 1987;26:416-423.
170. Snock W, Weber H, Jorgensen B. Double blind evaluation of extradural methyl prednisolone for herniated lumbar discs. *Acta Orthop Scand*. 1977;48:635-41.
171. Dilke TFW, Burry HC, Grahame R. Extradural corticosteroid injection in management of lumbar nerve root compression. *British Medical Journal*. 1973;2:635-637.
172. Weber H, Ljunggren AE, Walker L. Traction therapy in patients with herniated lumbar intervertebral discs. *J. Oslo City Hospitals*. 1984;34:61-70.
173. Ljunggren AE, Walker L, Weber H, Amundsen T. Manual traction versus isometric exercises in patients with herniated intervertebral lumbar discs. *Physiotherapy Theory and Practice*. 1992;8:207-213.
174. Ljunggren AE, Weber H, Larsen S. Autotrraction versus manual traction in patients with prolapsed lumbar intervertebral discs. *Scand J Rehab Med*. 1984;16:117-124.
175. Porsman O, Friis H. Prolapsed lumbar disc treated with intramuscularly administered dexamethasonephosphate. *Scand J Rheumat*. 1979;8:142-144.
176. Weber H. Lumbar disc herniation, a controlled, prospective study with ten years of observation. *Spine*. 1983;8:131-140.
177. Larsson U, Choeler U, Lidstroem A, et al. Auto-traction for treatment of lumbago-sciatica. *Acta Orthopaed Scand*. 1980;51:791-798.
178. Mathews W, Morkel M, Mathews J. Manipulation and traction for lumbago and sciatica: physiotherapeutic techniques used in two controlled trials. *Physiotherapy Practice*. 1988;4:201-206.
179. Moher D, Jadad AR, Nichol G, et al. Assessing the quality of randomized controlled trials: an automated bibliography of scales and checklists. *Controlled clinical trials*. 1995;16:62-73.

180. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *The Lancet*. 1991;337:867-872.
181. Pocock SJ. Clinical trials: a practical approach. Chichester: Wiley; 1983.
182. Carette S, Leclaire R, Marcoux S, et al. Epidural corticosteroid injections for sciatica due to herniated nucleus pulposus. *New England Journal of Medicine*. 1997;336:1634-40.
183. Orlando MP, Sherman MO. Corticosteroid injections for sciatica (letter). *N Engl J Med*. 1997;337:1242.
184. Gillies JH, Ward JH, Griesdale DE. Corticosteroid injections for sciatica (letter). *N Engl J Med*. 1997;337:1242.
185. Cowan NC, Bush K, Katz DE, Gishen P. The natural history of sciatica: a prospective radiological study. *Clinical radiology*. 1992;46:7-12.
186. Deyo RA. Conservative therapy for low back pain: distinguishing useful from useless therapy. *JAMA*. 1983;250:1057-1063.
187. Vroomen PCAJ, Krom MCTFMD, Slofstra DP, Knottnerus JA. A systematic review of the conservative treatment of sciatica due to disc herniation. (submitted). 1998.
188. Teasell R, Dittmer DK. Complications of immobilization and bed rest. Part 2: Other complications. *Can Fam Physician*. 1993;39:1440-2.
189. Bonde Petersen F, Suzuki Y et al. . Effect of 20 days bed rest upon peripheral capillary filtration rate, venous compliance and blood flow in arms and legs. *Acta Physiol Scand Suppl*. 1994;616:65-9.
190. Convertino VA, Bloomfield SA, Greenleaf JE. An overview of the issues: physiological effects of bed rest and restricted physical activity. *Med Sci Sports Exerc*. 1997;29:187-90.
191. Dittmer DK, Teasell R. Complications of immobilization and bed rest. Part 1: Musculoskeletal and cardiovascular complications. *Can Fam Physician*. 1993;39:1428-32.
192. Ferrando AA, Stuart CA, Brunder DG, Hillman GR. Magnetic resonance imaging quantitation of changes in muscle volume during 7 days of strict bed rest. *Aviat Space Environ Med*. 1995;66:976-81.
193. Levine BD, Zuckerman JH, Pawelczyk JA. Cardiac atrophy after bed-rest deconditioning: a nonneural mechanism for orthostatic intolerance. *Circulation*. 1997;96:517-25.
194. Pedersen BJ, Schlemmer A, Hassager C, Christiansen C. Changes in the carboxyl-terminal propeptide of type I procollagen and other markers of bone formation upon five days of bed rest. *Bone*. 1995;17:91-5.
195. Suzuki Y, Murakami T, Haruna Y, et al. Effects of 10 and 20 days bed rest on leg muscle mass and strength in young subjects. *Acta Physiol Scand Suppl*. 1994;616:5-18.
196. Takenaka K, Suzuki Y, Kawakubo K, et al. Cardiovascular effects of 20 days bed rest in healthy young subjects. *Acta Physiol Scand Suppl*. 1994;616:59-63.
197. Beurskens AJHM, Vet HCWd, Kke AJA. Responsiveness of functional status in low back pain. A comparison of different instruments. *Pain*. 1997.
198. Melzack R, Katz J. The McGill Pain Questionnaire: appraisal and current status. In: Turk DC, Melzack R, eds. *Pain*; 1992:152-168.
199. Deyo RA. Comparative validity of the sickness impact profile and shorter scales for functional assessment in low-back pain. *Spine*. 1986;11:951-954.
200. 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*. 1982;8:141-143.
201. Hsieh CY, Phillips RB, Adams AH, Pope MH. Functional outcomes of low back pain: comparison of four treatment groups in a randomized controlled trial. *J Manipulative Physiol Ther*. 1992;15:4-9.
202. Beurskens AJ, Vet HCd, Koeke AJ, Heyden GJvd, Knipschild PG. Measuring the functional status of patients with low back pain; assessment of the quality of four disease specific questionnaires. ; 1994.
203. Cox DR. Regression models and life tables. *J. Roy. Statist. Soc. B*. 1972;34:187-220.
204. Fordyce WE. Pain and suffering. *American Psychologist*. 1988;43:276-283.
205. Linton SJ, Hellsing AL, Andersson D. A controlled study of the effects of an early intervention on acute musculoskeletal pain problems. *Pain*. 1993;54:353-9.
206. Vroomen PCAJ, Krom MCTFMD, Wilmlink JT, Kester ADM, Knottnerus JA. A randomized controlled trial of 14 days of bed rest versus watchful waiting for a lumbosacral radicular syndrome. (submitted). 1998.
207. Saal JA, Saal JS, Herzog RJ. The natural history of lumbar intervertebral disc extrusions treated nonoperatively. *Spine*. 1990;15:683-686.
208. Nykvist F, Alaranta H, Hurme M, Karppi SL. Clinical findings as outcome predictors in rehabilitation of patients with sciatica. *Int J Rehabil Res*. 1991;14:131-44.
209. Barrios C, Ahmed M, Arroategui JJ, Bjrnsson A. Clinical factors predicting outcome after surgery for herniated lumbar disc: an epidemiological multivariate analysis. *J Spinal Disord*. 1990;3:205-9.

REFERENCES

210. Herron LD, Turner J, Clancy S, Weiner P. The differential utility of the Minnesota Multiphasic Personality Inventory. *Spine*. 1986;11:847-850.
211. Wiltse LL, Rocchio PD. Preoperative psychological tests as predictors of success of chemonucleolysis. *J Bone Joint Surg*. 1975;57A:478-483.
212. Tullberg T, Svanborg E, Isacson J, Grane P. A preoperative and postoperative study of the accuracy and value of electrodiagnosis in patients with lumbosacral disc herniation. *Spine*. 1993;18:837-42.
213. Modic MT, Masaryk T, Boumprey F, Gormastic M, Bell G. Lumbar herniated disk disease and canal stenosis: prospective evaluation by surface coil MRI, CT, and myelography. *AJR*. 1986;147:757-765.
214. Knottnerus JA. Medical decision making by general practitioners and specialists. *Family practice*. 1991;8:305-307.
215. Roland M, Morris T. A study of the natural history of low-back pain (part II) development of guidelines for trials of treatment in primary care. *Spine*. 1982;3:145-150.
216. 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.
217. Kim KY, Kim YT, Lee CS, Shin MJ. MRI classification of lumbar herniated intervertebral disc. *Orthopedics*. 1992;15:493-7.
218. 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.
219. Larequi-Laubert T, Vader J-P, Burnand B, et al. Appropriateness of indications for surgery of lumbar disc hernia and spinal stenosis. *Spine*. 1996;22:203-209.
220. Komori H, Shinomiya K, Nakai D, Yamaura, Takeda S, Furuya K. The natural history of herniated nucleus pulposus with radiculopathy. *Spine*. 1996;21:225-229.
221. Stam J. Consensus over het lumbosacrale radiculair syndroom. *Ned Tijdschr v Geneesk*. 1996;140:2661.
222. Waddell G, Turk DC. Clinical assessment of low back pain. In: Turk DC, Melzack R, eds. *Pain*; 1992:16-36.
223. Naftulin S, Fast A, Thomas M. Diabetic lumbar radiculopathy: sciatica without disc herniation. *Spine*. 1993;18:2419-2422.
224. Watson T, Benson D. Nondiscogenous back and leg pain. In: Youmans, ed. *Textbook of Neurological Surgery*. 3rd ed. Philadelphia: WB Saunders; 1990.
225. Trowbridge WV, French JD. The "false positive" lumbar myelogram. *Neurology*. 1954;4:339-344.
226. Deyo RA, Loeser JD, Bigos SJ. Herniated lumbar intervertebral disk. *Ann Int Med*. 1990;112:598-603.
227. Takahashi H, Suguro T, Okazima Y, Motegi M, Okada Y, Kakiuchi T. Inflammatory cytokines in the herniated disc of the lumbar spine. *Spine*. 1996;21:218-224.
228. O'Donnell JL, O'Donnell AL. Prostaglandin E₂ content in herniated lumbar disc disease. *Spine*. 1996;21:1653-1656.
229. Kawakami M, Tamaki T, Weinstein JN, Hashizume H, Nishi H, Meller ST. Pathomechanism of pain-related behavior produced by allografts of intervertebral disc in the rat. *Spine*. 1996;21:2101-2107.
230. Kang JD, Georgescu HI, McIntyre-Larkin L et al. Herniated lumbar intervertebral discs spontaneously produce matrix metalloproteinases, nitric oxide, interleukin-6, and prostaglandin E₂. *Spine*. 1996;21:271-277.
231. Vroomen PCAJ, Hapert SJMv, Acker REHv, Beuls EAM, Kessels AGH, Wilmink JT. The clinical significance of Gadolinium enhancement of lumbar herniations and nerve roots in pre-operative MRI. *Neuroradiology (accepted for publication)*. 1998.
232. Nguyen C, Houghton VM, Ho K-C, et al. Contrast enhancement in spinal nerve roots: an experimental study. *AJNR*. 1995;16:265-268.
233. Jenkins JR. MR of enhancing nerve roots in the unoperated lumbosacral spine. *AJNR*. 1993;14:193-202.
234. Ross JS, Robertson JT, Frederickson RCA, et al. Association between peridural scar and recurrent radicular pain after lumbar discectomy: Magnetic Resonance Evaluation. *Neurosurgery*. 1996;38:855-863.
235. Plewes JL, Jacobson I. Sciatica caused by sacral-nerve-root cysts. *Lancet*. 1970;2:799-802.
236. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part III: 1-year outcomes of surgical and non-surgical management of lumbar spinal stenosis. *Spine*. 1996;21:1787-1795.
237. Atlas SJ, Deyo RA, Patrick Dlet al. The Quebec Task Force classification for spinal disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. *Spine*. 1996;21:2885-2892.
238. Frymoyer JW. Lumbar disk disease: epidemiology. *Instr Course Lect*. 1992;41:217-23.

Summary

The first part of this thesis focusses on the value of history and physical examination in the diagnosis of sciatica due to disc herniation. The second part addresses the value of the conservative treatment of sciatica.

In chapter two the literature is systematically reviewed to investigate the value of history and physical examination in the diagnosis of sciatica due to disc herniation. The literature for the period 1965-1994 was searched. 37 studies meeting the selection criteria were retrieved and were systematically and independently read by three readers to determine diagnostic test properties using a standard scoring list to determine the methodological quality of the diagnostic information. Few studies investigated the value of the history. Pain distribution seemed to be the only useful history item for determination of the level of disc herniation. Of the physical examination signs the SLR was the only sign consistently reported to be sensitive for sciatica due to disc herniation. However, the sensitivity values varied greatly, the pooled sensitivity and specificity being 0.85 and 0.52, respectively. The crossed SLR was the only sign shown to be specific; the pooled sensitivity and specificity were 0.30 and 0.84 respectively. There was considerable disagreement on the specificity of the other neurological signs (paresis, sensory loss, reflex loss). All studies were biased in one way or another with three major types of bias being predominant: sequential-ordering bias, spectrum bias, and diagnostic/test review bias. Several other methodological drawbacks limited the generalizability of the study results.

In chapter three the technique of the history and physical examination in patients with radiating pain into the leg is described. The reproducibility of these signs and symptoms had been either investigated in populations of low back pain patients or had not been investigated at all. A random selection of 91 patients was investigated by a neurologist-resident couple. These two clinicians disagreed on the presence of nerve root involvement in 1 of 4 patients after history and in 1 of 5 patients after physical examination. Asking for an increase of pain on coughing/straining/sneezing, a feeling of coldness in the legs and urinary incontinence and the investigation of muscle strength, sensory loss, reflex changes, SLR and Bragard's sign provide the most consistent results.

In chapter four we report on a diagnostic study designed to investigate the association between clinical findings and a Magnetic Resonance Imaging (MRI) diagnosis of nerve root compression. Patients with pain radiating into the leg, and of sufficient severity for bed rest therapy to be considered were investigated. 274 patients were selected. In 2 by 2 tables, nerve root compression was associated with six history

and ten physical examination findings (depressed reflexes, four nerve root tension signs, three signs of decreased muscle strength, and two signs of decreased spinal mobility). When the contributions of other findings were accounted for, six symptoms and two signs (increased finger-floor distance and paresis) significantly predicted nerve root compression. The straight leg raising test was not predictive, unless the requirement of a restriction of the angle below 45 degrees was added. The Receiver Operating Characteristic (ROC) showed that the area-under-the-curve increased from 0.80 to 0.83 when the physical examination was added to the history. The bivariate results agree best with our ideas of valuable signs and symptoms of nerve root compression. However, most tests had lower sensitivity and specificity values than previously reported. Moreover, the multivariate analysis and the ROC showed that most of the physical examination findings, including the straight leg raising test, add little diagnostic information to the history.

In chapter five clinical data are correlated to details of MRI anatomy. Anatomical studies have offered many theories on how clinical findings are produced without clearly demonstrating the clinical relevance of the observations. Clinico-anatomical series are scarce and are affected to a great extent by selection bias. Clinical data showed that the higher the level of symptomatic disc herniations, the more likely the compression will be more laterally situated. Medial herniations will more likely give rise to low back pain. Classical findings of sciatica (typically dermatomal pain, pain increase on coughing, sneezing or straining, and positive straight leg raising) were most likely with compression of the nerve root in the axilla and with mediolateral disc herniations. L3, L4, L5 and S1 nerve roots each tend to be compressed at different sites along the rostral-caudal course of the nerve root. Disc herniations become symptomatic at different sites for each disc level. The drawing produced *a priori* could well explain these findings. The occurrence of particular clinical findings can be predicted by specific patho-anatomical findings.

In chapter six randomized controlled trials of conservative therapies for sciatica are reviewed. A standardized literature search for randomized controlled trials (RCTs) was done. Trial methodology and outcome were registered independently by three observers. If possible, trial results were statistically pooled (= meta-analysis). 18 RCTs were finally selected. The methodological quality of these trials varied considerably. Neither traction, exercise therapy, manipulation nor drug therapy were unequivocally shown to be beneficial. Epidural steroids seem superior to placebo. No RCT of bed rest was found. Furthermore, clinical heterogeneity and methodological quality aspects seemed to affect review results. Small sample sizes in most studies may partly explain the absence of significant effects in many studies. From the review results the natural course in secondary care patients was estimated. Overall, 51% of patients with sciatica were shown to have a favourable natural course over a period of 10 days.

Chapter seven features a report on a randomized controlled trial of bed rest versus

watchful mobilizing for a duration of 14 days. Patients with sciatica of sufficient severity to justify a treatment with 14 days of bed rest were included. They were recruited from a primary care setting. Primary outcome measures were global measures of improvement over 2 and 12 weeks. Secondary measures were a variety of functional and pain scales (at 2, 3 and 12 weeks), work absenteeism and surgical intervention. No significantly higher rate of improvement was found in the 92 patients in the bed rest group than in the 91 in the watchful mobilizing group neither at two nor at twelve weeks. Visual Analogue Scales of pain intensity and bothersomeness of complaints, McGill Pain Questionnaires, and functional status questionnaires revealed no significant differences between the two groups. Work absenteeism and surgical intervention rates were similar for both groups. It was concluded that bed rest does not provide a better chance of recovery for patients with sciatica and other signs and symptoms of a lumbosacral radicular syndrome. Therefore, such patients may rest in bed, but may mobilize when possible.

In chapter eight the prognostic value of history and physical examination was described. Outcome was registered as the improvement reported by the patient after two weeks and after three months. Furthermore, the prediction of Roland disability scale scores, and visual analogue scale scores for pain in the leg and low back at three months was investigated. In 183 patients, a variety of clinical findings were shown to independently predict outcome. The most consistent predictors of a favourable outcome were previous sporting activities, more pain on sitting, and increased finger-floor distance. The most consistent predictors of an unfavourable outcome were a positive straight leg raising test, a higher Quetelet index, and numbness/subjective sensory loss.

In chapter nine we wondered whether patients with a very high likelihood of eventually undergoing surgery could be recognized at an early stage? In 183 sciatica who underwent a standardized history, physical examination and Magnetic Resonance Imaging surgery carried out in the next 6 months was recorded. Two models were constructed. Reduced model A was based on baseline findings, only, and contained 6 variables. Model B incorporated change over time as well and contained 10 variables. The area-under-the-curve (in a Receiver Operating Characteristic) for these models was 0.86 and 0.92, respectively. It was shown that at a probability cut-off of 0.60, model B predicted eventual surgery with a sensitivity of 57% and a specificity of 100%. Given the requirement of a high specificity, eventual operation could be adequately predicted by model B. The application of the model in sciatica patients was hypothesized to lead to earlier operation in a subset of patients resulting in a reduction of duration of illness and associated indirect costs.

Chapter ten offered a general discussion of the material of this thesis.

Samenvatting

Het eerste deel van dit proefschrift behandelt de waarde van anamnese en lichamelijk onderzoek bij het lumbosacrale radiculare syndroom. Het tweede deel behandelt de waarde van de conservatieve behandeling van dat syndroom.

In hoofdstuk 2 wordt de literatuur betreffende de diagnostische waarde van anamnese en lichamelijk onderzoek bij het radiculare syndroom op een systematische wijze samengevat. Literatuur uit de periode 1965 t/m 1994 werd gelezen door drie reviewers met aandacht voor de kwaliteit en inhoud van de diagnostische informatie. Er was nauwelijks literatuur die de anamnese belichtte. Van de testen uit het lichamelijk onderzoek was alleen de proef van Lasègue frequent onderwerp van studie geweest. De gerapporteerde diagnostische eigenschappen voor deze proef varieerden echter sterk. Gepoold was de sensitiviteit 0.85 en de specificiteit 0.52. Voor de gekruiste Lasègue waren deze cijfers respectievelijk 0.30 en 0.84. Er was weinig overeenstemming over de waarde van andere testen (bv. parese, sensibiliteitsuitval, reflexverschil). Alle studies waren overigens op de een of andere wijze gebiast. De belangrijkste vertekeningen traden op ten gevolge van sequential-ordering bias, spectrum bias en diagnostic/test review bias. Bovendien waren er tekortkomingen die met name de externe validiteit van de informatie beperkten.

In hoofdstuk 3 wordt de (inter-observer) reproduceerbaarheid van anamnese en lichamelijk onderzoek besproken. Voorheen is deze onderzocht in patiënten populaties met overwegend lage rugklachten. Opvallend was dat twee artsen het in 1 van de 4 patiënten oneens waren over de diagnose na de anamnese alleen en in 1 van de 5 patiënten over de diagnose na anamnese én lichamelijk onderzoek. Vragen naar pijn bij drukverhogende momenten, een koud gevoel, of urine-incontinentie en het onderzoek van de kracht, de sensibiteit en de reflexen en de proef van Lasègue en Bragard leverden de meest consistente informatie op.

In hoofdstuk 4 staat een studie naar de diagnostische waarde van anamnese en lichamelijk onderzoek centraal. Zenuwwortelcompressie bij Magnetic Resonance Imaging (MRI) vormde de Gouden Standaard. Patiënten met uitstralende pijn in het been werden gestandaardiseerd onderworpen aan zowel anamnese en lichamelijk onderzoek als aan MRI onderzoeken. De uitkomsten van zes anamnestiche vragen en van tien fysisch-diagnostische testen bleken geassocieerd met zenuwwortelcompressie. Dit resultaat stemt goed overeen met het heersende idee van wat waardevolle testen zijn bij verdenking op een radiculair syndroom. De diagnostische eigenschappen waren in het algemeen wel slechter dan voorheen beschreven. Indien -onafhankelijk van de andere testen- de voorspellende waarde van testen bekeken werd bleken zes vragen te discrimineren. Van de testen uit het lichamelijk onderzoek droegen enkel

een parese en een beperkte vinger-vloer afstand significant bij. Opvallend was dat de proef van Lasègue niet behoorde tot deze acht significante voorspellers. Het lichamelijk onderzoek voegde relatief weinig toe aan het discriminerend vermogen van de anamnese; de oppervlakte onder de Receiver Operating Characteristic verbeterde slechts van 0.80 naar 0.83.

Het gestandaardiseerd onderzoek bij 274 patiënten biedt een inzicht in de klinische bevindingen bij deze patiënten enerzijds en de MRI-bevindingen anderzijds. De MRI bevindingen kunnen gezien worden als een weergave van de anatomische situatie. Hoofdstuk vijf is een beschrijvende studie van de klinische bevindingen in het licht van bepaalde anatomische observaties. Eerdere vergelijkbare studies zijn zowel schaars als gebiast. Opvallend was dat de “typische” radiculaire verschijnselen vooral optraden indien de zenuwwortel in de oksel tussen wortelmanchet en durale zak gecompriëerd wordt en bij compressie door mediolateraal gehernieerd discusmateriaal. Bovendien viel op dat de zenuwwortels L3, L4, L5 en S1 ieder op een andere voorkeursplaats langs de rostraal-caudale as van de wortel gecompriëerd lijken te worden.

Hoofdstuk 6 biedt een systematisch overzicht van gerandomiseerde gecontroleerde trials met betrekking tot conservatieve behandelingen voor het radiculaire syndroom. Van de geselecteerde trials werd de methodologie beoordeeld en de resultaten samengevat. Dit gebeurde waar mogelijk door een meta-analyse. Van de 18 geselecteerde trials varieerde de methodologie sterk. Opvallend was het kleine aantal patiënten in de studies. Aspecten van de methodologische kwaliteit en de klinische heterogeniteit leken de uitkomsten van de studies mede te beïnvloeden. Van tractie, oefentherapie, manuele therapie of medicamenteuze behandeling werd geen onomstreden nut aangetoond. Epidurale steroïden leken beter te werken dan placebo. Het nut van bedrust bij het radiculaire syndroom bleek nooit eerder in een gerandomiseerde trial vergeleken met een controle groep. Ongeveer de helft van de patiënten lijkt gemiddeld na 10 dagen verbeterd.

Bedrust wordt frequent voorgeschreven bij het lumbosacrale radiculaire syndroom. Hoofdstuk 7 biedt een verslag van een gerandomiseerde trial van bedrust gedurende 14 dagen versus een afwachtend beleid bij dit syndroom. Patiënten werden in de eerste lijn geworven op basis van een uitstralende pijn onder het been. Er was geen verschil tussen beide groepen in het percentage patiënten dat na 2 of 12 weken verbetering aangaf. Ook met betrekking tot pijnscore op een visueel analoge schaal of de McGill Pain Questionnaire, score op functionele status vragenlijsten en werkverzuim waren er geen significante verschillen. Het percentage patiënten dat een operatieve ingreep moest ondergaan was gelijk in beide groepen. Kortom, de kans op herstel is dus niet hoger bij bedrust dan bij een afwachtend beleid. Patiënten met een radiculair syndroom hoeven dus geen bedrustkuur te houden.

In hoofdstuk 9 wordt de prognostische waarde van de klinische verschijnselen voor

patiënten met een radiculair syndroom aan de orde gesteld. Een reeks verschijnselen werd afgezet tegen verschillende uitkomsten (ervaren verbetering, pijnscores en functioneren). Gunstige prognostische factoren bleken verergering van pijn bij zitten, voorgaande sportieve activiteiten en een vinger-vloer afstand meer dan 25 cm te zijn. Een positieve proef van Lasègue, een hoge Quetelet index en subjectief gevoelsverlies bleken ongunstige prognostische factoren te zijn.

Had de uiteindelijke operatie bij een deel van de patiënten in een vroege fase voorspeld kunnen worden? Hoofdstuk 9 belicht een poging om bij 183 patiënten met een radiculair syndroom de waarschijnlijkheid van operatie in de daarop volgende 6 maanden te voorspellen aan de hand van anamnese en lichamelijk onderzoek. Testen uit de anamnese en lichamelijk onderzoek bij de eerste consultatie van de patiënt bleken de waarschijnlijkheid van operatie niet voldoende nauwkeurig te kunnen aangeven (model A). Wanneer echter ook informatie van een consultatie 2 weken later werd meegenomen, bleek een dergelijke voorspelling wel voldoende nauwkeurig mogelijk te zijn (model B). De voorspellende eigenschappen van model B dienen in een andere onderzoekspopulatie bevestigd te worden. Toepassing van een werkzaam model B in de praktijk zou tot een reductie van ziekteduur en geassocieerde indirecte kosten kunnen leiden.

Het proefschrift wordt afgesloten met een algemene discussie, een overzicht van de geraadpleegde literatuur en een Engelse en Nederlandse samenvatting.

History** sucks the lifeblood of the present, the little time man has to himself
James Joyce, *Ulysses*, 1938

**Patrick Vroomen werd als Carolus Andreas Johannes geboren op 11 juli 1969 te Brunssum, doorliep de kleuterschool spelenderwijs, bleef niet zitten op de lagere school te Born, behaalde zijn Gymnasium B diploma *summa cum laude* te Sittard, werd met goed gevolg probleem gestuurd onderwezen aan de Maastrichtse Faculteit der Geneeskunde, en wordt momenteel ingewijd in de soft- en hardware van de menselijke existentie tijdens zijn opleiding tot neuroloog bij de afdeling Neurologie van het Academisch Ziekenhuis Maastricht

Walk of life

The future accumulates like a weight upon the past. The weight upon the earliest years is easier to remove to let that time spring up like grass that has been crushed. The years following childhood become welded to their future, massed like stone, and often the time beneath cannot spring back into growth like new grass: it lies bled of its green in a new shape with those frail bloodless sprouts of another, unfamiliar time, entangled one with the other beneath the stone.

Janet Frame, *An Angle at my table*, 1984

Looking forward to the unfamiliar sprouts with hope and excitement