

From Center for Family and Community Medicine,
Department of Neurobiology, Care Sciences and Society,
Karolinska Institutet, Stockholm, Sweden

ON THE REHABILITATION OF NON-ACUTE, NON-SPECIFIC SPINAL PAIN

Odd Lindell



**Karolinska
Institutet**

Stockholm 2010

Supervisor: Lars-Erik Strender

Cover picture: 'The Thinker' by Auguste Rodin, a French sculptor, 1840–1917. The original is placed in Musée Rodin, Paris. This copy in Prins Eugens Waldermarsudde, Stockholm, was snapped by Odd Lindell.

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Universitetsservice US-AB

© Odd Lindell, 2010
ISBN 978-91-7457-178-3

“...Jag har sagt något litet och fattigt av det
som brunnit hos mig och så snart brinner ner,
men den kärlek, där fanns, ej förgängelse vet...”

“... I have said something slightly and poorly of things,
which were burning in me and so soon will burn down,
but the love that exists has perpetual wings...”

(From “Epilogue” by Dan Andersson, a Swedish poet, 1888–1920;
translation by Odd Lindell)

To my mother Inga-Britt; to Caroline, Axel and Gunvor

ABSTRACT

Background: Non-specific spinal pain (NSP), comprising back and/or neck pain, is one of the leading disorders behind long-term sick-listing. The general aim was to study the rehabilitation of non-acute (=leading to full-time sick-listing > 3 weeks) NSP as regards **epidemiology** ((Study) I), **reliability** (II), **treatment** (III), and **return-to-work prediction** (IV).

Specific aims: **I:** To compare living conditions associated with long-term sick-listing for NSP in patients with non-acute NSP with a population-based sample of non-patients. **II:** To answer the question “given a 10-test package of function tests for patients with non-acute NSP, could an examiner without formal medical education be used without loss of quality?” **III:** For patients with non-acute NSP, a programme of cognitive-behavioural rehabilitation was compared with traditional primary care. The specific aim was to answer the question “within an 18-month follow-up, will the outcomes differ in respect of sick-listing and number of health-care visits?” **IV:** For patients with non-acute NSP, to answer the question “which are the predictors at baseline for stable (= lasting ≥ 1 month) return-to-work during a 2-year period after baseline: objective variables from function tests, socioeconomic, subjective and/or treatment variables?”

Methods: **I** (cross-sectional study): For the 125 patients of study III, living conditions were compared with 338 non-patients by logistic regression. **II** (methodological study): Examination by a physiotherapist (A) in performing the 10-test package was compared with that by a research assistant (B) without formal medical education. The reliability, including inter- and intra-rater reliability, was assessed. In the inter-rater reliability study, 50 participants (30 patients + 20 healthy subjects) were tested once each by A and B. In the intra-rater reliability study, the 20 healthy subjects were tested twice by A or B. One-way ANOVA intra-class-correlation coefficient (ICC) was calculated. **III** (randomized controlled trial): After stratification by age (≤ 44 / ≥ 45 years) and subacute / chronic (= full-time sick-listed 3–12 / > 12 weeks) NSP, 125 primary-care patients were randomized to cognitive-behavioural rehabilitation (rehabilitation (rehab) group) or continued primary care (primary-care group). Outcomes: *Return-to-work share* (percentage) and *Return-to-work chance* (hazard ratios) over 18 months; *Net days* (crude sick-listing days x degree), and the number of *Visits* (to physicians, physiotherapists etc) over 18 months and the 3 component 6-month periods. Descriptive statistics, Cox regression and mixed-linear models were used. **IV** (prospective cohort study): *Stable return-to-work* was the outcome variable, the above-mentioned factors were the predictive variables in multiple-logistic regression models, one per follow-up at 6, 12, 18 and 24 months. The predictors which were represented in ≥ 3 follow-ups were finally considered.

Results: **I:** In the univariate analyses, 13 of the 18 living conditions had higher odds for the patients with a dominance of physical work strains and *Indication of alcohol over-consumption*, (odds ratio (OR)) 14.8 (95% CI)[3.2–67.6]. Five conditions remained in the multivariate model: *High physical workload*, 13.7 [5.9–32.2]; *Hectic work tempo*, 8.4 [2.5–28.3]; *Blue-collar job*, 4.5 [1.8–11.4]; *Obesity*, 3.5 [1.2–10.2]; and *Low education*, 2.7 [1.1–6.8]. **II:** All 5 tests requiring no manual fixation had acceptable reliability (ICC > 0.60 and no indication of systematic error). The 5 tests that required manual fixation had poor reliability except cervical rotation. The

difference (5 vs 1) was significant ($p = 0.01$). **III: All patients:** *Return-to-work share* and *Return-to-work chance* were equivalent between the groups. *Net days* and *Visits* were equivalent over 18 months but decreased significantly more rapidly for the rehab group over the 6-month periods ($p < 0.05$). **Subacute patients:** *Return-to-work share* was equivalent. *Return-to-work chance* was significantly greater for the rehab group (hazard ratio 3.5 [1.001–12.2]). *Net days* were equivalent over 18 months but decreased significantly more rapidly for the rehab group over the 6-month periods and there were 31 days fewer in the 3rd period. *Visits* showed similar though non-significant differences and there were half as many in the 3rd period. **Chronic patients:** *Return-to-work share*, *Return-to-work chance* and *Net days* were equivalent. *Visits* were equivalent over 18 months but tended to decrease more rapidly for the rehab group and there were half as many in the 3rd period (NS). **IV:** Three variables qualified: *Low total prior sick-listing* (including all diagnoses) was the strongest predictor in 2 follow-ups, 18 and 24 months, (OR) 4.8 [1.9–12.3] and 3.8 [1.6–8.7] respectively, *High self-prediction* (the patients' own belief in return-to-work) was the strongest at 12 months, 5.2 [1.5–17.5] and *Young age* (≤ 44) the 2nd strongest at 18 months, 3.5 [1.3–9.1].

Conclusions: Epidemiology: In the univariate analyses, the patients vs the non-patients had higher odds for most of the conditions. In the multivariate analysis, 5 conditions qualified, indicating work strains, lower social class and life-style. As these cross-sectional data makes causal conclusions impossible, they should be complemented by prospective research. **Reliability:** Given a 10-test package for patients with non-acute NSP, an examiner without formal medical education could be used without loss of quality, at least for the 5 tests that require no manual fixation. To make our results more generalizable, a similar study should be conducted with 2 or more examiners with and without formal medical education, and the intra-rater reliability study should also include patients and involve more participants. **Treatment:** Though the results were equivalent over 18 months, there were indications that cognitive-behavioural rehabilitation in the longer run might be superior to primary care. For subacute NSP, in terms of both sick-listing and health-care visits; for chronic NSP, in terms of health-care visits only. More conclusive results concerning this possible long-term effect might require a longer follow-up. **Return-to-work prediction:** The strong predictors of stable return-to-work were 2 socioeconomic variables (*Low total prior sick-listing* and *Young age*), and 1 subjective variable (*High self-prediction*). Objective variables from function tests and treatment variables were non-predictors.

Keywords: Non-specific; non-acute; subacute; chronic; spinal pain; back pain; neck pain; sick-listing; cross-sectional; methodological; randomized controlled; prospective cohort; epidemiology; reliability; treatment; return-to-work; prediction.

LIST OF PUBLICATIONS

This thesis is based on the following original articles, which will be referred to in the text by their Roman numbers.

- I Lindell O, Johansson SE, Strender LE: **Living conditions, including life style, in primary-care patients with non-acute, non-specific spinal pain compared with a population-based sample: a cross-sectional study.** *Clinical Epidemiology*. Accepted 2010-10-12.
- II Lindell O, Eriksson L, Strender LE: **The reliability of a 10-test package for patients with prolonged back and neck pain: could an examiner without formal medical education be used without loss of quality? A methodological study.** *BMC Musculoskelet Disord* 2007;8:31.
- III Lindell O, Johansson SE, Strender LE: **Subacute and chronic, non-specific back and neck pain: cognitive-behavioural rehabilitation versus primary care. A randomized controlled trial.** *BMC Musculoskelet Disord* 2008;9:172.
- IV Lindell O, Johansson SE, Strender LE: **Predictors of stable return-to-work in non-acute, non-specific spinal pain: low total prior sick-listing, high self-prediction and young age. A two-year prospective cohort study.** *BMC Fam Pract* 2010;11:53.

ABBREVIATIONS AND DEFINITIONS

Acute NSP	NSP leading to full-time sick-listing 0–21 days (3 weeks).
ANOVA	Analysis of variance.
Chronic NSP	NSP leading to full-time sick-listing > 12 weeks.
CI	Confidence interval. Usually the 95% CI is used. It indicates within which values the true value lies with a probability of 95%.
Epidemiology	The study of disease patterns in human populations.
Hazard ratio	A complex measure of the probability of changes over time compared between two groups.
ICC	Intra-class-correlation coefficient.
Multidisciplinary	A physician's consultation in addition to psychological, social or vocational intervention or a combination of these.
Net days	Days of sick-listing expressed in whole days (crude days X the degree). For example, sick-listing half-time 60 days = 30 net days.
Non-acute NSP	NSP leading to full-time sick-listing > 3 weeks
Non-specific	No need for specific treatment, i.e., treatment by hospital specialists, e.g., orthopedist or neurologist.
NSP	Non-specific spinal pain.
Odds	The probability of an event divided by the probability it does not occur.
OR	Odds ratio.
Odds ratio	The odds for one event in one group, divided by the odds for that event in another group.
Prevalence	The percentage of people in a known population who have the symptom (e.g., pain) during a specified period of time. Point prevalence concerns the day of the interview. Lifetime prevalence is the percentage of those who have the symptom at some times in their lives.
Rehabilitation	Any method by which people with a sickness or injury that interferes with their work ability can be returned to work. This can involve medical treatment as well as vocational measures as retraining etc.
Reliability	Acceptable reliability of an assessment method includes acceptable inter- and intra-rater reliability, i.e., it requires that the measurements are comparable when performed (a) on the same subject by numerous examiners and (b) on several occasions by the same examiner.
Sick-listing	Includes all form of work absence due to sickness, including disability pension.
SIO	Social Insurance Office (In Swedish: Försäkringskassan).

Spinal pain	Back and/or neck pain.
Subacute NSP	NSP leading to full-time sick-listing 42–84 days (12 weeks)
<i>p</i> -value	<i>p</i> = probability. A statistical significance often requires a <i>p</i> -value < 0.05 which means that the probability that the difference is by random is < 5%.

CONTENTS

List of publications	6
Abbreviations and definitions.....	7
Introduction.....	12
Pain.....	12
Acute pain.....	12
Chronic pain	12
Chronic pain and disability	12
Sick-listing in Sweden.....	13
Spinal pain	14
Prevalence of lumbar, thoracic and cervical pain.....	14
Back and neck.....	15
Prevalence of spinal pain	15
Cost to society	16
Non-specific spinal pain (NSP).....	17
Non-specific back pain.....	17
Non-specific neck pain.....	17
Sick-listing and NSP	18
Epidemiology	19
Reliability	19
Treatment.....	19
Return-to-work prediction.....	20
Our project	20
The clinical core	20
The scientific core	21
Aims	23
General aim.....	23
Specific aims.....	23
Study I.....	23
Study II	23
Study III	23
Study IV	23
Methods.....	24
Ethical approval	24
Study areas	24
Studies I, III and IV	24
Study II	24
Study I – a cross-sectional study	24
Subjects.....	24
Living conditions associated with long-term sick-listing for NSP	25
Outcome measure	27
Statistics	27
Study II – a methodological study	28
Setting and examiners	28
Subjects.....	28
The 10-test package.....	28

Examination procedure	31
Statistics.....	31
Study III – a randomized, controlled trial	32
Patients	32
Interventions.....	32
Outcome measures	35
Statistics.....	35
Data collection.....	36
Power calculation.....	36
Inclusion procedure.....	36
Premature cessation of recruitment	36
Follow-up	37
Study IV – a prospective cohort study.....	37
Patients	37
Outcome variable	37
Predictive variables	37
Statistics.....	44
Results.....	45
Study I.....	45
Patients and non-patients	45
Outcome	45
Study II	47
Patients and healthy subjects	47
Inter-rater reliability	47
Intra-rater reliability	48
Reliability	48
Study III.....	52
Response rate and analysis of missing data.....	52
Baseline characteristics and participant flow	53
Outcome measures	54
Interventions.....	57
Study IV.....	59
Loss to follow-up	59
Stable return-to-work	59
Predictors of Stable return-to-work	61
General discussion.....	64
Epidemiology	64
Work strains	64
Indicators of lower social class.....	64
Life style.....	65
Reliability	66
Treatment.....	66
Sick-listing.....	66
Health-care visits.....	67
Return-to-work prediction.....	68
Predictors in study IV compared with prior research	68
Non-predictors compared with prior research.....	68
Non-predictors that have previously been insufficiently studied	68

Objective versus subjective variables	69
Treatment as a predictor of return-to-work	69
Strengths of the project.....	69
General strengths	69
Specific strengths	69
Limitations of the project	70
General limitations	70
Specific limitations	70
Conclusions.....	71
Clinical implications.....	72
A personal reflection	72
Future research.....	73
Acknowledgements	74
Summary in Swedish / Sammanfattning på svenska	75
References.....	80
Study I.....	93
Study II.....	107
Study III	119
Study IV	137

INTRODUCTION

PAIN

Pain is the most common symptom for which patients seek health care [174]. Its complexity is mirrored in the definition: “Pain is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage” [129]

Acute pain

Because it is a powerful stressor, acute pain very effectively drives us to behaviours aimed at protecting the injured area so the tissues might heal, and at preventing new damage. To fight or fly from the pain stimulus, stress hormones are released, leading to increased heart rate, more rapid breathing and sweating palms. These changes are also characteristic of anxiety: acute pain and anxiety are closely linked [174]. Pain signals from the periphery are constantly modulated within the computer-like network of the central nervous system [115]. Pain, emotions and behaviour are integrated and work *in both directions*: pain might change behaviours and behaviours might change the pain [134].

Chronic pain

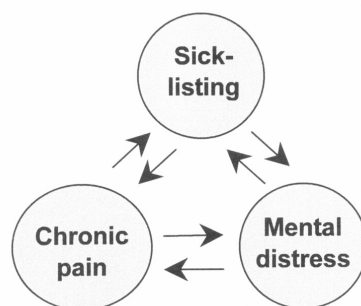
Chronic indicates a duration of at least 3 months [174], but is far from just acute pain with a prolonged duration. It might include plastic and partially irreversible changes in the pain tracts and the peripheral tissues [29]. When it is established, the chances of total pain relief are very small [132]. In a long-term study, 85% of subjects with chronic pain after 12 years were still suffering [5]. Chronic pain loses its biological meaning and becomes counterproductive. The activity-driving response is replaced by passivity, hopelessness and withdrawal from social activities. This is also characteristic of depression, which is often linked to chronic pain [174].

Chronic pain and disability

Pain and disability often go together, but are not the same. Pain is a symptom. *Disability* is restricted activity. Some patients manage to lead surprisingly normal lives despite severe pain, while ordinary backache may totally and permanently disable other patients. Many patients (and doctors!) assume it is simply a question of pain causing disability, and if we treat the pain the disability will disappear. Too often, that just does not work. The connections between pain and disability are complex [174]. Mental distress is an equal co-actor. A large, cross-national study showed that pain predicted mental distress and mental distress predicted pain. Disability, expressed as work absence, was the strongest predictor for both chronic pain and mental distress [46]. In the long run, they all interact: increased pain worsens the inability to work, which worsens mental distress, leading to increased pain, inability to work, etc [176]. A prerequisite for support during sick absence is *sick-listing*, including all kinds of sick

absence, including disability pensions. Figure 1 shows the co-acting. The problem of passive and counterproductive sick-listing was noticed early [161,170]. Sweden is no exception.

Figure 1. Sick-listing, chronic pain and mental distress.



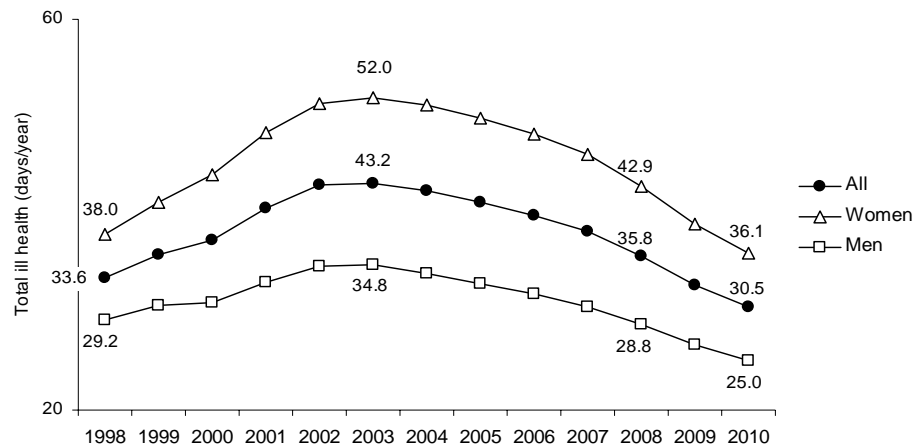
Sick-listing in Sweden

In Sweden, publicly provided, tax-financed social insurance compensates loss of income due to sickness. Sick-listing includes absenteeism with sickness or rehabilitation benefit, temporary disability pension and disability pension (the temporary form was abolished in 2008). For sick-listing > 7 calendar days, a doctor's certificate is required with a detailed description of symptoms and signs and a recommendation of the degree, 0.25, 0.50, 0.75 or 1.00 (= full-time), and duration of sick-listing. The ultimate approval or disapproval of benefits are made by the Social Insurance Agency (SIO). The employer has the financial responsibility for the 2 initial weeks, SIO for sick-listing > 14 days. The main rehabilitation actors (except the patient) are the employers (in cases of unemployment, the Public Employment Services), health care and SIO. *Rehabilitation* is any method by which people with a sickness or injury that interferes with their work ability can be returned to work. This can involve medical treatment as well as vocational measures as retraining, etc [178]. Since 1992 the employer has had the responsibility for noticing if the sick-listed employee needs rehabilitation and to take appropriate measures, and SIO has had the responsibility for the comprehensive coordination of the rehabilitation [9].

Since the late 1980s the need of better coordination has been indicated in several official reports [33,44,84] and studies [109,110,141,142,71,35,111,7,143,8,9,90]. However, the sick-listing rose precipitously after 1998 with a paradoxical peak in 2003, when the Swedes were the most sick-listed people in Europe, probably the world, but simultaneously one of the healthiest with a mean life expectancy at birth of 81.0 years vs 79.1 in comparable countries [57] and 66.1 in the world [91] (Figure 2). The Swedish National Audit Office (In Swedish: Riksrevisionen) published a crushingly negative report of SIO in 2004 [155]. The failure of gate-keeping was obvious: it was far too easy to enter a passive and costly sick-listing [151]. A governmental project was initiated in 2002 to halve sick-listing [40], focussing on increased restrictiveness. Thereafter it decreased to the European mean [10]. In 2008 the sick-listing rules

became very restrictive: “After 90 days you have the right (to sickness benefit) exclusively if you can’t do any work for your employer. After 180 days... exclusively if you can’t do any work at all...” [145]

Figure 2. Total ill health* in Sweden. The start of the increase, the peak, the change in sick-listing rules, and the current situation are indicated; 2010 is 31 August [3], otherwise 31 December [135].



* Includes all sick-listing. E.g., in a total ill health of 30 days, the absenteeism is 1 of a possible 12 months for the entire work force. One day of total ill health \approx 6 million sick-listing days [3] \approx a loss of production of 24 billion SEK (2.6 billion €) (from data provided by Assistant Professor Paula Liukkonen, 2010-11-04).

Diagnoses behind sick-listing

For many years, *musculoskeletal disorders* dominated sick-listing in Sweden. However, following international trends [78], it has been outflanked by *mental disorders* since 2005. Those two groups together constitute 7 out of every 10 new disability pensions [120]. The most common diagnoses among mental disorders are depression, stress-related disorders and anxiety. The most common musculoskeletal disorder is spinal pain.

SPINAL PAIN

Most of the chronic-pain cases suffer from spinal pain [46]. *Spinal pain* is pain arising from various parts of the spine, i.e., the lumbar, thoracic and/or cervical spine [108].

Prevalence of lumbar, thoracic and cervical pain

Prevalence is the percentage of people in a known population who have the symptom during a specified period of time. *Point prevalence* concerns the day of the interview. *Lifetime prevalence* is the percentage of those who have pain at some times in their lives [119].

Lumbar pain (usually denoted ‘low back pain’) is most common, followed by cervical pain; thoracic pain is least common. Their lifetime prevalences, according to a large Danish study, are 57%, 40% and 17%, respectively [89].

Back and neck

There is a widely used dichotomy of spinal pain between *back pain* and *neck pain* [117,22,53,51,179]. *Back* includes the gluteal regions [130,175,55] and *neck* the shoulders and the upper parts of the arms [74,73,55]. Concerning thoracic pain, we have found no clear delimitation. We have chosen to define *back pain* as pain from the lower half of the thoracic spine and downwards and *neck pain* as pain from the upper half of the thoracic spine and upwards (Figure 3). This is consistent with the predominance of cervical and lumbar pain and its widespread and radiating image [162,45,179], and is suitable in clinical practice.

Prevalence of spinal pain

Lifetime prevalence

According to different studies, the lifetime prevalence of spinal pain is 54–80% [108].

Women are overrepresented

The clearly elevated rate of pain, including spinal pain, in women relative to men has been reported in many studies [107,4,168,45].

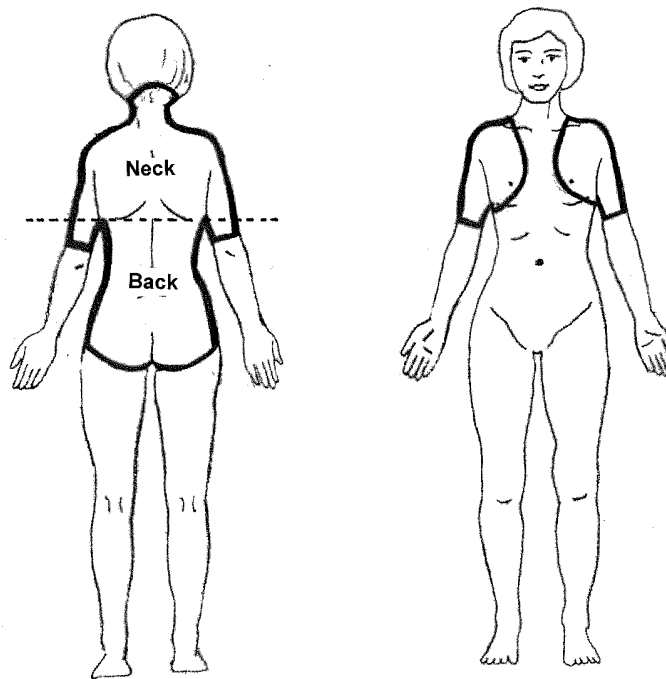
A dramatic increase with a possible peak

A comparison of two large British cross-sectional studies, from the middle of the 1950s and the 1990s, showed an up to 4-fold increase in spinal pain, particularly among women [55]. The peak has possibly passed, at least in Sweden. From 1980/81 to 2000/01 there was an increase, especially in women, and then a decrease, especially among men (Table 1).

Table 1. Point prevalence of severe musculoskeletal pain (mostly spinal pain) in Sweden 1980–2006 among patients aged 16–84 years. Peak value in bold text. From the ULF surveys [101].

Years	1980–81	1988–89	1994–95	2000–01	2004–05	2006
Women	19.8	21.2	23.0	25.4	23.5	22.1
Men	16.7	15.6	16.6	18.0	15.2	14.3

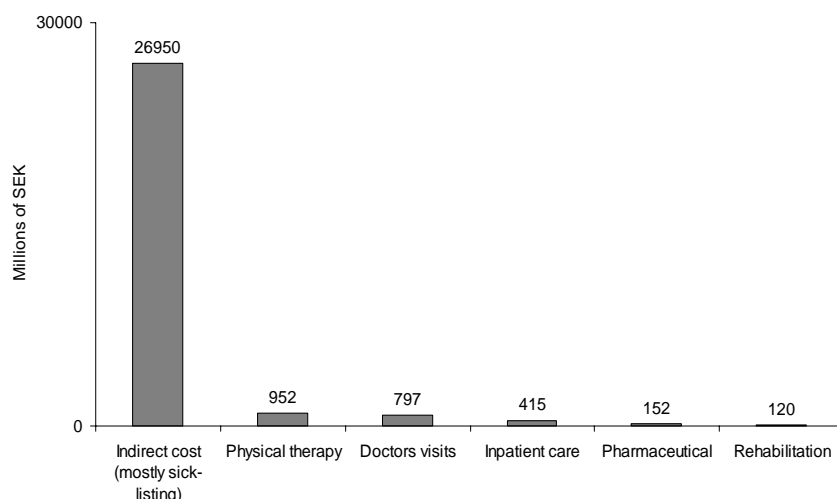
Figure 3. The components of spinal pain. Back pain is pain in the area below an imaginary line connecting the lower tips of the shoulder blades. Neck pain is pain in the area on and above this line. The line crosses the middle of the thoracic spine [133].



Cost to society

In addition to individual suffering, spinal pain causes great societal costs. In Sweden in 1995 the total cost was estimated at 29.4 billion SEK = 3.0 billion € (euro) yearly or 1.7% of the Gross National Product [124], nearly double the cost of health and social care [156] (Figure 4). The overwhelming portion, 91.7%, comprised the indirect costs dominated by sick-listing, in particular on a long-term basis [124]. The lowest portion, 0.4%, was for rehabilitation. No national cost analysis for spinal pain is available for Sweden after 1995. However, as the sick-listing of today is about on the same level as in 1995 [10], the cost should be approximately the same. In 1995 the expenditure on rehabilitation for spinal pain was extremely low. However, a Swedish government report of 2006 showed a *further decrease* in rehabilitation measures [9]. The vast majority of cases of spinal pain are non-specific.

Figure 4. The societal cost of spinal pain.



NON-SPECIFIC SPINAL PAIN (NSP)

NSP constitutes > 95% of spinal pain. We define *NSP* as spinal pain with no need for specific treatment, i.e., treatment by hospital specialists.

Non-specific back pain

Among patients with back pain, < 1% have serious spinal diseases such as tumours or infections (e.g., tuberculosis) and need urgent treatment by oncologists, etc; < 1% have inflammatory disorders (e.g., Bechterew's disease) and require rheumatological management; < 5% have a symptom-giving slipped disc, but only 1 in 10 of those needs surgery (the vast majority are cured spontaneously) [173]. The rest are cases for primary care or no professional treatment at all.

Non-specific neck pain

There has been less research on neck pain than back pain [74,167], but it has increased substantially during the last decade. Out of every 100,000 individuals, up to 20,000 will experience neck pain during the coming year, 8,000 will seek care for it, 6 will have neck pain with neurological manifestations and < 10 a serious instability, spinal infection etc [47]. Thus, the overwhelming majority of neck pain is also non-specific.

Post-traumatic neck pain

Neck pain after traumas, especially Whiplash-Associated Disorders (WAD), is between specific and non-specific. Eighty percent of WAD cases are associated with car

accidents [146]. In Sweden every year, about 30,000 individuals (< 0.5% of the general population) report symptoms, mostly neck pain, after road accidents. The large majority recover within a couple of months, but 18–60% have persistent symptoms 6 months after the accident; 5–8% suffer substantial impact on work capacity and need more profound treatment. The guidelines for treating prolonged WAD are practically identical to those for management of chronic, musculoskeletal pain in general, including NSP [180]. However, as our project also concerned non-chronic NSP, patients with WAD were excluded.

Sick-listing and NSP

The great majority with disabling NSP recover quickly [26]. Around 50% have returned to work after full-time sick-listing for one week and 90% after 12 weeks. Thereafter, however, the recovery speed evidently levels off. As time off work has such a devastating impact, we categorized disabling NSP from the time period of full-time sick-listing [176] (Figure 5).

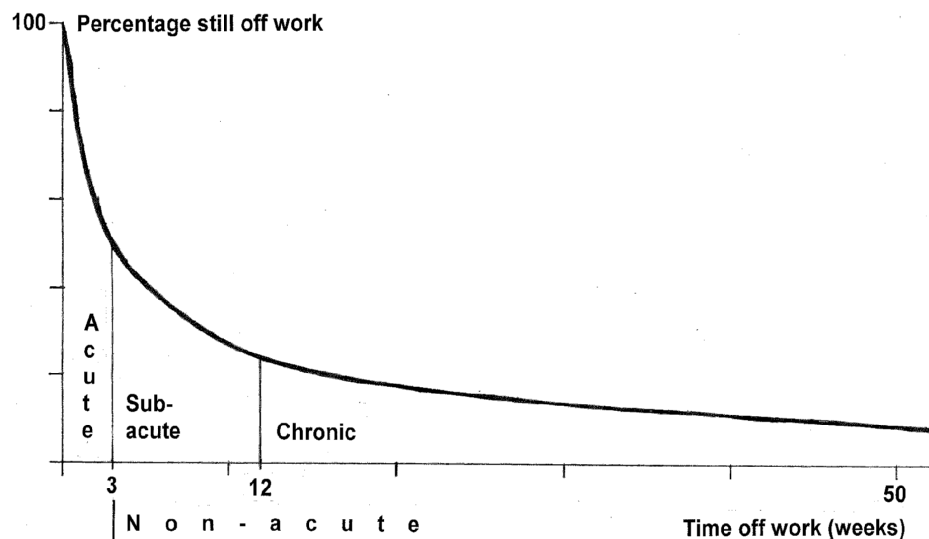
Acute NSP and non-acute NSP

Pain leading to full-time sick-listing 0–21 days (3 weeks) and >3 weeks, respectively.

Sub-acute NSP and chronic NSP

Pain leading to full-time sick-listing 22–84 days (3–12 weeks) and > 12 weeks (3 months), respectively.

Figure 5. The course of disabling NSP. Modified after Waddell [176].



Epidemiology

Clinical guidelines emphasize the necessity of early intervention to prevent long-term sick-listing caused by NSP [176]. This requires the identification of patients at risk. Socio-economic and medical factors are associated both with the onset of acute NSP and the progression to non-acute NSP [117,171,96,167]. Sweden has a unique tradition of keeping population statistics, starting in 1749, earliest in the world. This provides an exceptional opportunity for epidemiological research. However, the research within the area has been seriously limited with, e.g., an under-representation of women [53].

Reliability

Acceptable *reliability* of an assessment method includes acceptable inter- and intra-rater reliability, i.e., it requires that the measurements are comparable when performed (a) on the same subject by numerous examiners and (b) on several occasions by the same examiner [102]. In the rehabilitation of patients with non-acute NSP, it is necessary to assess the *physical impairment*, i.e., the pathological, anatomical or physiological abnormality of structure or function leading to loss of normal ability [169]. As these patients suffer from non-specific pain, the focus is on dysfunction [173]. The assessment is made by function tests, i.e., tests in which the patient performs some kind of physical activity [103].

Despite an immense amount of research, no gold standard has been established for which function tests to use for which patients for the assessment of NSP [103]. Several prior studies also have elucidated the problem of achieving agreement between different examiners [19,152,37]. For example, in an 8-test package for patients with NSP, only 1 test had acceptable reliability [63]. In some reliability studies, chiropractors [56], naprapaths [15] or physicians [17,15,113,152,121] have been represented. The vast majority of reliability studies, however, have been performed with physiotherapists as examiners [102,25,136,65,63,75,19,37].

Treatment

Acute NSP is managed by continuing ordinary activities as normally as possible, and manual therapy, if necessary. *Manual therapy* includes manipulation, mobilisation and stabilizing training [43]. In cases of non-acute NSP, multidisciplinary rehabilitation should be considered [177]. *Multidisciplinary* treatment includes a physician's consultation in addition to psychological, social or vocational intervention or a combination of these [74]. In the treatment of non-acute NSP return-to-work is crucial [178]. Another important issue is the health-care utilization needed to achieve certain treatment results. In that respect, a frequently-used outcome measure is the number of health-care visits [98,100].

The 3 key components of successful rehabilitation programmes for NSP are: reactivation and progressive increase in activity levels, addressing dysfunctional beliefs and behaviour by a cognitive-behavioural therapeutic approach, and occupational interventions [178].

Cognitive-behavioural therapy for pain has been developed to be an integral part of rehabilitation programs. It was conceptualized as a way of enhancing treatment by addressing pertinent cognitive aspects, e.g., negative emotions and thoughts, and

behavioural aspects, e.g., altered activity and medication taking. In addition, the cognitive-behavioural approach offers an educational concept, whereby learning encompasses the entire rehabilitation process [97]. To sum up: ‘cognitive’ is *doing by learning*, ‘behavioural’ is *learning by doing*.

Concerning back pain, programmes including these items have shown good results in several studies [92,68,50,105,60]. Randomized controlled trials have concerned patients with subacute back pain only [92,68,105,60,100,150,6], mixed groups with subacute or chronic back pain [50,72] or patients with chronic back pain only [128]. There is a serious lack of evidence concerning the rehabilitation of neck pain [74].

The high frequency of relapses after rehabilitation of NSP is associated with inadequate follow-ups. A short program might fail to achieve long-standing behavioural changes [161]. In the 1990s the vast majority of rehabilitation programs in Sweden were comparatively short, with a fixed duration averaging 6 weeks [109].

Primary care is the appropriate source of treatment for NSP [177]. In Sweden, however, notwithstanding clinical guidelines, only a small minority of individuals with non-acute NSP receive multidisciplinary rehabilitation [142].

Return-to-work prediction

Cost-effectiveness in allocating treatment resources requires predictors of return-to-work to be collected by means of both questionnaires and function tests. While the former are cheap, the latter require substantial personnel resources.

The cross-sectional design of most population statistics makes conclusions about causes and effects impossible. For example, anxiety, depression and low physical activity could be both explanatory and responding variables for non-acute NSP [118,167]. Thus prediction of return-to-work requires prospective data. As return-to-work is often followed by recurrences of work absence, longitudinal data are required to denote a stable return-to-work, i.e., data that is collected from several time points [5].

OUR PROJECT

The clinical core

The clinical core of the project was a rehabilitation centre for patients with non-acute NSP (the STRONG unit). The centre operated in Haninge, a semi-urban district 25 kilometres south-east of Stockholm city, during 1991–2006. From 1996 the centre used a cognitive-behavioural programme with the aim of achieving the maximal degree of work ability lasting for at least 30 consecutive days; possible relapses were met by individual and, when needed, long rehabilitation periods. *Work ability* was inversely proportional to sick-listing, which is the definition used by SIO. Work abilities of 1.00 (= full-time), 0.75, 0.50 and 0.25 corresponded to sick-listings of 0, 0.25, 0.50 and 0.75, respectively. Zero work ability = full-time sick-listing. The physiotherapists at the rehabilitation centre used a package of 10 function tests. Most of the tests had been validated in previous studies by comparing the results obtained by medically trained examiners [17,106,114,93,25,65,17,113,93].

For the first 10 years the rehabilitation centre was run within Stockholm County Council. From 2002 it operated as a private company and the number of rehabilitation

teams was decreased from 4 to 1, comprising 4 team members: a physician (Odd Lindell), a physiotherapist trained in manual therapy, a psychologist or a social worker trained in cognitive-behavioural therapy, and a health-care adviser.

In 2006 the centre was closed down as a result of decreasing demands for rehabilitation services from SIO and employers.

The scientific core

We wanted to elucidate different aspects on the rehabilitation of patients with non-acute NSP. The starting points were our rehabilitation programme and the then-existing need of further research. The scientific core was a randomized controlled trial, running 2000–2006, with the objective of comparing the programme with traditional primary care (study III). Data from study III were re-assessed in a cross-sectional study (study I) and in a prospective cohort study (study IV). In study II the reliability of the 10-test package, which were used in studies III and IV, was evaluated.

Epidemiology (study I)

We found no previous study where primary-care patients with non-acute NSP were compared with a population-based sample. The representation of women in our patient sample was satisfactory.

Reliability (study II)

At the time of inclusion and 1 year later, each patient in the randomized controlled trial (study III) met a research assistant at that patient's health centre. Among other items, the patients performed the 10-test package. For practical and economic reasons it was appropriate for the person who administrated the study and visited the different health centres also to execute the tests. Although the research assistant had no formal medical education, this seemed reasonable, since the tests were standardized and easy to perform. However, we found no study of reliability in which examiners without formal medical education were engaged. Still, the evaluation of rehabilitation efforts might be less biased if performed by personnel standing outside the treatment work itself. It seemed economically unrealistic for ordinary clinics to keep medically-trained personnel only for assessment tasks. Therefore, if medically untrained examiners could be used without decreased quality, this might produce a better assessment of outcome at defensible cost and could also be useful in a research context.

Treatment (study III)

We found no previous randomized controlled trial in which the same rehabilitation programme was offered to patients who were stratified by subacute and chronic NSP. This might be interesting as the subacute phase implies that most of the spontaneous recovery in NSP has passed, but still the rehabilitation potential is good. In the chronic phase the potential often is substantially decreased [176].

Return-to-work prediction (study IV)

In a large 2004 review, as far as we know the hitherto most extensive work in its genre, a total of 133 possible risk factors for sick-listing for spinal pain was elucidated. Although the review indicated some factors, the vast majority of the investigated factors proved to be insufficiently studied. The review concluded that, despite the large

consequences of sick-listing in NSP, the research within this area is surprisingly meagre with, among other scarcities, to few longitudinal studies [53].

AIMS

GENERAL AIM

The general aim of this project was to elucidate different aspects of the rehabilitation of patients with non-acute NSP: epidemiology (study I), reliability (study II), treatment (study III) and return-to-work prediction (study IV).

SPECIFIC AIMS

Study I

The aim of study I was to compare living conditions associated with long-term sick-listing for NSP in patients with non-acute NSP, with a non-patient population-based sample.

Study II

The aim of study II was to answer the question “given a 10-test package for patients with non-acute NSP, could an examiner without formal medical education be used without loss of quality?”

Study III

For patients with non-acute NSP, a programme of cognitive-behavioural rehabilitation was compared with traditional primary care. The aim of study III was to answer the question “within an 18-month follow-up, will the outcomes differ in respect of sick-listing and number of health-care visits?”

Study IV

Patients in study III completed the 10-test package and a questionnaire at baseline. The aim of study IV was to answer the question “which are the predictors at baseline for stable return-to-work during a 2-year period after baseline: objective variables from function tests, socioeconomic, subjective and/or treatment variables?”

METHODS

ETHICAL APPROVAL

Approval was given by The Research Ethics Committee, Karolinska University Hospital, Huddinge. Studies I, III and IV: Dnr 170/99; complements 2000-05-29 and 2000-08-02. Study II: Dnr 443/00.

STUDY AREAS

Studies I, III and IV

The study area was the Southern part of Stockholm County, including 5 urban districts (Enskede-Årsta-Vantör, Farsta, Älvsjö, Skarpnäck and Hägersten-Liljeholmen) and 4 semi-urban districts (Huddinge, Nynäshamn, Tyresö and Haninge). The number of inhabitants (31 December 2001) in the county totalled about 1,830,000, of whom 1,100,100 were of the same age as the patients of these studies (18–59 years). The study area had about 467,000 inhabitants, of whom 281,000 were aged 18–59 years and constituted the source population.

Study II

The study area was Haninge, geographically near the middle of the study area of studies I, III and IV, and with about 70,000 inhabitants.

STUDY I – A CROSS-SECTIONAL STUDY

Subjects

Patients

One hundred and twenty-five patients with non-acute NSP, between August 2000 and January 2004, were included in a randomized controlled trial (study III). For these patients, a cross-sectional study was carried out with baseline data.

The criteria for inclusion: 1. Vocationally active, up to and including 59 years of age. 2. Sick-listed full-time for NSP for at least 6 weeks (42 days) and for at most 2 years (730 days). 3. Able to fill in forms. *The criteria for exclusion:* Temporary disability pension, or disability pension being paid or in preparation. 2. A primary need for action by a hospital specialist (e.g., operation for intra-vertebral slipped disc). 3. Pregnancy and diseases (other than NSP) that would probably make rehabilitation impracticable (e.g., advanced pulmonary disease). 4. Whiplash associated disorders as a primary obstacle to working. 5. Previous rehabilitation at the rehabilitation centre. 6. Other multidisciplinary rehabilitation ongoing or planned.

One of the 125 patients failed to complete the questionnaire and was excluded. The remaining 124 patients were included in this study. They were recruited by 41 family

doctors at 13 primary-care health centres in a non-systematic way, i.e., dependent on the motivation and available time of the family doctor. To ensure that all the study patients, including those who were allocated to continued primary care, received a high minimum level of treatment, only permanently employed or long-term substitute doctors were engaged. For the patients who fulfilled the criteria, the family doctor gave verbal and written information. Each patient who gave the consent to participate was interviewed by telephone by a research assistant within 2 days. The patients who still qualified for the study saw the assistant at the health centre within 5 days. After signing an informed consent to participate, the patient with support of the assistant, completed a questionnaire of baseline data.

Non-patients

Statistics Sweden conducts The Survey of Living Conditions annually (In Swedish: Undersökningarna av levnadsförhållanden (ULF))[101,154]. To reach an acceptable power, 2 years of ULF data, 2000 + 2001, were combined. Most of the patients (81/124) were recruited during that period. ULF 2000/2001 was a simple, random sample of 7,465/7,459 individuals, aged 16–84 years. They were invited to be interviewed in their homes. Non-responders and those who declared that they did not want to be visited were offered a telephone interview. From the interviewed subjects we selected those of the same age as the patients except those with disability pensions. This resulted in a nationwide sample, of which 371 subjects were living in the study area. By exclusion of the vocationally inactive (e.g., students and housewives) and the full-time sick-listed subjects, a comparison group of 338 non-patients was achieved.

Living conditions associated with long-term sick-listing for NSP

The cross-sectional design made conclusions about causes and effects impossible. We therefore limited the analyses to living conditions that could reasonably be supposed to have existed before the start of the current sick-listing and excluded comparisons of, e.g., anxiety, depression, pain and exercise habits. For a majority of the living conditions, the questions in the patient and the ULF questionnaires were identical or nearly identical. As regards the non-identical questions, we made modifications so they were reasonably comparable. Questions concerning alcohol consumption were put only to the ULF subjects of 2001, of whom 169 belonged to the non-patients. Questions regarding work conditions were put exclusively to the 325 non-patients in employment. The questions concerning the other living conditions were put to all non-patients. The 18 living conditions are shown in Table 2.

Table 2. Living conditions. Univariate analyses. One hundred and twenty-four patients with non-acute NSP compared with 338 non-patients by logistic regression, adjusted for gender and age. If not otherwise stated, results are shown as number (in case of missing data, the total number is also shown) with percentage in parenthesis; 95% CI within brackets.

	Patients (n=124)	Non-patients (n=338)	Odds ratio	p-value
<i>Woman</i> [52,49,75,32]	68 (54.8 [46.0–63.7])	161 (47.6 [42.3–53.0])	1.3 [0.9–2.0]	NS
<i>Older age</i> (= ≥45 years) [80,28]	57 (46.0 [37.1–54.9])	107 (31.7 [26.7–3.6])	1.8 [1.2–2.8]	0.006
<i>Immigrant</i> (= born outside Sweden) [22]	34 (27.4 [19.5–35.4])	43 (12.7 [9.2–16.3])	2.6 [1.6–4.4]	<0.001
<i>Single life</i> (= living alone without children) [112]	22 (17.7 [10.9–24.6])	101 (29.9 [25.0–34.8])	0.5 [0.3–0.9]	0.02
<i>Living with children at home</i> [112]	69 (55.7 [46.8–64.5])	167 (49.4 [44.1–54.8])	1.3 [0.9–2.0]	NS
<i>Low education</i> (= at most junior high school) [182]	44 (35.5 [26.9–44.0])	41 (12.1 [8.6–15.6])	3.8 [2.3–6.3]	<0.001
<i>Unemployed</i> [143]	29 (23.4 [15.8–30.9])	13 (3.9 [1.8–5.9])	8.2 [4.0–16.5]	<0.001
<i>Blue-collar job</i> ^{a,b} [171,94]	83 (87.4 [80.6–94.2])		15.0 [7.7–29.1]	<0.001
Physical work strains^a: [94]				
<i>High physical workload</i> [52]	79 (83.2 [75.5–90.8])	51/325 (15.7 [11.7–19.7])	30.4 [15.9–58.3]	<0.001
<i>Monotonous work tasks</i> [52]	61 (64.2 [54.4–74.0])	134/324 (41.4 [36.0–46.7])	2.7 [1.7–4.3]	<0.001
<i>Difficult work postures</i> [52]	76 (80.0 [71.8–88.2])	107/324 (33.0 [27.9–38.2])	9.0 [5.1–15.9]	<0.001
<i>Vibrations in work</i> [159]	35 (36.8 [27.0–46.7])	15/324 (4.6 [2.3–6.9])	18.6 [8.7–39.9]	<0.001
Psychosocial work strains^a: [157]				
<i>Hectic work tempo</i> [62]	88 (92.6 [87.3–98.0])	239/324 (73.8 [68.9–78.6])	4.5 [2.0–10.1]	<0.001
<i>Low decision latitude</i> [52]	30 (31.6 [22.1–41.1])	42/321 (13.1 [9.4–16.8])	3.2 [1.8–5.5]	<0.001

	Patients (n=124)	Non-patients (n=338)	Odds ratio	p-value
<i>Smoking</i> (daily + not daily) [80]	49 (39.5 [30.8–48.2])	118/336 (35.1 [30.0–40.2])		NS
<i>Indication of alcohol over-consumption</i> ^c [30]	17 (13.7 [7.6–19.8])	2/164 (1.2 [-0.0–2.9])	14.8 [3.2–67.6]	0.001
<i>Obesity</i> (= BMI \geq 30 [79]) [58]	30 (24.2 [16.6–31.8])	23/332 (6.9 [4.2–9.7])	4.3 [2.3–7.7]	<0.001
<i>Comorbidity</i> ^d [123]	45 (36.3 [27.7–44.9])	105 (31.1 [26.1–36.0])	1.1 [0.7–1.7]	NS

^a Concerning the subjects in employment: 95/124 patients and 325/338 non-patients.

^b According to Socio-Economic Classification (In Swedish “Socioekonomisk indelning (SEI)”) [http://www.scb.se/statistik/LE/LE0101/_dokument/SEIstandard.pdf]. **Modification:** the subjects in the group “Entrepreneur” were considered *Blue-collar job* starting from their probable level of education.

^c The alcohol questions were put to 169/338 non-patients.

^d Any other prolonged disease except NSP and obesity.

Outcome measure

As the outcome variable of logistic regression, being either a patient or a non-patient.

Statistics

We first estimated the distribution of the living conditions for the patients and the non-patients. The results are shown as proportions with 95% confidence intervals (CI). Differences between the groups were evaluated by univariate-logistic regression [64], adjusted for gender and age divided in 2 classes: *Old age* ≥ 45 and *Young age* ≤ 44 years. The dependent variable was the sample class, i.e., patient or non-patient. The predictive variable was the living condition. The results are presented with odds ratios (OR), 95% CI and *p*-values. To find the most discriminative living conditions we used multiple-logistic regression, adjusted for gender and age, with the sample class as the outcome variable and the living conditions as the explanatory variables. Subjects with missing data were excluded. This left 249 subjects (95 patients and 154 non-patients) for this analysis. We first explored univariate analyses. The variables with a *p*-value of at most 0.10 are presented with OR, *p*-values and 95% CI. They were included in a multiple model, from which the variables with *p*-values ≥ 0.05 were excluded stepwise to yield a model comprising only variables with *p*-values < 0.05 . The final multivariate model is presented with OR, *p*-values, 95% CI, a goodness-of-fit test by Hosmer-Lemeshow, the percentage of correctly predicted patients, and the area under the ROC (receiver operating characteristic) curve [64]. Stata, version 10.1, was used for the analyses [149].

STUDY II – A METHODOLOGICAL STUDY

Setting and examiners

The study was performed at the rehabilitation centre and a physiotherapy centre situated next door. In appraising the assessment work of a medically untrained examiner it seemed logical to use an experienced physiotherapist as the gold standard. Examiner A (Lars Eriksson) had the highest Swedish degree in orthopaedic manual therapy and had been working as a physiotherapist for 10 years. Examiner B had a Bachelor of Arts in psychology but no formal medical education. She had been working as a research assistant with purely administrative tasks for 2½ years and had no previous vocational experience of manual contact with patients. B was prepared for study II by 4 hours' training in the performance of the 10-test package and practising the package during the autumn of 2000 on barely 40 patients who were included in study III.

Subjects

Fifty participants were included and gave their written consent to participate in the study: 30 patients with prolonged (≥ 4 weeks) NSP, and 20 healthy subjects.

Patients

From March up to and including August 2001, a total of 30 patients were recruited at the physiotherapy centre in Haninge. They were supplied with verbal and written information. Thirty-one consecutive patients fulfilling the criteria were asked to participate. All but 1 agreed.

Inclusion criteria: 1. NSP for > 4 weeks. 2. The patient was considered able to execute the whole 10-test package.

Exclusion criteria: 1. Such severe pain or dysfunction that it might be harmful for the patient to participate. 2. Whiplash-associated disorders. 3. Inability to read the written information.

Healthy subjects

From February up to and including August 2001, 20 healthy subjects were recruited among the staff at the rehabilitation and the physiotherapy centre. Twenty staff members (physiotherapists, physicians and receptionists) were asked consecutively and all of them agreed to participate.

The 10-test package

Four tests included motion in one direction only. Four tests comprised motion to the right and to the left, and 1 involved motion forward and backward. A lifting test included a lumbar and a cervical sub-test. This resulted in 10 tests composed of 16 sub-tests.

Five of the 10 tests required that the examiner kept a firm hold against the foundation of those parts of the participant's body that were not supposed to move during the test. This manual fixation was done to eliminate misleading co-movements from those parts.

The package followed the protocol of previous studies, with some modifications. The total examination time of the package was approximately 30 minutes. A detailed description:

1. Forward bending

The participant (P) stood barefoot with the heels together. P bent forward, keeping the knees straight and with the arms straightened out downwards the floor. When P had bent maximally, the examiner (E) measured the distance between the middle-finger tip and the floor, to within 1 cm, with a wooden stick. If the floor was reached, the distance was noted as 0 cm [17].

2. Modified Schober

P stood with the feet together. Three dots were marked: dot a between the lowest lumbar spinal process and sacrum, dot b 10 cm above and dot c 5 cm beneath a. P bent forward, keeping the knees straight. The distance b–c when P was bent maximally forward was measured with a tape to within 1 cm. The difference of b–c when maximally bent forward and standing was noted. Normally, b–c increases by at least 5 cm [106].

3. Lateral bending (right/left)

P stood with 20 cm between the feet and with the back, neck, back of the head and shoulders against a wall and the arms loosely against the sides of the body. The middle-finger tip positions on the outside of the thighs were marked with dot a. P bent to the right side, keeping the knees straight and without losing contact between the shoulders and the wall. In the maximally bent position, the middle-finger tip position on the right thigh was marked by dot b. The same procedure was performed on the left side. The distances a–b on the right and left thighs were measured with a tape to within 1 cm [114].

4. Trunk rotation (right/left)

P sat on a stool with the knees together holding a rod horizontally in the frontal plane across the upper sternum and the front of the deltoid muscles. From the ends of the rod, a line with a plumb weight hung down pointing at a semicircular protractor lying on the floor under and in front of P. In the initial position, the base line of the protractor was in the same frontal plane as the rod and the middle of the base line was directly below the middle of the rod. E stood behind P holding the lower part of Ps body still by firmly pressing the iliac crests down towards the seat of the stool. P rotated the trunk maximally to the right. The maximally rotated position was read, to within 5 degrees, where the plumb weight pointed at the protractor. The same procedure was performed on the left side [93].

5. Active-straight-leg raise (right/left)

P was lying supine on a couch with the knees straight. An MIE meter was placed on the lower part of the right leg at the tuberositas tibiae. While the left leg was held in its initial position by E, P raised the right leg, keeping the knee straight. When the leg was maximally raised, the angle between the leg and the horizontal plane was read to within 1 degree. The same procedure was performed with the right leg fixed to the couch and the left leg raised [93].

6. Cervical bending (forward/backward)

P sat on a chair with the head in a neutral position. A CROM meter was placed on the head. E held Ps thoracic and lumbar spine fixed to the back support of the chair. P bent the head forward and then backward. In the maximally bent positions, the angle between the head and the vertical line was read to within 1 degree [25].

7. Cervical rotation (right/left)

The same procedure as in test 6, except that P rotated the head to the right and then to the left. The angle between the head in neutral and in maximally rotated position was read to within 1 degree [25].

8. Abdominal endurance

P was lying supine on a couch with the knees bent at 90°, the soles of the feet on the couch and the palms resting on the front of the thighs. P performed a sit-up, with the fingertips touching the upper part of the patellae, and sustained this position as long as possible. The maximal sit-up time, until the fingers lost contact with the patellae, was measured with a stop-watch to within 1 second [65].

9. Modified Biering-Sørensen

P was lying prone with the lower part of the body, from the upper part of the iliac crest downwards, placed on a couch. The upper part of the body hung down from the short side of the couch, resting on the seat of a chair 2 dm beneath the level of the couch. E held Ps feet fixed to the couch. P lifted the upper body from the seat and held it straight out from the edge of the couch, with the arms folded across the chest. The maximal time for which P was able to keep the unsupported upper body horizontal was measured with a stop-watch to within 1 second.

Modifications: In the original Biering-Sørensen, the buttocks and legs are fixed by 3 canvas straps and there is an upper time limit of 240 seconds [17].

10. Modified PILE (lumbar/cervical)

PILE = Progressive Iso-inertial Lifting Evaluation.

Modified PILE lumbar: P lifted a tray with weights (plastic bottles filled with sand) from the floor to a 75-cm-high table and back again to the floor. The table was placed 90° to the left of P, which added a twisting factor. An electronic pulse-counter was attached to Ps thorax. The starting weight was 4 kg. E added 2 kg after each successful attempt. Each attempt had to be carried out within 20 seconds. The weight managed during the last lifting moment was recorded as the test result. The test was discontinued if the heart rate reached 85% of the estimated maximal heart rate or if the load reached 55% of the body weight.

Modified PILE cervical: This sub-test was carried out as described above, except that P stood in front of the table and lifted the tray from the table up to a 50-cm-high platform (i.e., 125 cm above the floor). The platform was placed on the left side of the table, which added a twisting factor.

Modifications: In the original PILE, the table is 76 cm high, the platform is 137 cm above the floor, men and women have different weights at the start (3.6 vs 5.9 kg) and different weights are added to men and women (2.25 vs 4.5 kg), and the result is adjusted for the body weight [113]. Our modifications are in line with Lindström et al [93](+ personal communication Ingall Lindström, 2000).

Examination procedure

The test package was performed at different times of day. Along with the agreement to participate, the participants received identical instructions, both verbally and in written form, from a manual produced for this study.

They were to wear training clothes or underclothes, not to do any warming up, and to perform the tests to their maximum capacity within the limits of exertion and pain; they could discontinue whenever they wanted. The participants were also informed that the examiners were a physiotherapist and a research assistant. The patients were not informed about which of the 2 examiners they were seeing. The healthy subjects could not be blinded to the examiner because they were co-workers of one or both of the examiners. Whether A or B would conduct the first examination was randomized by envelopes, which were prepared by an independent statistician and opened immediately before the first test. Close to the start of the examination the participant was once again verbally instructed to perform the tests to her or his maximum capacity within the limits of exertion and pain, and was reminded that the tests could be discontinued whenever she or he wanted. The test package was then conducted straight through without a break and without further verbal communication, except for purely technical instructions on how to perform the test. Before the first and after the last test of the package, the participants were asked to estimate their exertion on Borg's 20-point scale [20] and their level of pain on Borg's 10-point scale [21].

The participants and the examiners were given no results on any occasion until all the tests were completed. The participants were asked not to tell the second examiner their experiences at the first examination.

Inter-rater reliability study

The 30 patients and 20 healthy subjects were first tested by one of the examiners (examination 1). After a break for 30 minutes, they were re-tested by the other examiner (examination 2).

Intra-rater reliability study

The 20 healthy subjects participated. Examiners A and B tested 10 healthy subjects each. After examination 2, the subjects rested for another 30 minutes and were then re-tested (examination 3) by the same examiner as at examination 1. The reason for including only healthy subjects in the intra-rater reliability study was that we considered 3 consecutive examinations too much of a strain for the patients to be ethically defensible; it would also have made the results of the 3rd examination difficult to interpret.

In total, the patients and the healthy subjects were occupied in the study for approximately 1½ and 2½ hours respectively.

Statistics

Intra-class correlation coefficient (ICC) is the basic measure in most reliability studies involving continuous data (degrees, centimetres, etc) [136,19,116,121,63,163]. The ICC increases with the degree of reliability up to a maximum of 1.00 for identical ratings [48]. We calculated the one-way ANOVA (analysis of variance) ICC, random-effects model, and its 95% CI [48]. We also calculated the standard error of

measurement (SEM) of the ICC [139]. The 95% CI is a band of values that, with 95% confidence, contains the true reliability. A narrow CI suggests a more precise estimate of reliability. The SEM enables the reliability of a measurement expressed in the units of the measurement of interest, such as degrees or centimetres, to be assessed. As such, it is valuable for the clinician because it provides guidance on whether the measured change is due to measurement error or to real change [163]. There is a lack of consensus concerning the cut-off values for ICC [136,63]. We chose to consider an ICC > 0.60 to indicate acceptable reliability and an ICC ≤ 0.60 to indicate poor reliability [83,27].

For each sub-test, the mean difference between the measurements and its 95% CI were calculated. The possible systematic error of the ICC was calculated, using a *t*-test to evaluate the mean difference [121]. We considered a sub-test to have acceptable inter- or intra-rater reliability when ICC was > 0.60 and there was no significant, systematic error. A test was considered to have acceptable reliability when it had (1) acceptable inter-rater reliability for the 50 participants, (2) acceptable intra-rater reliability for both examiners A and B and (3), for tests comprising 2, when both sub-tests had acceptable inter- and intra-rater reliability.

The proportions of tests that showed acceptable inter-rater reliability were calculated for the patients and for the healthy subjects, and for the 5 tests that required manual fixation and the 5 that did not. The proportions of tests with acceptable intra-rater reliability were calculated for A and B and for the tests that did and did not require manual fixation. The proportions of tests with acceptable reliability were calculated for the tests that did and did not require manual fixation. The mutual proportions were then compared by a *z*-test [1].

The exertion and pain before and after each examination were analysed. The difference between examinations 1 and 2 of the 50 participants was compared by the Wilcoxon sign-rank test. The differences between examinations 1 and 3 and the differences between the healthy subjects of examiners A and B were compared by the Wilcoxon rank-sum test [2].

A *p*-value < 0.05 was considered statistically significant. Stata, 9.1, was used for the analyses [149].

STUDY III – A RANDOMIZED, CONTROLLED TRIAL

Patients

For the 125 patients in study I a randomized, controlled trial was carried out.

Interventions

One treatment group was allocated to cognitive-behavioural rehabilitation at the rehabilitation centre (rehabilitation group). The other group to continued primary care (primary-care group).

Cognitive-behavioural rehabilitation

The patients of the rehabilitation group received a cognitive-behavioural program of graded activity [92] as described in Table 3.

Participation in the rehabilitation group did not exclude the patient from seeking other care, including primary care, during the follow-up period.

Primary care

The hubs of Swedish primary care are the health centres. Overall medical responsibility belongs to the family doctor. The 13 health centres in this study engaged about 85 family doctors (expressed in full-time duty) and served around 165,000 individuals. Besides family doctors, their staff consisted of physiotherapists, nurses, assistant nurses, occupational therapists and social workers. Besides management at the health centre, primary care could include referral to consultation by, e.g., an orthopedist or a neurologist.

Participation in the primary-care group excluded the patient from turning to the rehabilitation centre during the follow-up period but not from any other health care, including multidisciplinary rehabilitation at units other than the rehabilitation centre.

Table 3. Cognitive-behavioural rehabilitation.

Staff category	Investigation and treatment phase, 2–8 weeks	Frequency
Physician	Mapping out of medical obstacles to working. Handling of the sick-listing. If needed, prescription of drugs (antidepressants, analgesics etc.) and injections of cortisone (in shoulder- or hip-muscle attachments etc.) [43].	1–2 (consultations) per week.
Physiotherapist	Mapping out of biomechanical obstacles to working including a visit to the work place [92]. Start of graded activity: the patient first carried out an activity measurable in minutes, metres, etc., for example a walk, until the pain increased. The starting level was about 25% below that. A gradual increase of the activity was decided on check-ups, the final aim being to manage the load in a job, for the unemployed an imaginary one [92]. If needed, manual therapy [43].	2–3 consultations. 1 / week. 1 / week.
Psychologist or social worker	Mapping out of psychosocial obstacles to working. Cognitive-behavioural therapy focussed on anxiety and depression [97].	1 / week.
Health-care adviser	Start of education in applied relaxation [97].	1 / week for 6–8 w.
Action phase, 2–8 months		
Team	Conference that produced a written rehabilitation plan with: <i>1. Final aim</i> = the optimal degree of work ability that could be achieved and maintained for at least 30 consecutive days. <i>2. Partial aims</i> concerning functioning only (e.g., increase of vocational training by five hours/week); symptom aims, for example, pain reduction, were excluded [92]. <i>3. Means of reaching the aims</i> (e.g., increase of vocational training ½ hour/day week 1, 1 hour/day w. 2 etc.).	At the start of the action phase.
Team	Check-up conferences produced fresh partial aims.	1 / 3–4 weeks.
Team member (usually the physiotherapist)	Vocational conferences with the employer and a clerk from the SIO or, for unemployed patients, the Public Employment Services.	
Physician	Handling of the sick-listing.	1 / 3–4 weeks.
Physiotherapist	Completion of graded activity. Check-ups less frequent.	1 / 3–4 weeks.
Health-care adviser	Completion of education in applied relaxation.	1 / week (f. 6–8 w.)
Psychologist or social worker	If needed: cognitive-behavioural therapy as support during the retraining process [97].	1 / week.
	When the final aim was reached, or when it was obvious that return-to-work would not be achieved.	The end of rehabilitation.

Outcome measures

Return-to-work share

The percentage of patients who regained any degree of work ability for at least 30 days in succession over 18 months; the primary outcome measure. Secondary measures were:

Return-to-work chance

The chance, as expressed in hazard ratios [77], of achieving any degree of work ability over 18 months, irrespective of the duration of that work ability.

Net days

Sick-listing, expressed in whole days, over 18 months and the 3 component 6-month periods. *Net days* = crude days x degree [8].

Visits

The total number of health-care visits over 18 months and over the 3 component 6-month periods. *Visits* comprised consultations at the rehabilitation centre, within primary care and other care, including alternative-care providers, but excluded consultations relating to multidisciplinary rehabilitation at units other than the rehabilitation centre.

Statistics

Return-to-work chance was compared by a Cox regression analysis for recurrent events with event dependence and a time interaction with the exposure variable (i.e., rehabilitation group or primary-care group) and is presented as hazard ratios with 95% CI [77]. It was analysed at 6, 12 and 18 months.

Net days and *Visits* in the 1st, 2nd and 3rd 6-month periods were outcome variables in 2 separate mixed-linear models. In the models, the main effects of 3 explanatory variables and two interaction terms were compared using a random intercept model of the unstructured covariance type on the group level and time as repeated factor [24]. The explanatory variables were time (i.e., 6-month period 1, 2 or 3), rehabilitation group or primary-care group, and subacute or chronic patient. The interaction terms were time x rehabilitation group or primary-care group and time x rehabilitation group or primary-care group x sub-acute or chronic. The models were also adjusted for possible baseline characteristics with significant differences between the groups. The analyses were performed using PROC MIXED in SAS, version 9.1, and the results are presented as separate graphs for the subacute and chronic patients and as means with 95% CI and *p*-values, adjusted for all parameters (main effect and interactions). The 2 patients who died (Figure 7) were excluded from the outcome analyses except from the Cox regression [77]. *Visits* at 18 months were analysed for those patients who had completed all the follow-up forms, while the mixed-linear model also included incomplete responders. To evaluate their possible influence on the treatment results, we also analysed the days of hospital care, the use of surgery for musculoskeletal disorders and multidisciplinary rehabilitation at units other than the rehabilitation centre.

The analyses were performed on an intention-to-treat basis. The total percentage of withdrawals and drop-outs was calculated. This sum should not exceed 30% [166].

Baseline characteristics were compared. A p -value < 0.05 or, concerning the Cox regression, a 95% CI not including 1.00, was considered statistically significant. Except for the mixed linear models, analyses were performed using Stata, 9.1 [149].

Data collection

The sick-listing data were provided by the Stockholm County SIO. As the employer has the financial responsibility for the 2 initial weeks, the available data included only the sick-listing periods > 2 weeks. For the unemployed subjects, however, those data included all periods. Data concerning the rehabilitation centre were collected from the medical records of the centre. Primary care and other health-care data were obtained from follow-up forms. Although these self-report measures have been used successfully in previous research, their reliability has not been established. However, because the patients were free to seek treatment anywhere, the only comprehensive sources of health-care data were self-ratings [100]. The data were fed into a specially designed database using Access, version 2000.

Power calculation

To calculate the power, a preliminary study was performed. In this retrospective study, 172 consecutive patients with non-acute NSP, who completed rehabilitation at the centre during the period 1996–2000, were included. The mean rehabilitation period was 266 (SD±170) days. The *Return-to-work share* was 76%; for subacute and chronic NSP 89% and 73%, respectively ($p < 0.05$). The power calculation was based on this preliminary study and a forecast of the probability of return-to-work after traditional care for NSP [170]. The forecast probability for the patients in the preliminary study was calculated from their current sick-listing at baseline. It proved to be 49%, i.e., 27 percentage units less than the actual rate of 76%. Including an uncertainty about the application of this forecast to our patient sample, we expected to reach a difference between the rehabilitation group and the primary-care group of at least 22 percentage units. With an alpha of 0.05 and a power of 80%, this should require the inclusion of 154 patients; or, to allow a reasonable dropout rate, 170 patients.

Inclusion procedure

Together with the research assistant, the patient completed a questionnaire (see study I) and was categorized as having back and/or neck pain, based on how the patient completed a pain drawing [183, 16] and a short interview. Then, after performing the 10-test package (study II) and a stratification by age ($\leq 44 / \geq 45$ years) and subacute / chronic NSP, the assistant performed the randomization. The 2 treatment alternatives were distributed in opaque envelopes by a computerized block-randomization procedure produced by an independent statistician. The assistant opened the remaining envelope with the lowest random number and presented the content to the patient.

Premature cessation of recruitment

The recruitment started in August 2000 and was discontinued in January 2004, when 125 patients were included. The reason was the opening in April 2004 of a large spine-

rehabilitation centre in a neighbouring municipality (Nacka) on the initiative of the Stockholm County SIO and Stockholm County Council. We presumed that many primary-care group patients might be referred to that centre, which might contaminate the study.

Follow-up

Six, 12 and 18 months after inclusion, the patients completed forms concerning, among other items, health-care utilization. If necessary, a postal reminder was sent after 2 weeks and a telephone reminder after another 2 weeks. If the forms were not returned despite these measures, the data were considered missing. The patient who was last to be included completed the 18-month follow-up period in July 2005.

STUDY IV – A PROSPECTIVE COHORT STUDY

Patients

For the 125 patients in studies I and III, a prospective cohort study was conducted, with a re-assessment of the data with the 2 treatment groups considered as a single cohort.

Outcome variable

Stable return-to-work

Stable return-to-work required a duration of at least 1 month. The reference to *Stable return-to-work* was *Non-return-to-work*, including non-return-to-work a specific day and return-to-work that day but with recurrence of work absence the following month. Due to the employer responsibility, *Stable return-to-work* possibly contained a period of work absence of a maximum of 14 days during the follow-up month. *Stable return-to-work* was analysed in 4 specific days during a 2-year period, selected as 6, 12, 18 and 24 months after baseline.

Predictive variables

Objective variables

The 6 reliable function tests from study II were used as objective variables. Two of those tests included flexion to the right and to the left and rotation to the right and to the left, and a lift test comprised a lumbar and a cervical subtest. Nine subtests in total are given in Table 4.

Table 4. Objective variables. Results of univariate-logistic regression, adjusted for gender and age, with *p*-values of at most 0.10. Footnotes to the right of the table.

Subtests	Class limits	n	Prediction for <i>Stable return-to-work</i>					
			6 months			12 months		
			OR ¹	<i>p</i> -value	95% CI ²	OR	<i>p</i> -value	95% CI
<i>Forward flexion</i> (centimetres (cm))	25–64	41						
	8–24	42	–	–	–	–	–	–
	0–7	41	–	–	–	–	–	–
<i>Modified Schober</i> (cm)	1–3	18						
	4–6	83	–	–	–	–	–	–
	7–19	23	–	–	–	–	–	–
<i>Lateral flexion right</i> (cm)	3–10	41						
	11–15	39	–	–	–	–	–	–
	16–28	44	–	–	–	–	–	–
<i>Lateral flexion left</i> (cm)	2–11	41						
	12–15	38	–	–	–	–	–	–
	16–27	45	–	–	–	–	–	–
<i>Cervical rotation right</i> (degrees)	0–50	44						
	51–60	43	–	–	–	–	–	–
	61–80	37	–	–	–	–	–	–
<i>Cervical rotation left</i> (degrees)	0–50	47						
	51–60	39	–	–	–	–	–	–
	61–80	38	–	–	–	–	–	–
<i>Abdominal endurance</i> (seconds)	0	46						
	1–14	40	–	–	–	–	–	–
	15–75	38	–	–	–	–	–	–
<i>PILE lumbar</i> (kilogram)	0–6	33						
	8–12	45	–	–	–	–	–	–
	14–44	46	–	–	–	–	–	–
<i>PILE cervical</i> (kilogram)	0–6	37	Ref.					
	8–12	47	1.4	NS	0.5–4.4	–	–	–
	14–44	40	2.9	0.09	0.9–9.5	–	–	–

Socioeconomic variables

These were collected from the questionnaire except sick-listing data, which were collected from the SIO. The sick-listing variables were: *Subacute NSP* with the reference *Chronic NSP* and *Low total prior sick-listing* = at most 183 net days during the 2 years prior to baseline, including all diagnoses, with the reference *High total prior sick-listing* ≥ 184 net days (7,176]. In total, 23 socioeconomic variables are presented in Table 5.

Table 4 continued.

Prediction for <i>Stable return-to-work</i>					
18 months			24 months		
OR	<i>p</i> -value	95% CI	OR	<i>p</i> -value	95% CI
Ref. ³			Ref.		
3.4	0.01	1.3–8.8	2.6	0.05	1.0–6.5
2.1	NS	0.8–5.6	1.3	NS	0.5–3.2
–	–	–	–	–	–
–	–	–	–	–	–
Ref.					
2.3	0.09	0.9–6.2	–	–	–
1.9	NS	0.8–4.9	–	–	–
Ref.					
2.9	0.03	1.1–7.6	–	–	–
1.8	NS	0.7–4.7	–	–	–
–	–	–	Ref.		
–	–	–	2.6	0.04	1.0–6.6
			2.7	0.05	1.0–7.1
–	–	–	–	–	–
–	–	–	–	–	–
–	–	–	–	–	–
–	–	–	–	–	–
–	–	–	–	–	–
Ref.			Ref.		
1.1	NS	0.4–2.9	1.4	NS	0.5–4.4
2.8	0.06	1.0–8.4	2.9	0.09	0.9–9.5

¹ = Odds ratio.

² = Confidence interval.

³ = Reference, which always has OR = 1.0.

NS = Non-significant (*p* > 0.10).

Subjective variables

The subjective variables were collected from the questionnaire. They comprised, among other items, a question about the probability of return-to-work: “What do you believe, honestly, is the probability that you will become so much better that you will be able to work at some time in the future?” [34]. *High self-prediction* included the answering alternatives ‘rather probable’, ‘probable’ and ‘very probable’, and *Low self-prediction* ‘rather improbable’, ‘improbable’ and ‘very improbable’. A total of 16 subjective variables are shown in Table 6.

Table 5. Socioeconomic variables. Footnotes to the right of the table. Further explanations in Table 4.

	n	Prediction for <i>Stable return-to-work</i>					
		6 months			12 months		
		OR	p-value	95% CI	OR	p-value	95% CI
<i>Man</i> [49,52,75,32]	56	–	–	–	–	–	–
<i>Young age</i> (≤ 44 yrs) [80,28]	67	–	–	–	–	–	–
<i>Non-immigrant</i> [22] ¹	90	–	–	–	–	–	–
<i>Co-habiting</i> [18] ²	85	–	–	–	–	–	–
<i>Living without children</i> [112]	55	–	–	–	–	–	–
<i>Non-bad economy</i> [82] ³	68	–	–	–	–	–	–
<i>Non-low education</i> [182] ⁴	80	2.2	0.07	0.9–5.6	2.9	0.02	1.2–7.1
<i>White-collar job</i> [171] ⁵	125	–	–	–	–	–	–
Physical work conditions ⁶ :							
<i>No vibrations</i> [159]	84	3.3	0.03	1.2–9.4	2.9	0.04	1.0–7.0
<i>Light physical workload</i> [52]	21	–	–	–	–	–	–
<i>Varied work tasks</i> [52]	46	–	–	–	–	–	–
<i>Non-sedentary work</i> [94]	88	–	–	–	–	–	–
<i>Comfort. work postures</i> [52]	27	–	–	–	–	–	–
Psychosocial work cond. ⁷ :							
<i>No job strain</i> [157]	90	–	–	–	–	–	–
<i>Good social support</i> [62]	94	4.5	0.02	1.2–16.2	–	–	–
<i>Non-unemployed</i> [143]	95	0.5	0.10	0.2–1.2	–	–	–
<i>No work trauma litigation</i> ⁸	978	–	–	–	–	–	–
<i>Non-smoking</i> [80]	75	–	–	–	–	–	–
<i>No indication of alcohol overconsumption</i> [30] ⁹	107	–	–	–	–	–	–
<i>High physical activity</i> [126] ¹⁰	86	–	–	–	–	–	–
<i>Non-obese</i> (BMI ≤ 30) [38]	94	0.4	0.05	0.2–1.0	–	–	–
<i>Subacute NSP</i> ¹¹	38	3.4	0.006	1.4–8.0	2.8	0.02	1.2–6.3
<i>Low total prior sick-listing</i> ¹²	57	3.1	0.005	1.4–6.9	3.1	0.005	1.4–6.9

Table 5 continued.

Prediction for <i>Stable return-to-work</i>						
18 months			24 months			
OR	p-value	95% CI	OR	p-value	95% CI	
–	–	–	–	–	–	¹ = Born in Sweden. Reference: <i>Immigrant</i> (n=34).
2.9	0.006	1.3–6.1	2.6	0.01	1.2–5.4	² Includes living single with children. Reference: <i>Single</i> = living alone, without children (n=39).
–	–	–	–	–	–	³ = 'neither bad nor good', 'good' or 'very good'. Reference: <i>Bad economy</i> = 'very bad' or 'bad' (n=56).
–	–	–	–	–	–	⁴ Reference: <i>Low education</i> = at most junior high school (n=44).
–	–	–	–	–	–	⁵ Out of the 94 non-unemployed patients. Reference: <i>Blue-collar job</i> (n=82).
3.0	0.01	1.3–6.9	3.5	0.004	1.5–8.0	⁶ "State the conditions that you regularly (not occasionally) are exposed to: ...Vibrations (from tools, vehicles etc.) ...Heavy lifting or greater muscle efforts ...Monotonous work tasks ...Sedentary work ...Difficult work postures (bent, twisted, locked etc.)": answer 'no'. References: 'yes' [94].
–	–	–	–	–	–	⁷ Psychological demands (5 items), decision latitude (6 items) and social support (6 items), total scores 5–20, 6–24 and 6–24 respectively; <i>No job strain</i> = non-scoring demands above the midpoint (> 13) and decision latitude below the midpoint (< 15); reference: <i>Job strain</i> = demands above + decision latitude below the midpoint (n=34). <i>Good social support</i> = above the midpoint; reference: <i>Bad social support</i> = below the midpoint (n=30) [39].
–	–	–	–	–	–	⁸ Out of 115 patients (9 patients scored 'I don't know'). "Have you reported your pain as a work trauma?": answer 'no'. Reference: 'yes' (n=18) [160].
–	–	–	–	–	–	⁹ To drink alcohol corresponding to at least 1/2 bottle (= 37.5 centilitres) of strong spirits on one and the same occasion, less than 2–3 times monthly. Reference: <i>Indication of alcohol overconsumption</i> = at least 2–3 times monthly (n=17) [138](+ personal communication Anders Romelsjö 27 Aug 2007).
–	–	–	–	–	–	¹⁰ Physical activity, including walking > 30 minutes, twice/week or more. Reference: <i>Low physical activity</i> : once/week or less (n=38).
5.4	<0.001	2.2–13.0	3.1	0.008	1.4–7.2	¹¹ = a current, full-time sick-listing at baseline for NSP 42–84 days. Reference: <i>Chronic NSP</i> = a corresponding sick-listing of 85–730 days (n=84) [172].
7.7	<0.001	3.3–18.1	4.9	<0.001	2.2–11.0	¹² = a prior 2-year sick-listing for all diagnoses of at most 183 net days. Reference: <i>High total prior sick-listing</i> ≥ 184 net days [22].

Table 6. Subjective variables. Further explanations in Table 4.

	Class limits	n	Prediction for <i>Stable return-to-work</i>					
			6 months			12 months		
			OR	p-value	95% CI	OR	p-value	95% CI
<i>Pain just now</i> (VAS 1–100) [164]	70–100 48–69 0–47	41 43 40	Ref. 2.4 1.5	 0.09 NS	 0.9–6.9 0.5–4.3	 – –	 – –	 – –
<i>Pain worst last week</i> [164]	81–100 68–80 0–67	42 43 39	Ref. 2.5 1.4	 0.09 NS	 0.9–6.8 0.5–4.2	 – –	 – –	 – –
<i>Intermittent pain</i> [95] ¹	–	39	–	–	–	–	–	–
<i>Non-radiating pain</i> [82] ²	–	32	–	–	–	–	–	–
<i>Local pain</i> [164] ³	–	24	–	–	–	–	–	–
<i>Back-pain domination</i> [12] ⁴	–	86	–	–	–	–	–	–
<i>Time since start of NSP</i> (years) [75]	> 5 1.5–5 < 1.5	53 34 37	 – –	 – –	 – –	 – –	 – –	 – –
<i>No surgery for spinal pain</i> [81] ⁵	–	116	–	–	–	–	–	–
<i>No anxiety/depression</i> [95] ⁶	–	26	–	–	–	–	–	–
<i>Tired seldom</i> [153] ⁷	–	59	3.1	0.01	1.3–7.6	–	–	–
<i>No comorbidity</i> [123] ⁸	–	79	–	–	–	–	–	–
<i>Non-severe functional impairment</i> (ODI) ⁹	–	78	2.1	0.09	0.9–4.9	2.9	0.01	1.3–6.8
<i>Health-related quality of life</i> (EQ-5D) [51]	0–0.359 0.360–0.629 0.630–1.0	42 46 36	Ref. 2.8 2.9	 0.06 0.06	 1.0–8.3 0.9–8.9	Ref. 2.1 2.6	 NS 0.06	 0.8–5.4 1.0–7.1
<i>State of health</i> (EQ-VAS) [51]	0–35 36–49 50–100	44 33 47	Ref. 2.2 3.6	 NS 0.02	 0.7–7.0 1.3–10.3	 – –	 – –	 – –
<i>Non-catastrophizing</i> [70] ¹⁰	–	67	2.2	0.08	0.9–5.1	–	–	–
<i>High self-prediction</i> [34]	–	95	4.2	0.03	1.2–15.2	6.4	0.002	1.9–21.0

Table 6 continued.

Prediction for <i>Stable return-to-work</i>						
18 months			24 months			
OR	p-value	95% CI	OR	p-value	95% CI	
–	–	–	–	–	–	¹ Reference: <i>Continual pain</i> = pain whenever awake (n=95).
–	–	–	–	–	–	² Reference: <i>Radiating pain</i> = radiation of pain/numbness to the leg beneath the knee and/or the arm beneath the elbow (n=92).
–	–	–	–	–	–	³ Pain in the back or the neck. Reference: <i>Widespread pain</i> = pain in both the back and the neck (n=100).
–	–	–	–	–	–	⁴ Reference: <i>Neck-pain domination</i> (n=38).
–	–	–	–	–	–	⁵ Reference: Surgery for back and/or neck pain at least once (for example, for a slipped disc) (n=8).
2.3	0.04	1.0–5.4	–	–	–	⁶ Item 5 in EQ-5D [51], alternative 1 = ‘I am not anxious or depressed’. Reference: alternative 2: ‘... moderately...’ or 3: ‘... extremely...’.
–	–	–	–	–	–	⁷ One item from SF-36 [153]: ‘Tired during the last four weeks:’ ‘some of the time’, ‘a little bit of the time’ or ‘none of the time’.
–	–	–	–	–	–	Reference: <i>Tired often</i> = ‘all of the time’, ‘most of the time’ or ‘a good bit of the time’.
Ref.			Ref.			⁸ Reference: <i>Comorbidity</i> = any other, chronic disease except NSP or obesity (n=45).
2.9	0.03	1.1–7.4	2.2	0.09	0.9–5.5	⁹ ODI (Oswestry Disability Index) scores general functional disability associated with back pain, 0–100%: 0–20% = minimal, 21–40% = moderate, 41–60% = severe, 61–100% = extremely severe to crippling disability [36]. Reference: <i>Non-severe functional impairment</i> = ODI < 41%.
1.5	NS	0.6–3.6	1.1	NS	0.5–2.7	¹⁰ Six catastrophizing thoughts, never–always, 0–6, are summarized, 0–36. <i>Non-catastrophizing</i> ≤ 15. Reference: <i>Catastrophizing</i> > 15 (n=39).
–	–	–	–	–	–	
–	–	–	–	–	–	
1.9	0.09	0.9–4.2	–	–	–	
–	–	–	–	–	–	
2.5	0.02	1.2–5.4	–	–	–	
Ref.			–	–	–	
2.1	NS	0.8–5.1	–	–	–	
3.0	0.03	1.1–7.9	–	–	–	
Ref.			–	–	–	
2.0	NS	0.7–5.4	–	–	–	
3.1	0.01	1.3–7.7	–	–	–	
3.6	0.002	1.6–8.0	2.3	0.04	1.1–4.9	
4.4	0.005	1.5–12.4	4.4	0.005	1.4–10.2	

Treatment variables

Sixty-three of the 125 patients received *Cognitive-behavioural rehabilitation* and 62 patients *Traditional primary care*. The treatment options are described in study III.

Statistics

Stable return-to-work

Stable return-to-work for 6, 12, 18 and 24 months, and of disability pension in 24 months were calculated. The proportions were compared between the genders by univariate-logistic regression, adjusted for age and are given with p -values [64]. In the logistic regression *Stable return-to-work* might have the values 1, including the degrees 0.25, 0.50, 0.75 and 1.00, or 0 = *Non-return-to-work*.

Multiple-logistic regression

We built 4 multivariate models for each of the follow-ups at 6, 12, 18 and 24 months with *Stable return-to-work* as outcome variable and the objective, socioeconomic, subjective and treatment variables as predictive variables. Ordinal and continuous variables were divided into classes. The models were adjusted for gender and age. We first explored univariate analyses. The variables with a p -value of ≤ 0.10 are presented with OR, p -values and 95% CI.

They were included in a multiple model, from which the variables with p -values ≥ 0.05 were excluded step-wise to yield a model comprising only variables with p -values < 0.05 . However, in the choice between a model with a larger number of variables including p -values of 0.05 or slightly above and a smaller model with p -values exclusively < 0.05 , the larger model was tested against the smaller. If that test produced a p -value < 0.05 , the larger model was chosen, otherwise the smaller [64]. All possible 1st-order interaction terms were tested. The final models are shown as OR with p -values and 95% CI with goodness-of-fit tests by Hosmer-Lemeshow, the percentages of correctly predicted patients and the areas under ROC-curves [64]. We chose to take into final consideration only the variables that were represented in at least 3 of the 4 follow-ups. Stata, 10.1, was used for the analyses [149].

RESULTS

STUDY I

A flowchart of the study is shown in Figure 6.

Patients and non-patients

A majority of the patients were recruited by a minority of the doctors: 15 doctors (36.6%) recruited in all 94 patients (75.2%). Twenty-one doctors recruited only 1 patient each. Ninety-nine patients (79.2%) were living in 3 of the 9 districts, this number of inhabitants corresponding to 30.8% of the total number of inhabitants in the study area.

The mean age of the 124 patients was 42.6 (r (range) 18–59) years. The proportion of Old age was significantly higher than among the non-patients (Table 2). Women predominated slightly. The mean age of the 338 non-patients was 39.3 (r 19–59) years. Males predominated slightly, but the gender difference versus the patients was non-significant (Table 2).

Outcome

In the univariate analyses, 13 of the 18 conditions had higher odds for the patients with a dominance of physical and psychosocial work strains, and *Indication of alcohol over-consumption* (OR 14.8); only 1 condition, *Single life* (OR 0.5), had lower odds (Table 2). Five conditions qualified for the final multivariate model, the proportion of correctly classified subjects was high and the area under ROC-curve was large (Table 7).

Figure 6. Flowchart of study I.

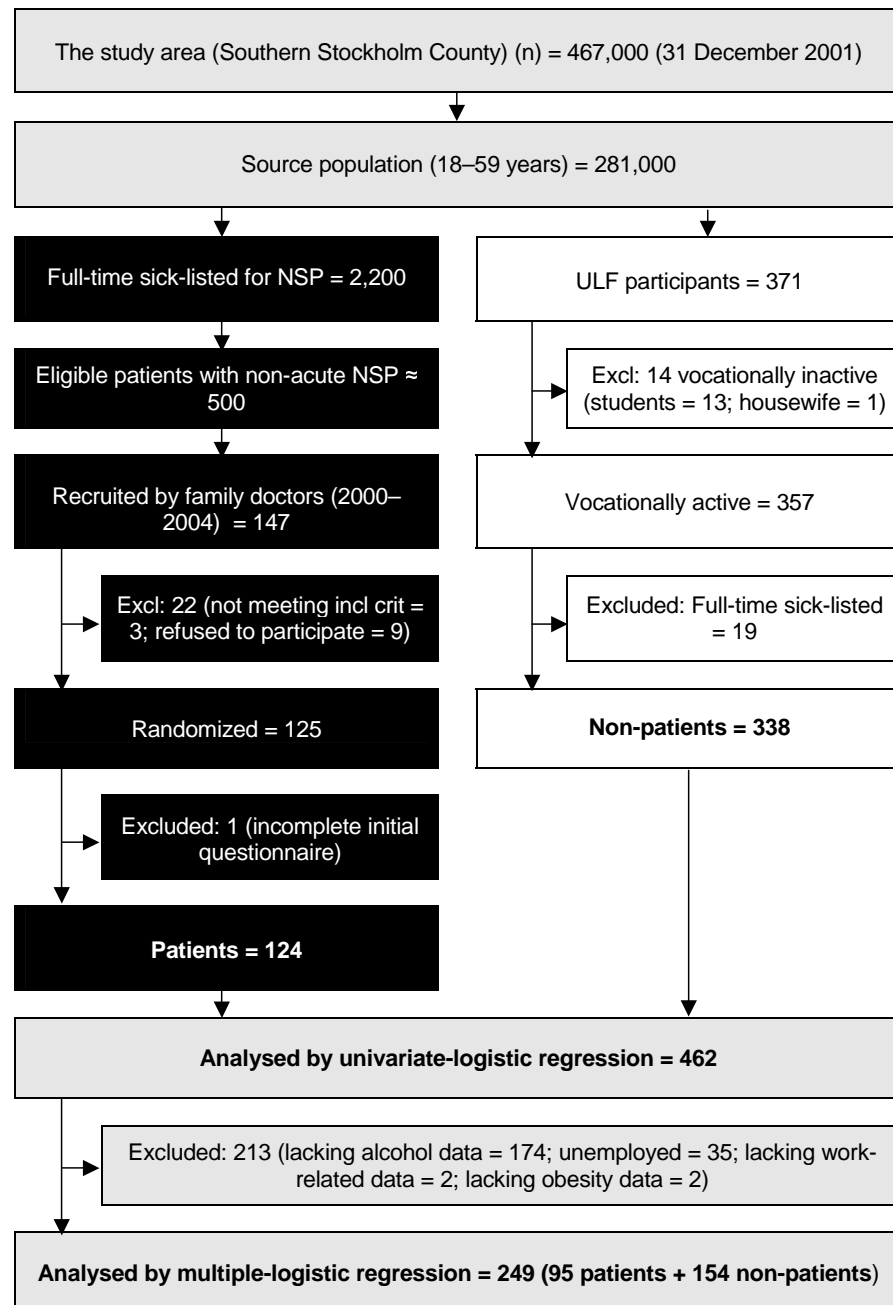


Table 7. Living conditions. Multivariate analysis. Ninety-five patients with non-acute NSP compared with 154 non-patients by logistic regression. Ranking by odds ratios.

	Odds ratio	<i>p</i> -value	95% CI
<i>High physical workload</i>	13.7	< 0.001	5.9–32.2
<i>Hectic work tempo</i>	8.4	0.001	2.5–28.3
<i>Blue-collar job</i>	4.5	0.003	1.8–11.4
<i>Obesity</i>	3.5	0.02	1.2–10.2
<i>Low education</i>	2.7	0.04	1.1–6.8

Goodness-of-fit: Hosmer-Lemeshow 0.57; correctly classified 85.5%; area under ROC 0.92.

STUDY II

A flowchart of the study is shown in Figure 7.

Patients and healthy subjects

Of the 30 patients 17 were women (mean (*m*) 41.5 (*r* 28–60) years) and 13 men (*m* 42.4 (*r* 20–63) years). Of the 20 healthy subjects, 14 were women (*m* 36.2 (*r* 22–55) years) and 6 men (*m* 40.2 (*r* 28–53) years). All 50 participants completed all the tests.

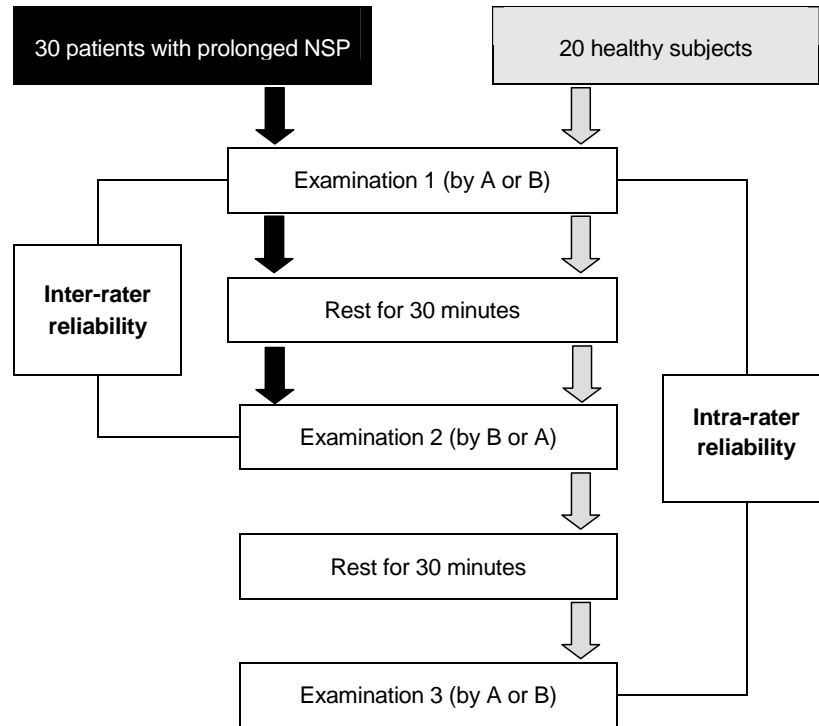
Inter-rater reliability

Seven of the 10 tests had acceptable inter-rater reliability (Table 8). Three tests had poor inter-rater reliability: active-straight-leg raise, cervical bending and modified Biering-Sørensen. For the patients and the healthy subjects, 7 and 4 of the 10 tests respectively had acceptable inter-rater reliability (non-significant (NS)).

All 5 tests that required no manual fixation by the examiner had acceptable inter-rater reliability, compared with 2 of the 5 tests that required such fixation. The difference in proportion (5 vs 2 out of 5 tests) was significant ($p = 0.04$).

The exertion and the pain before and after examination 1 did not differ significantly from those before and after examination 2 (data not shown).

Figure 7. Flowchart of study II.



Intra-rater reliability

For examiner A (the physiotherapist), all 10 tests had acceptable intra-rater reliability (Table 9). For examiner B (the research assistant), 8 tests, i.e., all but the trunk rotation and the modified Biering-Sørensen, had acceptable intra-rater reliability (NS).

All the tests requiring no manual fixation had acceptable intra-rater reliability for both A and B. Of the 5 tests that required manual fixation, 5 and 3 tests had acceptable intra-rater reliability for A and B, respectively (NS).

The exertion and the pain before and after examinations 1 and 3 did not differ significantly between the 10 healthy subjects of A and B (data not shown).

Reliability

All 5 tests requiring no manual fixation had acceptable reliability (forward bending, modified Schober, lateral bending, abdominal endurance and modified PILE). The 5 tests that required manual fixation (trunk rotation, active-straight-leg raise, cervical bending, cervical rotation and modified Biering-Sørensen) had poor reliability except cervical rotation. The difference (5 vs 1) was significant ($p = 0.01$).

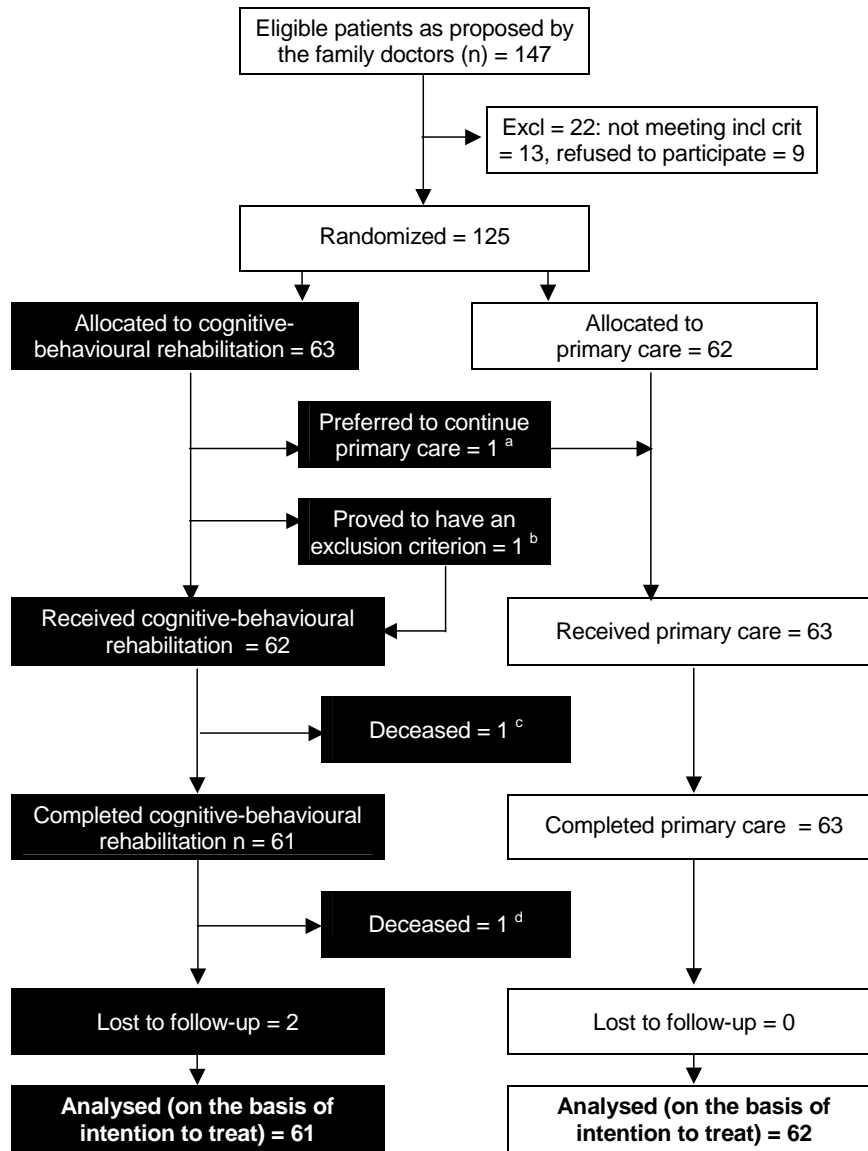
Table 8. Inter-rater reliability. Fifty participants tested by A and B. The 5 tests that required manual fixation are italicized. ICC (= Intra-class-correlation coefficient) in bold indicates acceptable ICC (>0.60). The mean difference between the measurements by A and B is compared. *p*-value in bold text indicates a significant difference ($p < 0.05$). + indicates acceptable, – indicates poor inter-rater reliability. NS = Non-significant. SE = Standard error.

10-test package (including 16 sub-tests)	Forward bending (cm)	Modified Schober (cm)	Lateral bending (cm)		Trunk rotation (degrees)		Active-straight- leg raise (degrees)	
			Right	Left	Right	Left	Right	Left
All 50 participants								
ICC	0.99	0.79	0.93	0.95	0.82	0.85	0.94	0.90
95% CI of ICC	0.98–1.00	0.67–0.88	0.89–0.96	0.91–0.97	0.70– 0.89	0.75– 0.91	0.91–0.97	0.86– 0.95
SE of measurement	1.2	0.7	1.3	1.1	6	6	4	6
Mean	6.4	6.8	17.9	18.1	48	47	68	70
Mean difference	–0.1	0.2	0.3	0.4	1	–1	3	4
95% CI of mean diff.	–0.6–0.4	–0.1–0.5	–0.2–0.8	–0.1–0.9	–1–3.7	–2.8–1.8	1.2–4.6	1.6–6.0
p-value	NS	NS	NS	NS	NS	NS	0.002	0.001
Inter-rater reliability	+	+	+	+	+	+	–	–
30 patients								
ICC	0.99	0.94	0.98	0.97	0.85	0.88	0.96	0.96
95% CI of ICC	0.98–1.00	0.90–0.97	0.93–0.98	0.95–0.98	0.74– 0.91	0.81– 0.93	0.95–0.98	0.94– 0.98
SE of measurement	1.4	0.4	1.0	0.9	6	5	4	4
Mean	9.2	6.6	16.4	16.8	46	43	64	65
Mean difference	0.0	0.2	0.1	–0.2	1	2	2	2
95% CI of mean diff.	–0.8–0.8	–0.1–0.4	–0.5–0.6	–0.7–.3	–1.6–4.3	–0.9–4.3	0.1–3.9	0.2–4.2
p-value	NS	NS	NS	NS	NS	NS	0.04	0.04
Inter-rater reliability	+	+	+	+	+	+	–	–
20 healthy subjects								
ICC	0.95	0.22	0.79	0.85	0.75	0.75	0.84	0.70
95% CI of ICC	0.92–0.97	0.07–0.46	0.68–0.89	0.84–0.95	0.59– 0.85	0.64– 0.87	0.78–0.92	0.62– 0.86
SE of measurement	0.9	1.0	1.5	1.1	6	6	5	7
Mean	2.2	7.1	20.1	20.2	50	52	75	77
Mean difference	–0.3	0.3	0.8	1.4	2	–4	4	6
95% CI of mean diff.	–0.8–0.3	–0.4–0.9	–0.3–1.8	0.6–2.1	–2.4–5.4	–7.8–0.3	0.8–7.6	1.5–10.8
p-value	NS	NS	NS	0.001	NS	NS	0.02	0.01
Inter-rater reliability	+	–	+	–	+	+	–	–

Table 8 continued.

10-test package (including 16 sub-tests)	<i>Cervical bending (degrees)</i>		<i>Cervical rotation (degrees)</i>		Abdominal endurance (seconds)	<i>Modified Biering- Sørensen (seconds)</i>	Modified PILE (kilogram)	
	<i>Forward</i>	<i>Backward</i>	<i>Right</i>	<i>Left</i>			Lumbar	Cervical
All 50 participants								
ICC	0.61	0.84	0.70	0.69	0.92	0.91	0.97	0.97
95% CI of ICC	0.45–0.78	0.78–0.92	0.54–0.83	0.51–0.81	0.87–0.96	0.85–0.95	0.95–0.98	0.94–0.98
SE of measurement	7	5	6	6	8	16	2.2	1.8
Mean	52	65	65	68	32	79	27.8	19.3
Mean difference	4	3	2	1	–2	–8	0.5	0.4
95% CI of mean diff	1.2–6.7	1.3–5.1	–0.4–4	–1.0–3.9	–5.4–1.4	–14.3–1.1	–0.4–1.3	–0.3–1.2
<i>p</i> -value	0.006	0.001	NS	NS	NS	0.02	NS	NS
Inter-rater reliability	–	–	+	+	+	–	+	+
30 patients								
ICC	0.52	0.81	0.64	0.68	0.90	0.96	0.98	0.98
95% CI of ICC	0.36–0.74	0.69–0.89	0.44–0.78	0.49–0.80	0.85–0.95	0.92–0.98	0.96–0.99	0.96–0.99
SE of measurement	8	5	6	7	6	10	2.1	1.5
Mean	48	60	61	66	16	54	24.6	17.2
Mean difference	5	4	2	–1	–3	–2	0.3	–0.1
95% CI of mean diff	0.8–8.9	0.9–6.3	–1.7–4.9	–4.1–3.2	–6.0–0.2	–7.3–3.5	–0.8–1.4	–0.9–0.6
<i>p</i> -value	0.02	0.01	NS	NS	0.04	NS	NS	NS
Inter-rater reliability	–	–	+	+	–	+	+	+
20 healthy subjects								
ICC	0.59	0.86	0.66	0.63	0.86	0.69	0.95	0.94
95% CI of ICC	0.40–0.76	0.80–0.93	0.49–0.80	0.58–0.84	0.76–0.92	0.59–0.85	0.92–0.97	0.91–0.97
SE of measurement	6	4	5	4	12	22	2.3	2.1
Mean	58	72	70	72	55	116	32.5	22.4
Mean difference	3	3	2	4	0	–16	0.7	1.3
95% CI of mean diff	–1.2–6.3	0.0–5.3	–1.0–5.2	1.6–7.0	–8.0–7.3	–30.7–2.1	–0.8–2.2	–0.1–2.7
<i>p</i> -value	NS	0.047	NS	0.004	NS	0.03	NS	NS
Inter-rater reliability	–	–	+	–	+	–	+	+

Figure 8. Flowchart of study III.



^a Woman, 58, allocated to cognitive-behavioural rehabilitation, but continued primary care.

^b Male, 45, incorrectly included: except NSP, he suffered from whiplash-associated disorders that during the initial mapping out (Table 3) showed to be a primary obstacle to working.

^c Male, 55, died 12 months after inclusion from lung cancer.

^d Male, 53, died 11 months after inclusion of a reason which was unknown to us.

All these 4 patients had chronic NSP.

Table 9. Intra-rater reliability. Twenty healthy subjects tested twice by A or B. Further explanations in Table 8.								
10-test package (including 16 sub-tests)	Forward bending (cm)	Modified Schober (cm)	Lateral bending (cm)		Trunk rotation (degrees)		Active-straight- leg raise (degrees)	
			Right	Left	Right	Left	Right	Left
Examiner A								
ICC	0.95	0.87	0.99	0.94	0.92	0.96	0.99	0.97
95% CI of ICC	0.89–0.99	0.68–0.96	0.95–1.0	0.82–0.98	0.76–0.97	0.87–0.99	0.96–1.00	0.92–0.99
SE of measurement	.9	.3	.5	1.0	3	3	2	3
Mean	2.5	7.1	21.2	21.0	55	53	75	78
Mean difference	–0.7	0.2	–0.1	0.1	1	–1	0	1
95% CI of mean diff.	–1.6–0.1	–0.1–.5	–0.5–0.5	–0.9–1.1	–2.6–3.6	–3.8–1.8	–1.6–1.4	–1.7–3.3
<i>p</i> -value	NS	NS	NS	NS	NS	NS	NS	NS
Inter-rater reliability	+	+	+	+	+	+	+	+
Examiner B								
ICC	0.95	0.79	0.73	0.86	0.46	0.83	0.97	0.93
95% CI of ICC	0.86–0.98	0.46–0.93	0.37–0.91	0.61–0.95	0.13–0.85	0.54–0.94	0.90–0.99	0.78–0.97
SE of measurement	0.9	0.7	1.6	1.4	7	5	3	4
Mean	1.8	7.2	19.8	19.7	48	51	70	76
Mean difference	0.4	0.2	–0.7	–0.2	–8	–1	0	–1
95% CI of mean diff.	–0.5–1.3	–0.5–0.9	–2.3–0.9	–1.6–1.2	–14.7–0.3	–6.3–4.3	–2.8–3.0	–4.9–3.5
<i>p</i> -value	NS	NS	NS	NS	0.04	NS	NS	NS
Inter-rater reliability	+	+	+	+	–	+	+	+

STUDY III

A flowchart of the study is shown in Figure 8.

Response rate and analysis of missing data

Data for the baseline characteristics, sick-listing and care at the rehabilitation centre were complete. For other health-care data, the response rates for the 6-, 12- and 18-month forms in the rehabilitation group ($n = 61$) were 57 (93%), 56 (92%) and 55 (90%), respectively, and all forms were answered by 51 patients (84%)(complete answerers). The corresponding rates for the primary-care group ($n = 62$) were 50 (81%), 48 (77%), 50 (81%) and 42 (68%).

In the rehabilitation group, non-responders had: (1) at 6 and 12 months longer previous sick-listing (372 vs 215 days, $p = 0.008$, and 371 vs 214 days, $p = 0.01$, respectively) and longer current sick-listing at baseline (367 vs 158 days, $p < 0.001$, and 346 vs 156 days, $p < 0.001$, respectively); (2) at 12 months a higher prevalence of unemployment (60 vs 18%, $p = 0.03$; (3) for the non-complete answerers, a longer current sick-listing (275 vs 151, $p = 0.003$).

Table 9 continued.

Continuation: 10- test package (including 16 sub-tests)	<i>Cervical bending (degrees)</i>		<i>Cervical rotation (degrees)</i>		Abdominal endurance (seconds)	<i>Modified Biering- Sørensen (seconds)</i>	Modified PILE (kilogram)	
	<i>Forward</i>	<i>Backward</i>	<i>Right</i>	<i>Left</i>			Lumbar	Cervical
Examiner A								
ICC	0.86	0.98	0.94	0.86	0.90	0.92	0.93	0.95
95% CI of ICC	0.67–0.96	0.95–0.99	0.82–0.98	0.63–0.95	0.75–0.97	0.75–0.97	0.80–0.98	0.86–0.98
SE of measurement	2	2	2	3	9	16	2.3	1.5
Mean	58	75	72	74	66	117	31.8	20.8
Mean difference	2	1	1	–1	7	4	0.8	0.4
95% CI of mean diff.	–0.8–4.0	–0.5–2.5	–1.4–3.0	–3.8–1.4	–2.7–15.7	–12.3–19.3	–1.6–3.2	–1.1–1.9
<i>p</i> -value	NS	NS	NS	NS	NS	NS	NS	NS
Inter-rater reliability	+	+	+	+	+	+	+	+
Examiner B								
ICC	0.62	0.80	0.82	0.82	0.65	0.20	0.97	0.94
95% CI of ICC	0.12–0.85	0.53–0.94	0.53–0.94	0.52–.94	0.18–0.86	0.14–0.85	0.89–0.99	0.83–0.98
SE of measurement	6	5	4	4	17	17	2.4	2.9
Mean	57	67	68	67	46	104	32.8	23.5
Mean difference	–1	3	–1	–1	3	33	0.0	–1.8
95% CI of mean diff.	–6.7–5.6	–2.0–7.6	–5.4–2.6	–4.7–2.7	–14.3–20.5	15.3–50.5	–2.4–2.4	–4.7–1.1
<i>p</i> -value	NS	NS	NS	NS	NS	0.002	NS	NS
Inter-rater reliability	+	+	+	+	+	–	+	+

In the primary-care group, non-responders had: (1) at 6 months and for the non-complete answerers a lower age (35.8 vs 44.8 years, $p = 0.006$, and 38.3 vs 45.3, $p = 0.01$, respectively); (2) at 6 months a higher proportion of singles (58 vs 28%, $p = 0.046$); (3) at 12 months a lower health-related quality-of-life (EQ-5D)(0.357 vs 0.562, $p = 0.046$). At 18 months there were no significant differences between responders and non-responders.

Baseline characteristics and participant flow

Except for a higher prevalence of widespread pain in the rehabilitation group (55/63 (87[79–96]%) compared with the primary-care group (45/62 (73[61–84]%) ($p = 0.04$), there were no significant differences. When analysed separately, the subacute patients were mutually equal, while the chronic rehabilitation-group patients had a much higher prevalence of widespread pain: 93[85–100]% vs 68[54–82]% for the chronic primary-care-group patients ($p = 0.004$).

The rehabilitation-group patients started the programme within 1 week and 61 of them completed cognitive-behavioural rehabilitation; all primary-care-group patients completed primary care (Figure 8). The 2 deceased rehabilitation-group patients had passed the “red-flags” examinations [173] at the start without remark.

Outcome measures

Return-to-work share

There were no significant differences between the rehabilitation group and the primary-care group, or between the subacute and chronic patients considered separately (Table 10). In both groups, most of the patients who returned to work returned to full-time work: 20/35 (57%) and 25/35 (71%) respectively (NS). The mean degrees of work ability at return-to-work were 0.77 [0.67–0.87] and 0.85 [0.76–0.94] respectively (NS).

Table 10. Return-to-work share, Net days and Visits. Point estimates at 18 months. Descriptive statistics.

	Patients	Rehabilitation group	Primary-care group
<i>Return-to-work share (%)</i>	All	35/61 (57 [45–70])	35/62 (57 [44–69])
	Subacute	18/20 (90 [76–104])	15/18 (83 [64–102])
	Chronic	17/41 (42 [26–57])	20/44 (46 [30–61])
<i>Net days</i>	All	397 [354–440]	391 [345–436]
	Subacute	327 [261–392]	292 [194–391]
	Chronic	431 [377–486]	431 [383–478]
<i>Visits</i>	All	55.7 [49.3–62.2]	52.0 [38.1–66.0]
	Subacute	48.3 [38.5–58.1]	40.6 [23.1–58.1]
	Chronic	60.1 [51.6–68.7]	56.6 [38.1–75.2]

Return-to-work chance

The hazard ratio for the rehabilitation group increased over the three 6-month periods in comparison to the primary-care group, but the difference did not reach significance (Table 11). The subacute rehabilitation-group patients showed a substantial increase over these periods and achieved a significantly higher hazard ratio at 18 months than the subacute primary-care-group patients. There were no differences for the chronic patients.

Table 11. Return-to-work chance. Cox regression for recurrent events. Hazard ratios for the rehabilitation group as compared with the primary-care group with 95% CI. Significant differences in bold figures.

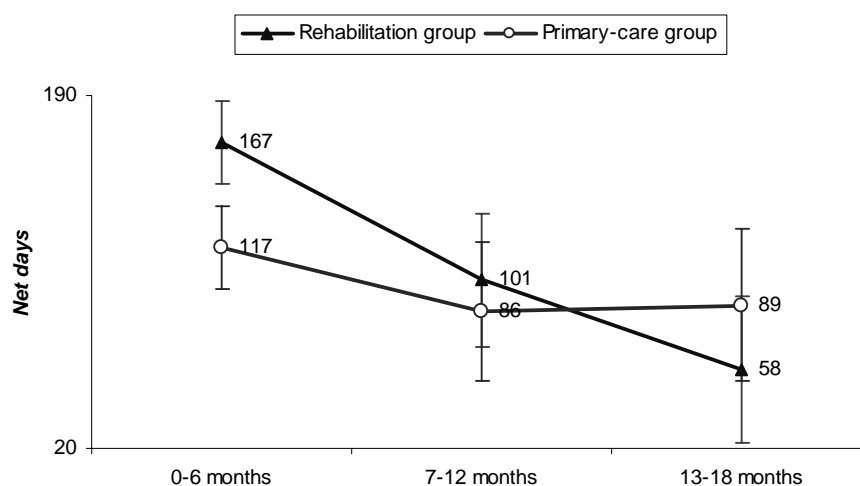
Rehabilitation group	6 months	12 months	18 months
All patients (n = 61)	0.9 [0.6–1.4]	1.2 [0.7–2.0]	1.6 [0.7–3.6]
Subacute patients (n = 20)	0.9 [0.5–1.6]	1.8 [0.8–3.9]	3.5 [1.001–12.2]
Chronic patients (n = 41)	0.9 [0.5–1.6]	0.9 [0.4–2.1]	1.0 [0.3–3.9]

Net days

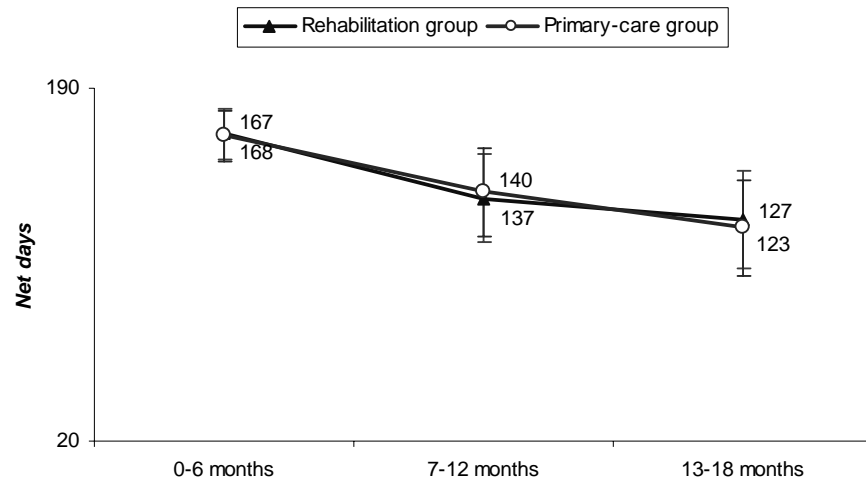
At 18 months there were no significant differences between the groups or the subacute and chronic patients considered separately (Table 10). Over the three 6-month periods, the decrease was significantly more rapid for the whole rehabilitation group and for the subacute rehabilitation-group patients considered separately (bottom of Figure 9). In the 1st 6-month period, there were 50 more *Net days* for the subacute rehabilitation-group patients; in the 3rd period there were 31 days fewer (Figure 9 a). There were no differences for the chronic patients (Figure 9 b). Adjustment for widespread pain showed no changes.

Figure 9 a-b. Net days. Mixed linear model with 95% CI. At the bottom the explanatory variables and their *p*-values are shown. Bold figures indicate a significant difference.

a. Subacute patients.



b. Chronic patients.



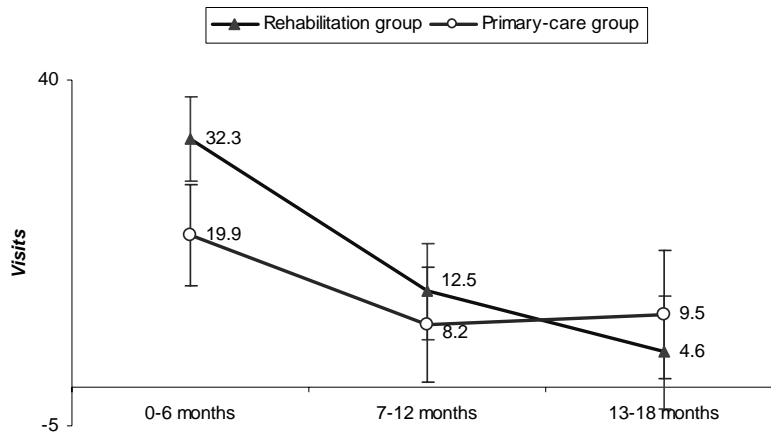
Time (1st, 2nd or 3rd 6-month period): $p < 0.001$; Rehabilitation group or primary-care group: NS; Subacute or chronic: $p < 0.001$; Time x rehabilitation group or primary-care group: $p = 0.008$; Time x rehabilitation group or primary-care group x subacute or chronic: $p < 0.001$.

Visits

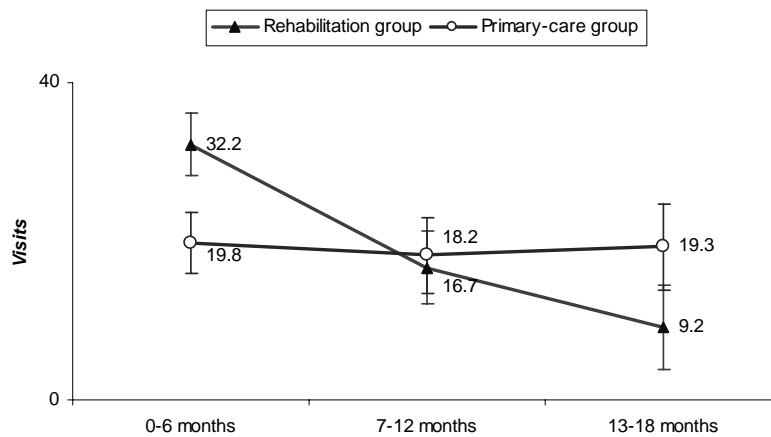
At 18 months there were no significant differences between the treatment groups or the subacute and chronic patients considered separately (Table 10). Over the three 6-month periods, the decrease was significantly more rapid for the whole rehabilitation group (bottom of Figure 10). For the subacute patients, the rehabilitation group showed a continuously decreasing trend while the primary-care group showed a substantial decrease between the 1st and 2nd 6-month periods but no further reduction (Figure 10 a). For the chronic patients, the rehabilitation group showed a continuous decrease while the primary-care group showed no reduction. *Visits* were substantially more numerous for both the subacute and chronic rehabilitation-group patients during the 1st period, but around half in the 3rd period (Figure 10 a-b). The difference in the rate of decrease between the subacute and chronic patients considered separately was NS (bottom). Adjustment for widespread pain gave no changes.

Figure 10 a-b. Visits. Mixed linear model. Further explanations in Figure 9 a-b.

a. Subacute patients.



b. Chronic patients.



Time (1st, 2nd or 3rd 6-month period): $p < 0.001$; Rehabilitation group or primary-care group: NS; Subacute or chronic: NS; Time x rehabilitation group or primary-care group: $p < 0.001$; Time x rehabilitation group or primary-care group x subacute or chronic: NS.

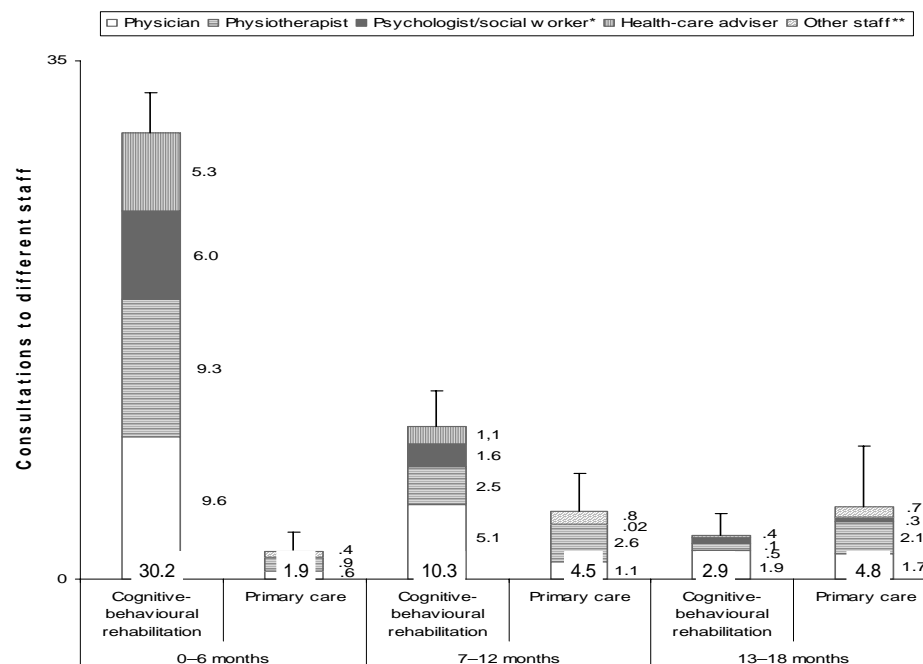
Interventions

Cognitive-behavioural rehabilitation

The total rehabilitation period was m 328 ((SD) \pm 195) days, the investigation and treatment phase 42 (\pm 18) and the action phase 287 (\pm 193) days. Over 18 months there were 45.1 [39.2–50.9] consultations, of which most took place in the 1st 6-month period, followed by a rapid reduction (Figure 11). Totalling 0–18 months, the most and

2nd most frequent consultations were with a physician (16.6 [14.4–18.7]) and a physiotherapist (12.3 [10.5–14.1]).

Figure 11. Consultations for the rehabilitation group. For the total number (presented at the bottom of the staples), 95% CI (upper part) are shown.



*Concerning primary care, social worker was the only option.

Primary care

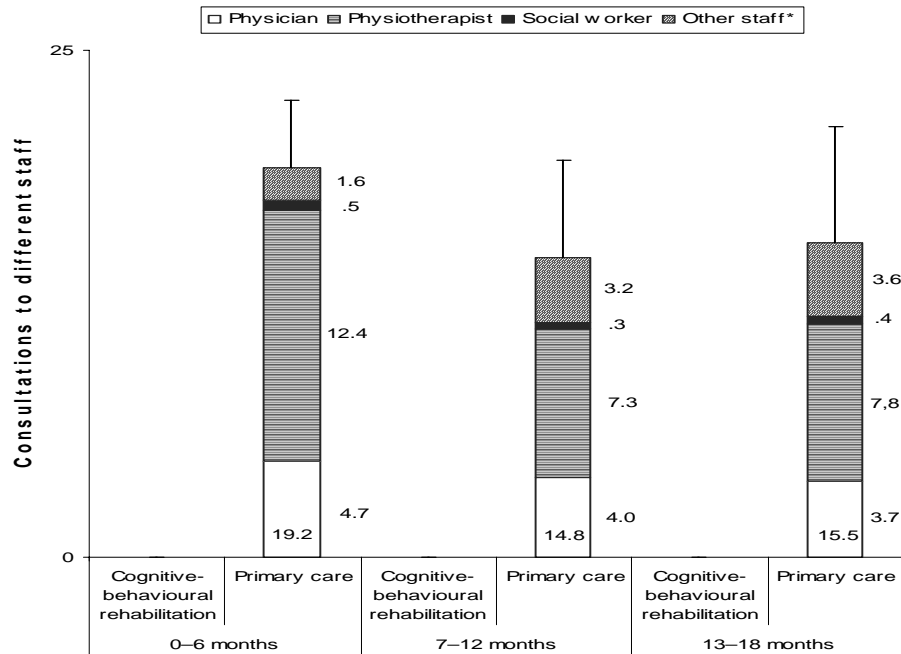
For the rehabilitation group, primary care over 18 months comprised 11.7 [6.7–16.7] consultations. After a slight increase from the 1st to the 2nd 6-month period, there was stagnation (Figure 12). During the 1st 6-month period most of the rehabilitation-group patients (41/57 (72%)) had no primary-care consultations at all.

For the primary-care group, care included 50.9 [37.5–64.3] consultations. After a slight decrease there was no further reduction (Figure 12). In total, the most frequent consultations were with a physiotherapist (28.9 [19.4–38.4]) and a physician (12.4 [10.2–14.7]).

Other treatment efforts

Hospital care was received by the rehabilitation and the primary-care group for 1.2 [–0.2–2.6] and 0.8 [0.1–1.6] days, surgery for musculoskeletal disorders by 1/51 (2[–2–6]%) and 3/43 (7[–1–15]%), and multidisciplinary rehabilitation at other units than the rehabilitation centre by 1/50 (2[–2–6]%) and 4/43 (9[0–18]%), respectively. The differences were NS.

Figure 12. Consultations for the primary-care group. Further explanations in Figure 10.



*= Occupational therapist, nurse, X-ray/MRI staff, laboratory personnel, and complementary-medicine staff (for example, masseur and “Chinese doctor”)

STUDY IV

A flowchart of the study is shown in Figure 13.

Loss to follow-up

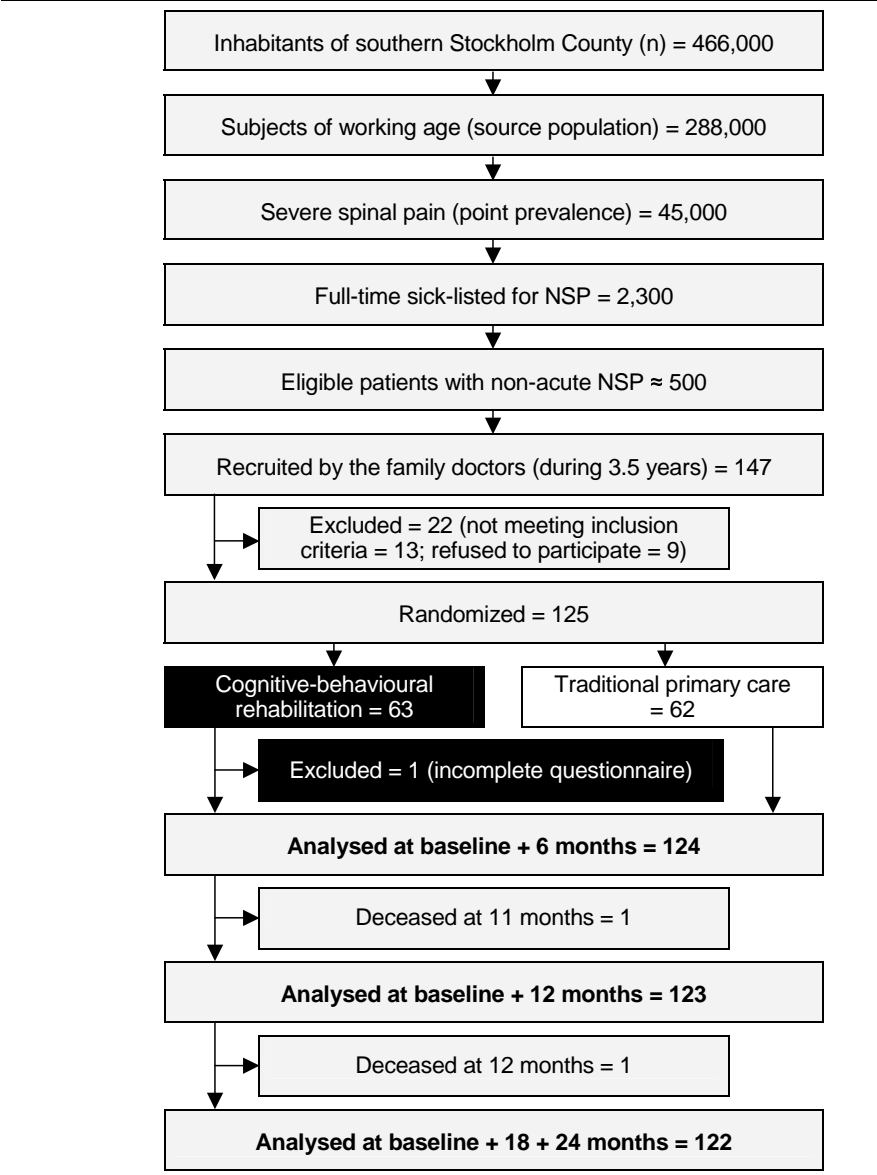
Three of the 125 patients, all males, deceased during the follow-up, 11, 12 and 22 months after baseline. The last deceased patient was excluded from the study because of an incomplete questionnaire. The other 2 subjects were analysed up to their possible follow-ups.

Stable return-to-work

Stable return-to-work gradually increased and was 58/122 (47.5%) at 24 months, a majority at full-time (43/58 = 74.1%). The proportions were generally higher for men, but the gender differences were non-significant (Table 12). At 24 months, disability

pension (temporary or permanent) was received by 30/122 (22 full- and 8 half-time pensions), with a significantly higher proportion of women, 22/68 (32.4%), than men, 8/54 (14.8%) ($p = 0.04$).

Figure 13. Flowchart of study IV.



Predictors of Stable return-to-work

In the univariate analyses, several objective, socioeconomic and subjective variables were associated with *Stable return-to-work* (Tables 4-6), while the treatment variables were not predictive in any of the follow-ups.

In the multiple-logistic models only socioeconomic and subjective variables remained. Three variables were finally considered, all represented in 3 follow-ups (Table 13): *Low total prior sick-listing* was the strongest predictor in 2 follow-ups, and *High self-prediction* and *Young age* were the strongest and 2nd strongest, respectively, in 1 follow-up. No interaction term was predictive. The model fit was generally good and the proportions of correctly classified patients were satisfactory (on average 74.1%).

Table 12. Stable return-to-work. All patients and gender. Comparisons by univariate logistic regression, adjusted for age.

	6 months			12 months			18 months			24 months		
	n	<i>Stable</i> <i>return-to-</i> <i>work</i> n (%)	<i>p</i>	n	<i>Stable</i> <i>return-to-</i> <i>work</i> n (%)	<i>p</i>	n	<i>Stable</i> <i>return-to-</i> <i>work</i> n (%)	<i>p</i>	n	<i>Stable</i> <i>return-to-</i> <i>work</i> n (%)	<i>p</i>
All	124	33 (26.6)	–	123	48 (39.0)	–	122	55 (45.1)	–	122	58 (47.5)	–
Men	56	19 (33.9)	NS	55	27 (49.1)	NS	54	29 (53.7)	NS	54	30 (55.6)	NS
Women	68	14 (20.6)		68	21 (30.9)		68	26 (38.2)		68	28 (41.2)	

Table 13. Predictors of *Stable return-to-work*. Multiple-logistic regression. The variables found in at least 3 follow-ups are in bold text.

	Prediction for <i>Stable return-to-work</i>											
	6 months			12 months			18 months			24 months		
	OR	<i>p</i>	95% CI	OR	<i>p</i>	95% CI	OR	<i>p</i>	95% CI	OR	<i>p</i>	95% CI
<i>Young age</i>	–	–	–	2.8	0.02	1.2–6.5	3.5	0.001	1.3–9.1	2.7	0.02	1.2–6.2
<i>Non-low education</i>	–	–	–	–	–	–	3.0	0.04	1.1–8.2	2.9	0.02	1.2–6.9
<i>Subacute NSP</i>	3.2	0.02	1.3–8.2	–	–	–	3.0	0.04	1.1–8.4	–	–	–
<i>Low total prior sick-listing</i>	–	–	–	2.7	0.02	1.2–6.4	4.8	0.001	1.9–12.3	3.8	0.002	1.6–8.7
<i>Back-pain domination</i>	9.5	0.004	2.0–44.4	2.9	0.04	1.1–7.7	–	–	–	–	–	–
<i>Non-catastrophizing</i>	–	–	–	–	–	–	3.4	0.01	1.3–9.1	–	–	–
<i>High self prediction</i>	4.1	0.02	1.1–15.7	5.2	0.009	1.5–17.5	–	–	–	2.7	0.06	0.9–7.8
Goodness-of-fit:												
Hosmer-Lemeshow	0.70			0.38			0.29			0.67		
Correctly classified (%)	78.2			71.5			73.0			73.8		
Area under ROC	0.79			0.79			0.85			0.79		

GENERAL DISCUSSION

The rehabilitation of non-acute, non-specific spinal pain was elucidated from 4 aspects:

Epidemiology (a cross-sectional study): Thirteen of 18 living conditions associated with long-term sick-listing had higher prevalence in the patients with non-acute NSP, versus non-patients, dominated by work strains and indications of alcohol abuse; in the multivariate analysis, 5 conditions qualified, touching upon work strains, lower social class and life-style.

Reliability (a methodological study): In the performance of function tests for patients with non-acute NSP, an examiner without formal medical education could be used without loss of quality, at least for function tests requiring no manual fixation.

Treatment (a randomized controlled trial): Though the results were equivalent over 18 months, there were indications that cognitive-behavioural rehabilitation in the longer run might be superior to primary care; for subacute NSP in terms of sick-listing and health-care visits; for chronic NSP, in terms of visits only.

Return-to-work prediction (a prospective cohort study): Three (of in total 50) variables predicted a stable return-to-work: low total prior sick-listing, young age, and the patients' own belief in return-to-work; function tests and treatment were non-predictive.

EPIDEMIOLOGY

Work strains

The association of *High physical* workload with NSP has been pinpointed in many studies [28,52,80,38]. Job strain, i.e., high demands, including among other items a high work tempo, and low control [158], has been associated with disabling NSP in several studies [140,66,39,54]. *Hectic work tempo* as an independent risk factor, however, is far less clear. In a review of risk factors for NSP, insufficient evidence was found for high work pace [62]. Despite occasional studies that indicate a relationship between high work tempo and a longer time to return-to-work [165], a rather recent review showed strong evidence for the recovery expectations of the patients, while stress/psychological strain were non-predictive [67]. This is also in line with study IV, where *High self-prediction* qualified as a predictor of stable return-to-work, while work-related variables did not.

Indicators of lower social class

Blue collar job and *Low education* are closely associated and might be looked upon as different aspects of belonging to a lower social class [171]. *Low education* limits the chances of getting a white-collar job, which explains the great dominance of work strains in the final model and the fairly low degree of variance for *Blue collar job* and *Low education* in themselves (Table 7). A possible association is probably a matter of social disadvantage, although it is not clear which aspects of the disadvantage are important [171]. In study IV, there were indications that the non-predictor *Low education*, may have qualified as a predictor with a longer follow-up than the 2 years of that study. *Blue-collar job*, however, was a clear non-predictor; with a prevalence of

87.4% of the patients versus 33.2% of the non-patients, it is logical that such a great difference qualifies for a multivariate analysis with the sample class as outcome variable. Study IV, however, exclusively involved patients with return/non-return-to-work as the outcome. A variable of such overwhelming frequency might be non-discriminative, although a powerful effect on sick-listing. Concerning the possible association between sick-listing for NSP and social class, and according to a large 2004 review, there is a lack of evidence [53]. Our project might contribute to the elucidation of this complex issue.

Life style

Associations between smoking and NSP have been found in several prior studies. A review of 1999 indicated smoking as a weak risk indicator but not a cause of NSP and causality were indicated only in the study with the largest sample, > 30,000 subjects [87]. In a 2000 review, a possible association was suggested, but the lack of prospective studies was emphasized [41]. In a recent meta-analysis, smoking was associated with NSP, though fairly modestly [144]. So the non-significance of *Smoking* in studies I and IV might be due to the small sample size.

Alcohol abuse constitutes one of the greatest health problems. In 2001, 10–15% of all men and about 5% of all women suffered from chronic alcohol dependency [11]. In Sweden the heavy abuse of alcohol has increased 20% since 2000 [85]. Among the subjects with chronic alcohol dependency about ¼ are in a phase of active abuse [11]. This corresponds to 2–2.5% of the non-patients of study I, which was approximately confirmed by the ULF data. The patients had a substantially higher prevalence. We have found nothing equivalent in any other study of NSP. The reason could be our use of one single binge-drinking question (Table 5, footnote 9) [138], which might decrease the risk for under-estimation of abuse in questions of total intake. We have found no previous study of NSP where this question has been used. However, in the multivariate context, the alcohol issue was eliminated by other closely-correlated variables. E.g., 15 of the 16 subjects in the multivariate analysis with *Indication of alcohol over-consumption* had a *Blue-collar job*. One prior study indicated alcohol over-consumption as a risk factor for long-term sick-listing for NSP [61], but this was contradicted by another study [30]. In study IV, *Indication of alcohol over-consumption* did not predict sick-listing during a 2-year follow-up. These conflicting results motivate further research.

During recent decades the prevalence of obesity has increased remarkably. It doubled in Sweden from 5% in 1980/81 to 10% in 2004/5 in both sexes [122]. Comorbidities with obesity include diabetes, cardiovascular diseases, pain in general and NSP in particular [69]. In our study, the prevalence of *Obesity* in the non-patients during the years 2000/01 corresponded well with the 7% in 1996/97 concerning all Swedes 16–84 years [122]. Among the patients it was more than 3-fold higher. *Obesity* remained significant in the multivariate model, though with a decreased OR, probably influenced by *Low education*, which is a risk factor for obesity [137]. According to a 2000 review, obesity should be considered a possible weak risk indicator, but with insufficient data to assess whether it causes spinal pain [88]. In a prospective study from 2002, obesity was a risk factor for the transition from acute to non-acute NSP, though with low OR (1.7) [38]. However, in a large review from 2004, there was insufficient evidence for obesity as a risk factor for non-acute NSP [53]. A quite recent,

very large, cross-sectional population-based study from Norway indicated associations between obesity and NSP and commented that further studies were needed to determine whether the association was causal [58]. Study IV, however, found no impact of *Obesity* on sick-listing. *Obesity* was found in 24.2% of the patients versus 6.9% of the non-patients. In line with the paragraph above, such a difference might qualify for a model with the sample class as outcome variable, but be eliminated in an analysis with return-to-work as outcome. It therefore remains unclear whether, how and why obesity and NSP are correlated [69].

To sum up: the patients were distinguished by higher odds of obesity, higher odds of indication of alcohol abuse that vanished in the multivariate analysis, and non-significant differences concerning smoking. Prospective research, including study IV, has yielded conflicting results. Therefore, the causal associations between life style and obesity and sick-listing for NSP, if any, are small.

RELIABILITY

Several prior studies have elucidated the problem of achieving agreement between medically skilled examiners [19,152,37]. It seems reasonable that an examiner without medical practice will experience even greater difficulties. In support of this, the tests in our package that required fixation tended to have a higher proportion of acceptable intra-rater reliability for the physiotherapist than for the research assistant (5 vs 3 tests), though the difference was non-significant. All the technically least advanced of our tests, i.e., the 5 tests that required no manual fixation by the examiner, had acceptable inter-rater reliability (5 out of 5 tests). This is consistent with the study of Bertilsson et al [15], in which a simple sensitivity test had acceptable inter-rater reliability while several more sophisticated tests had not. The abdominal endurance had acceptable reliability, as against the study of Moreland et al [116], in which the hands of the participant were held on the cheeks, while in study II the hands were stretched out towards the patellae. The test package was inexpensive and easy to perform; while abdominal endurance should be tested in the same way as in our study and the modified PILE used in this study could be recommended, the Biering-Sørensen test with our modification should not be used. All 5 tests requiring no manual fixation had acceptable reliability vs 1 of the tests which required fixation (cervical rotation). This difference (5 vs 1) was significant ($p = 0.01$). Notwithstanding its limitations (see below), study II indicates that even an examiner with no formal medical education could be used without loss of quality, at least for tests that require no manual fixation.

TREATMENT

Sick-listing

Why was *Return-to-work share* substantially lower than expected for the rehabilitation group (57 vs 76%) and higher than expected for the primary-care group (57 vs 49%)? The higher rate of widespread pain, which might complicate return-to-work [131, 164], in the rehabilitation group might be one explanation. On the other hand, that variable was a non-predictor in study IV. Some of the patients in the primary-care group may have been forced to return to work without recovery because of the greater restrictions

in sick-listing [40]. This inference is supported by the growing health problems in Sweden arising from sickness presenteeism, i.e., work despite a need for sick-listing [13,14,86]. Another explanation might be that the recruiting physicians supplied better return-to-work measures than the primary-care average. Anyhow, the low *Return-to-work share* in the rehabilitation group was disappointing, and even if the primary-care group had shown as low a *Return-to-work share* as predicted, the difference had remained non-significant.

However, when subacute and chronic patients were analysed separately, a different picture emerged: the *Return-to-work share* for the subacute rehabilitation-group patients was as expected, but the share for the chronic rehabilitation-group patients was far lower. The significantly better *Return-to-work chance* at 18 months and the more rapid reduction in *Net days* among the subacute rehabilitation-group patients highlighted this. Previous research supports the view that cognitive-behavioural interventions at an early stage of disabling NSP can prevent long-term disability [92,98,99,100], while the effect on sick-listing is more doubtful for chronic NSP [128]. Previous research on graded activity had an occupational-care setting and concerned subacute patients only [92,60,150,59,6,148]. Two earlier studies [92,60] found that graded activity decreases sick-listing. Two later studies [150,59] contradicted that; however, their follow-up period did not exceed 12 months. The better sick-listing trend for the subacute rehabilitation-group patients was not obvious until after 12 months. Thus, the possibility that a longer period of graded activity has a positive effect on sick-listing for subacute patients in a primary-care setting could not be excluded.

Unlike prior research on graded activity, we also included chronic NSP. Most of the rehabilitation-group patients (68%) had a current sick-listing > 12 weeks at baseline. Our programme did not reduce their sick-listing. Why? One reason could be its comparatively limited extent. Haldorsen et al [50] showed that for return-to-work light multidisciplinary treatment was adequate for moderately-disabled but not for highly-disabled patients. For the latter group, extensive multidisciplinary treatment totalling 120 hours was required; the light programme was no better than standard care. Jensen et al [72] showed that an extensive behavioural-rehabilitation programme (fully 120 hours) for long-term NSP in women reduced sick-listing while more limited efforts did not. Men, however, achieved no better results from the full-time programme than from a light programme or standard care. Staal et al [148] found that moderately disabled subjects benefited more from graded activity than those with higher disability scores. These studies indicate that return-to-work for patients with chronic NSP, if it is ever possible, requires a more extensive concept than our programme.

Health-care visits

The resources spent on the rehabilitation group in the 1st 6-month period were balanced by fewer consultations in primary care and a trend towards fewer *Visits* in the long run (Figures 11–12). Also, although the differences were non-significant, the rehabilitation group tended to experience less surgery and other multidisciplinary rehabilitation. For patients with subacute NSP, this agrees with Linton et al [100], whose cognitive-behavioural interventions were followed by a decrease in health-care utilization. For patients with chronic NSP, our findings are consistent with a large review showing that cognitive-behavioural programs have a substantial positive impact on psychological and medical function but only a small impact on sick-listing [97].

RETURN-TO-WORK PREDICTION

Predictors in study IV compared with prior research

Young age is in line with several previous studies and reviews [125,49,164,80,53,75]. Also *High self-prediction* is a well known predictor [34,95,31,32].

One of the most consistent predictors in previous research was low prior sick-listing for spinal pain [28,80,172]. According to one of the hitherto most extensive reviews of predictors of long-term sick-listing for spinal pain, prior sick-listing for all diagnoses has been insufficiently studied [53]. Our results indicate that it is very important to map prior sick-listing for all diagnoses, not only for spinal pain. This is also in line with some prior studies [95,22,12].

Non-predictors compared with prior research

Two non-predictors were in line with previous studies, *Comfortable work postures* [28,80] and *Good social support* [164,80]. The non-predictor *Non-smoking* is closer discussed above. Six of our non-predictors contradicted prior research:

Man and *Non-low education* were non-predictors, while prior research indicated them as predictors, at least for disability pension [22,53]. However, the proportion of disability pension at 24 months was significantly lower for men and *Non-low education* was close to qualify with a representation at both 18 and 24 months (Table 13). It is logical that a disability pension will be granted only after prolonged sick-listing and that education might influence return-to-work comparatively late in a rehabilitation process, when the medical efforts have been replaced by vocational measures. Consequently, our findings might be in line with prior research, although a longer follow-up than 2 years is required to confirm this. *High physical workload* were seen in 83.2% of the patients vs 15.7% of the non-patients. In line with the discussion above, a variable of such a high prevalence might be non-discriminative, despite a powerful impact on sick-listing [28,52,80,38]. *Non-severe functional impairment*, as measured by the Oswestry Disability Index [36,32], *Health-related quality of life*, according to EQ-5D [51] and *State of health*, as expressed by EQ VAS [51], were comparatively strong predictors in the univariate analyses, but non-predictors in the final multiple-logistic models. This is contrary to previous studies [36,164,32], for which we can offer no explanation.

Non-predictors that have previously been insufficiently studied

Many of our non-predictors that have been insufficiently studied in previous research might contribute to a widening of knowledge: *Non-immigrant*, *Co-habiting*, *Living without children*, *Non-unemployment*, *No work trauma litigation*, *Non-bad economy*, *Non-obese*, *No comorbidity*, *No surgery for spinal pain*, *Pain duration*, *Pain intensity*, *Local pain*, *Back-pain domination*, *High physical activity*, *Varied work tasks*, *No job strain*, *No depression/anxiety* and *No indications of alcohol over-consumption* [53].

Objective versus subjective variables

Our study strongly supports the predictive value of subjective predictors and might widen the knowledge of objective variables as non-predictors.

Treatment as a predictor of return-to-work

For the entire group of patients, treatment was non-predictive. However, a more detailed evaluation of the possible positive effect on return-to-work of our programme requires other analyses than in study IV – for example, survival analysis as in study III – and is a matter for future work.

STRENGTHS OF THE PROJECT

General strengths

One of the strengths of our project was the good representation of women.

The initial patient questionnaire was completed under the supervision of a research assistant, which might have contributed to a high quality of the patient data in studies I, III and IV, and increased the comparability between the patients and the non-patients and the patients mutually.

Because we used data from the SIO, no sick-listing data was missing, except the possible relapses of work absence < 14 days.

The use of reliable function tests is a major strength. One of the examiners in study II, the research assistant, also carried out the tests in study III and IV.

Specific strengths

Study I

The design of the nationwide ULF also allows local comparisons to be made. The responding rates of the ULF in 2000 and 2001 were practically 80%. These high-quality data concerning the comparison group was a strength.

Another strength of the study was the excellent model fit.

Study II

A strength was that all participants performed all tests.

Another strength was the complete collection of results, i.e., there were no missing data in the analyses.

Study III

The design, a randomized controlled trial, is the gold standard for evaluating treatment methods for spinal pain [117].

The health-care data was acceptably representative. The response rate for the questionnaire data was higher than 80% except at 12 months, when it was nearly 80% for the primary-care group. Even when the missing data for the 2 deceased patients were included, the rehabilitation group met drop-out criteria [166]. For the primary-care group, *Visits* over 18 months should be interpreted with some caution as 32% were

non-responders, but in other respects the follow-up rate of the primary-care group was also satisfactory. The non-responders in the rehabilitation group had characteristics that may have increased health-care use (longer sick-listing periods and higher unemployment). In the primary-care group the non-responders were younger, which could have decreased utilization, whereas the lower health-related quality of life could possibly increase utilization. However, for the great majority, there were no significant differences at baseline between the non-responders and responders.

Study IV

The prospective design, with a comparatively long follow-up period, is a major strength.

With the exclusion of 1 patient, also the questionnaire data were complete.

LIMITATIONS OF THE PROJECT

General limitations

Some circumstances might have decreased the representativeness of the patient sample of studies I, III and IV. The study population of 125 patients recruited over a period of 3.5 years constituted a very low percentage of the eligible subjects. As a comparison, Dionne et al [31] achieved a participation rate of 68.4% of eligible subjects. The inclusion was non-systematic: a family doctor with a local reputation of great skills in spinal pain might attract more complex cases, and have a higher motivation for research and the recruitment of study patients. Thus the patient sample might have been non-representatively complex, leading to spectrum bias [181]. We were overoptimistic concerning the recruiting propensity of the family doctors and lacked resources to increase the compliance. This contributed to a prolonged inclusion period that increased the risk of societal changes in rules and attitudes which might result in different return-to-work rates in identical NSP due to inclusion either early or late in the recruitment period. The problem with protracted inclusion is shared with other studies [92,104,72].

A closely-related limitation was the geographical imbalance in the recruitment; however, the greatest number of recruited patients were living in the district with the greatest number of inhabitants (Huddinge) (data not shown).

Specific limitations

Study I

A limitation was the non-prospective design. However, this study might contribute to a more detailed cross-sectional picture of the patients with non-acute NSP, which is also of value in the planning and interpretation of prospective research, e.g., study IV.

Study II

One limitation was that the gold standard consisted of 1 single physiotherapist. Also, the use of only 1 examiner without medical education is a limitation.

Another weakness was that the intra-reliability study only included a comparatively small number of healthy subjects. A way to overcome the ethical and methodological difficulties of using patients for as many as 3 examinations is to spread them out over

several days, as in the studies of Ljungquist et al [102] and Horneij et al [63]. This option, however, was beyond the limits of the resources of our study.

Study III

The inclusion plan was not fulfilled. A possible consequence may have been that some differences between the groups could not be demonstrated. However, certain differences in favour of the rehabilitation group were clear with the number of patients actually included.

Comparison of health-care visits gives only a limited idea of cost effectiveness. A complete health-economic evaluation includes a cost-benefit analysis in which the direct costs (mainly of the interventions themselves), the indirect costs (mainly of the sick-listing), and the health-related quality of life are compared [42]. This might be achieved in a future study.

The primary outcome measure showed no difference. Notwithstanding the positive trends in favour of the rehabilitation group, especially for the subacute patients, *Net days* and *Visits* were also equivalent over 18 months. As differences in the results of various interventions tend to even out after 12–18 months [72], more conclusive results might require a longer follow-up period than in study III.

Study IV

Work satisfaction as a separate variable was not included. Since it was indicated as a return-to-work predictor in several previous studies [164,147,81], it is a limitation.

CONCLUSIONS

The living conditions associated with long-term sick-listing of 124 patients with non-acute non-specific spinal pain were compared with 338 non-patients by applying logistic regression. In the univariate analyses, 13 of the 18 conditions had higher odds for the patients with a dominance of physical work strains and *Indication of alcohol over-consumption* (OR 14.8). Five conditions qualified for the multivariate analysis: *High physical workload* (OR 13.7), *Hectic work tempo* (OR 8.4), *Blue-collar job* (OR 4.5), *Obesity* (OR 3.5), and *Low education* (OR 2.7). As most of those living conditions have hitherto been insufficiently studied, our findings might help extend our knowledge of what distinguishes the individuals at risk for long-term sick-listing due to NSP. As the cross-sectional design makes causal conclusions impossible, our study should be complemented by prospective research.

Given a 10-test package for patients with prolonged back and neck pain, an examiner without formal medical education could be used without loss of quality, at least for the 5 tests that require no manual fixation. This might produce a better assessment of outcome at defensible cost and might also be useful in a research context. To make our results more generalizable and their implications more searching, a similar study should be conducted with 2 or more examiners with and without formal medical education, and the intra-rater reliability study should also include patients and involve more participants.

For patients with subacute and chronic NSP, cognitive-behavioural rehabilitation was compared with primary care. The results were equivalent over 18 months. However, there were indications that cognitive-behavioural rehabilitation in the longer

run might be superior. For subacute NSP, in terms of both sick-listing and health-care visits; for chronic NSP, in terms of health-care visits only. More conclusive results concerning this possible long-term effect might require a longer follow-up.

In primary-care patients with non-acute, non-specific spinal pain, the strong predictors of stable return-to-work were 2 socioeconomic variables, *Low total previous sick-listing* (including all diagnoses) and *Young age* (max 44 years), and 1 subjective variable, *High self-prediction* (the patients' own belief in return-to-work). Objective variables from function tests and treatment variables (a programme of cognitive-behavioural rehabilitation or traditional primary care) were non-predictors. Except for *Young age*, the predictors had been insufficiently studied in previous research. Hence, our study might contribute to a widening of knowledge within clinical practice, including the allocation of treatment resources.

CLINICAL IMPLICATIONS

Our project might help family doctors, supervisors in the work place, handling officers of the SIO, etc, to identify subjects at risk for long-term disability.

Our findings that medically untrained examiners could be used in function tests for patients with non-acute NSP, at least in function tests not requiring manual fixation, might contribute a better assessment at defensible cost in, e.g., physiotherapy and rehabilitation centres and within research.

Non-acute NSP might be difficult to handle by traditional care. In view of its possibly better effect in the long run, at least for subacute patients, our project indicates that referral for our concept of rehabilitation programme might be considered for those patients by family doctors, physiotherapists, psychologists and social workers in primary care.

A PERSONAL REFLECTION

Since the radical change of the sick-listing rules in 2008, the rehabilitation of patients with prolonged disorders, not least NSP, has been substantially complicated, both for the patients and those who, like the author of this thesis, are working within it. In my experience, it is now the rule, not the exception, that the interventions by SIO are counterproductive and in practice prolong the disabled citizen's way back to decent social functioning. Strangely enough, the change took place when sick-listing was already decreasing rapidly (Figure 2).

The shortcomings in the Swedish social insurance system came to a head in 2003. Concerning restrictiveness, the system now works at the other extreme: it has gone from doing too little, too late, to doing too much, too soon.

Really, the rehabilitation of non-acute non-specific spinal pain is not just a matter for physiotherapists, psychologists and doctors [178]. Before Sweden attains a balanced social insurance system, designed from evidence-based and human principles, that rehabilitation, in my opinion, will remain insufficient.

FUTURE RESEARCH

A 2-year follow-up of study III, including survival analyses, is planned.

We plan a complete health-economic evaluation of the 2 treatment alternatives (study III). This is enabled by the satisfactory responding rates of all follow-ups (6, 12, 18 and 24 months), and the detailed information of health-care consumption which is given from the follow-up questionnaires.

The 10-test package (study II) was performed with the research assistant as examiner also 1 year after baseline. The participation rate was fairly good, about 70%. We plan to investigate the validity of the package, i.e., if the 6 reliable tests have some correlations to socioeconomic, subjective and treatment variables at 1 year.

High self-prediction was one of the strong predictors in study IV. We plan a closer analysis of the variable. For example, does this variable have an impact on health-care consumption?

ACKNOWLEDGEMENTS

Where to begin?

Well, in this very moment my cohabitant *Caroline* and our little angle *Gunvor* are waiting for daddy. Believe me: I am waiting for you! My non-acute, non-specific spinal pain is near recovery. *Carolinska Institutet* will evict its name: our home! Without you, the pain would have persisted.

My son *Axel*, flying for himself for many years now, promised me once to never become a physician. Behind his promise, among other reasons, is daddy's NSP. However, from being a philologist, he recently received his nurse diploma...

My mother *Inga-Britt*, always there! Despite the agnostic delusions of the youngest son of a former vicar (In Swedish: kyrkoherde emerita), she has always asked our Lord to give me a chance. Mother: you succeeded!

All you who have passed before me: daddy *Allan*, bonus daddy *Tord*, grandmother *Millan* and grandfather *Helmer*, *Gunvor* (the original one!), *Anna-Karin*, *Stig*, *Sverker Rune* and *Rut*... Look down and enjoy a success which is also yours, in this moment I am looking up at you!

My supervisor *Lars-Erik*! You have never failed me. When we first met (in the beginning of time!), you told me: "A researcher should express himself shortly, literally and objectively, in other words, in precisely the ways you do not!" This penetrating vision made me roll up the sleeves. With your inspiration, sometimes artificial respiration and our mutual transpiration: we made it!

Sven-Erik, the indestructible rock in the statistical mine field!

The personnel of the Jordbro health centre, where it all began in our heads, and of the STRONG unit, where it all began in reality, especially *Ingegerd Eklund* (the research assistant); the staff of Haninge FysioCenter next door, especially *Lars Eriksson* (co-author of study II) and *Ragnar Faleij* (co-talker and co-thinker); the late 1980th and early 1990th directors of the primary care (especially *Lars-Bertil Arvidsson*, the other 'head' of the original spinal project); the Social Insurance Office and the municipality of Haninge (those were the days!)

Rod Stewart once told us in a Globen performance: what an audience! As an old amateur actor, I will tell all the coworkers in the Center for Family and Community Medicine: what an institution!

Last but not least: I will thank all the patient patients!

The project was supported by grants from the Stockholm County Social Insurance Office, Stockholm County Council, Ministry of Health and Social Affairs and Vårdal Foundation.

SUMMARY IN SWEDISH / SAMMANFATTNING PÅ SVENSKA

Rehabilitering av långvarigt ryggont

Vetenskapliga referenser inom hakparentes hänvisar till referenslistan. Figurerna finns i den föregående engelska texten.

BAKGRUND

Smärta

Smärta är det vanligaste skälet att söka vård [174]. Genom sitt starka obehag driver den akuta smärtan oss på ett effektivt sätt till beteenden som skyddar det skadade området så att vävnaden kan läka och ny skada förebyggas. Obehaget frigör stresshormoner som ger ångestliknande symptom: hjärtklappning, andfåddhet och handsvett. Akut smärta och ångest är nära besläktade [174].

Smärtsignalerna från periferin moduleras fortlöpande av det datorliknande centrala nervsystemet. Smärta, känslor och beteenden är integrerade och verkar i båda riktningar: *smärtan påverkar beteendet* och *beteendet påverkar smärtan* [134].

Kronisk smärta indikerar en varaktighet av minst 3 månader [174]. När den väl uppstått är chanserna till smärtfrihet mycket små [5]. Motsatt akut smärta förlorar den kroniska smärtan sin biologiska mening och blir kontraproduktiv. Aktivitetsdrivet ersätts av depressionsliknande passivitet, hopplöshet och att man drar sig undan från sociala aktiviteter. Kronisk smärta och depression hör ihop [174].

Frisk och sjuk sjukskrivning

Smärta kan leda till arbetsoförmåga. En förutsättning för ens försörjning vid arbetsoförmåga pga sjukdom är sjukskrivning. Sjukskrivning, kronisk smärta och psykiska besvär, framför allt ångest, depression och stressbesvär, samverkar [46]. *Frisk sjukskrivning*, dvs en aktiv sjukskrivning med vetenskapligt vedertagna insatser, är en investering i framtida stabil arbetsförmåga. En passiv sjukskrivning, utan aktiva åtgärder, förstärker den psykiska frustrationen, vilket ger än mer smärta och arbetsoförmåga i en ond cirkel (Figur 1) [176]. Problemet med sådan *sjuk sjukskrivning* beskrevs tidigt [161,170]. Sverige är inget undantag.

Den svenska sjukskrivningen

Sedan 1992 har i Sverige arbetsgivaren ansvaret för att sjukskrivna anställdas rehabiliteringsbehov uppmärksammas och tillgodoses; Försäkringskassan (FK) har ansvaret för rehabiliteringens övergripande samordning [9]. *Rehabilitering* är processen som gör att en individ med en arbetshindrande sjukdom eller skada kan återfå arbetsförmågan. Detta kan inkludera medicinsk behandling, arbetsträning och andra arbetslivsinriktade åtgärder [178]. Således kräver optimal rehabilitering insatser inte

bara av t ex läkare och sjukgymnast, utan i (minst) lika hög grad FK, arbetsgivaren och vid behov Arbetsförmedlingen [9].

Från slutet av 1980-talet påtalades den bristande samordningen i flera statliga utredningar [33,44,84] och forskningsstudier [109,110,141,142,71,35,111,7,143,8,9,90]. Dock fördjupades bristerna och från 1998 ökade sjukskrivningen lavinartat till ett paradoxalt maximum 2003 då svenskarna var mest sjukskrivna i Europa, sannolikt i världen, och samtidigt bland de friskaste (Figur 2) [57,91]. Riksrevisionen riktade svidande kritik mot FK, som bl a 4 av 5 år återbetalade statsmedel FK fått för köp av rehabilitering, men inte hunnit göra av med [155]; de år sjukskrivningen ökade som snabbast halverade FK sina aktiva åtgärder [9]; frånvaron av portvaktfunktion var tydlig: det var alldeles för lätt att komma in i passiv och dyrbar sjukskrivning [151]. Ett regeringsprojekt startade 2002 med fokus på ökad restriktivitet och målet att halvera sjukskrivningen [40], varefter den sjönk till genomsnittlig Europa-nivå [10]. 2008-07-01 ändrades FKs regler i mycket restriktiv riktning: ”Efter 90 dagars sjukskrivning har du bara rätt till sjukpenning om du inte kan utföra något arbete alls hos din arbetsgivare. Efter 180 dagar...bara... om du inte kan utföra något arbete alls....” [145].

Ryggont

Den vanligaste formen av kronisk smärta består av rygg- och/eller nackbesvär [46]. > 95 % av rygg och/eller nackbesvär är ospecifika och kallas här för *ryggont*. Ospecifik innebär att besvären inte kräver insats av sjukhusspecialist, t ex ortoped eller neurolog, utan bäst behandlas i primärvården, hos husläkaren, eller inte alls [173].

Ryggont är en av de vanligaste orsakerna till sjukskrivning, särskild långvarig sådan [120]. De flesta tillfrisknar snabbt [26]. Efter heltids sjukskrivning 1 vecka har hälften och efter 12 veckor 90 % återgått i jobb. Sedan planar förbättringstakten av betydligt (Figur 5).

Den totala samhällskostnaden för ryggont är mycket stor. Den har beräknats till drygt 29 miljarder kronor årligen eller nästan dubbla kostnaden för sjuk- och socialvård [156]. 92 % är sk indirekta kostnader, som utgörs framför allt av sjukskrivning. Den allra minsta kostnaden, 0,4 %, är för rehabilitering (Figur 4) [124,156].

Behandling av ryggont

Akut ryggont, definierat som ryggont ledande till heltids sjukskrivning 1–21 dagar (3 veckor) [176], går oftast snabbt över av sig självt, särskilt om man fortsätter sina vanliga vardagsaktiviteter så normalt som möjligt och möjligen tar hjälp av manuell behandling (muskeltöjning, manipulation, etc) [177].

Vid *långvarigt ryggont*, dvs ledande till heltids sjukskrivning längre än 3 veckor [176], kan multidisciplinär rehabilitering övervägas [177]. *Multidisciplinär* innebär insats av läkare tillsammans med t ex sjukgymnast och psykolog. Viktiga komponenter i sådana rehabprogram är gradvis ökning av fysisk och mental aktivitet, modifiering av negativa tankar och beteenden genom kognitiv beteendeterapi (KBT) samt direkt arbetslivsinriktade åtgärder [178]. *KBT* bygger på att *lära* och att *lära om* (dvs ersätta mindre bra beteenden, t ex fysisk passivitet, med bra beteenden, t ex fysisk aktivitet) [97].

Den långvariga fasen delas upp i *subakut ryggont* och *kroniskt ryggont* = heltids sjukskrivning 22–84 dagar (12 veckor) respektive mer än 12 veckor (3 månader) [176]. Under den subakuta fasen kommer det allra mesta av möjligheterna till spontantillfrisknande att tömmas ut samtidigt som rehabpotentialen ännu är god. I den kroniska fasen är ofta den sjukskrivna mera fast i onda cirklar av passivitet och nedstämdhet och ser mera pessimistiskt på framtiden [176]. De olika faserna ses i Figur 5.

Vårt projekt

Klinisk och vetenskaplig utgångspunkt

Den kliniska utgångspunkten var en rehabenhet för patienter med långvarigt ryggont (STRONG-enheten) som verkade i Haninge, en kommun 2 ½ mil sydöst om Stockholms city, 1991–2006. Fr o m 1996 använde enheten ett KBT-program med målet att återfå så hög och stabil arbetsförmåga som möjligt. Efter 10 år inom landstinget övergick enheten 2002 i privat regi och därmed minskades antalet rehabteam från 4 till 1: en läkare (undertecknad), en sjukgymnast, en psykolog eller socionom med utbildning i KBT samt en friskvårdskonsulent. 2006 stängdes enheten pga minskad efterfrågan på rehabilitering från FK och arbetsgivare.

Vi ville belysa rehabilitering av långvarigt ryggont ur olika aspekter. Den vetenskapliga kärnan blev en studie (randomiserad kontrollerad studie) som försiggick 2000–2006 med målet att jämföra rehabprogrammet med traditionell primärvård (studie 3). Sjukgymnasterna på enheten använde ett testpaket med 10 deltester för att bedöma patienternas funktionsnedsättning. 5 av testerna krävde att testledaren höll ett fast tag mot underlaget av de kroppsdelar som inte var avsedda att röra sig. Denna manuella fixering gjordes för att eliminera missvisande medrörelser. För att t ex kunna bedöma rörligheten i halsryggen, krävdes en fixering mot stolsryggen av den sittande patientens bröst- och ländrygg. Testpaketet användes även i studie 3 av vår forskningsassistent, som hade en universitetsexamen, men ingen medicinsk utbildning. För att undersöka tillförligheten (reliabiliteten) i detta gjordes studie 2 (metodologisk studie), som, vad vi vet, var den första reliabilitetsstudie som gjordes med en testledare utan medicinsk utbildning. En tanke var också att utvärderingen av funktionshöjande rehabinsatser skulle kunna bli mer objektiv om den utfördes av någon utanför själva behandlingsarbetet. Eftersom det är ekonomiskt orealistiskt med medicinskt utbildad personal endast för utvärdering, skulle det också kunna vara kostnadseffektivt med en testledare utan medicinsk utbildning. Data från studie 3 och de tester som i studie 2 visat sig vara tillförlitliga användes sedan i studie 4 (prospektiv kohortstudie). Data från studie 3 kunde också återanvändas i studie 1 (tvärsnittstudie).

Mål

Det övergripande målet med projektet var att belysa rehabilitering av långvarigt ryggont ur 4 olika aspekter, epidemiologi (läran om sjukdomsförekomst), reliabilitet (tillförlitlighet), behandling samt prediktion (försägelse) av framtida arbetsförmåga.

Metoder

Studie 1 (epidemiologi): Levnadsförhållanden med eventuellt samband med långvarig sjukskrivning för ryggont jämfördes mellan de 125 patienterna från studie 3 och 338

slumpvis utvalda ej sjukskrivna personer (icke-patienter) från samma hemkommuner och i samma ålder som patienterna (18-59 år).

Studie 2 (reliabilitet): Utförandet av de 10 funktionstesterna med en erfaren sjukgymnast som testledare jämfördes med den icke-medicinskt utbildade forskningsassistenten som testledare. Deltagare var 30 patienter med långvarigt ryggont från en sjukgymnastenhets som låg vägg i vägg med rehabenheten och 20 friska försökspersoner (personal på enheterna).

Studie 3 (behandling): 125 patienter med långvarigt ryggont fördelades slumpmässigt (randomiserades) till programmet vid rehabenheten (rehabiliteringsgruppen = 62 stycken) eller till fortsatt sedvanlig primärvård (primärvårdsgruppen = 63 stycken). De följdes sedan i 1 ½ år. Resultaten analyserades både för hela perioden och var och en av dess tre 6-månadersperioder och både för hela rehabiliteringsgruppen och primärvårdsgruppen och de subakuta patienterna för sig (20 och 18 i respektive rehabiliterings- och primärvårdsgruppen) och de kroniska för sig (42 och 45 i respektive grupp).

Studie 4 (prediktion): Faktorer som skulle kunna förutsäga stabil (= som varar i minst 1 månad) arbetsförmåga samlades in från de 125 patienternas basdata. Det var objektiva faktorer (resultaten på funktionstesterna vid studiens början), socioekonomiska faktorer (yrke, ålder, familjeförhållanden, sjukskrivning, livsstil osv), subjektiva faktorer (patientens uppgivna grad av smärta, psykologiska reaktioner på smärtan, etc, samt patientens egen grad av tro på att återfå arbetsförmågan) samt behandling (KBT-rehabilitering eller sedvanlig primärvård).

Statistiska metoder: Förutom deskriptiv statistik (*t*-test och *z*-test)[1,2], användes i studie 1 och 4 multipel logistisk regression [64], i studie 2 one-way ANOVA intra-class correlation coefficient [48] och i studie 3 Cox regression [77] och mixed linear models [24].

Resultat

Studie 1: 13 av 18 levnadsförhållandena som kan ha samband med långtidssjukskrivning var vanligare hos patienterna än hos icke-patienterna, med en överrepresentation av fysiskt ansträngande arbetsmoment samt indikation på alkoholmissbruk (8 gånger vanligare hos patienterna, 13,7 % jämfört med 1,7 % hos icke-patienterna). När faktorerna statistiskt vägdes samman, kvarstod 5 förhållanden som var vanligare hos patienterna: Fysiskt tunga arbetsmoment, hektiskt arbetstempo, arbetarjobb (till skillnad mot tjänstemannajobb), svår övervikt och låg utbildning. Svår övervikt, dvs BMI ≥ 30 , var 3 ggr vanligare hos patienterna, 24,2 % jämfört med 6,9 % hos icke-patienterna. Också låg utbildning, dvs som högst fullgjord grundskola, var 3 gånger vanligare, 35,5 % jämfört med 12,1 %.

Studie 2: Alla 5 tester som inte krävde manuell fixering hade acceptabel tillförlitlighet. Alla 5 testerna som krävde manuell fixering var otillförlitliga utom 1. Skillnaden (5 mot 1) var statistiskt signifikant.

Studie 3: Alla patienter analyserade tillsammans och för hela observationstiden på 18 månader: Andelen arbetsförmögna, chansen till arbetsförmåga, total sjukskrivning och antalet vårdbesök skiljde sig inte. Vid delanalys för var och en av de tre 6-månadersperioderna var dock minskningstrenden signifikant starkare för rehabiliteringsgruppens totala sjukskrivning och antal besök. Chansen till arbetsförmåga var signifikant högre för de subakuta rehabiliteringspatienterna jämfört

med de subakuta primärvårdspatienterna medan chansen var likvärdig mellan de kroniska patienterna i båda grupperna.

Studie 4: Tre (av sammanlagt 50) faktorer, mätta vid baslinjen, kunde förutse arbetsåtergång under de kommande 2 åren: låg total sjukskrivning (dvs sjukskrivning för alla diagnoser) de 2 åren före baslinjen, hög självprediktion (dvs patientens egen tro på att komma åter i jobb) och låg ålder (≤ 44 år). Objektiva faktorer och behandling predikerade inte arbetsåtergång.

Slutsatser och klinisk användbarhet

Epidemiologi och prediktion: Eftersom de flesta av de studerade levnadsförhållandena hittills varit ofullständigt utforskade, kan vårt projekt utöka kunskapen om vad som utmärker individer med framtida risk för långvarig sjukskrivning för ryggont. Detta kan vara till hjälp för t ex husläkare, arbetsledare och FK-handläggare.

Reliabilitet: För ett testpaket med 10 funktionstester kan användas en testledare utan medicinsk utbildning, åtminstone för de 5 tester som inte kräver någon manuell fixering. Resultatet kan bidra till förbättrad kvalitetssäkring till begränsad kostnad för t ex sjukgymnast- och rehabiliteringsenheter och även vara av värde inom forskningen.

Behandling: Fastän resultaten var likvärdiga för hela 18-månadersperioden, fanns indikationer på att vår typ av KBT-rehabilitering på längre sikt kan vara bättre än primärvård. För subakuta patienter vad gäller både sjukskrivning och antalet besök, för kroniska patienter vad gäller vårdkonsumtion. Mer slutgiltig värdering av denna eventuella långtidseffekt kräver dock längre uppföljningstid än 18 månader. Komplexiteten i långvarigt ryggont är ofta svår att handlägga inom den vanliga vården. Givet att KBT-rehabilitering av vår typ långsiktigt är bättre än traditionell primärvård, kan vårt projekt bidra till större satsningar än hittills på sådan verksamhet. Detta vore till hjälp för t ex husläkare, sjukgymnaster, psykologer och kuratorer inom primärvården.

En personlig reflektion

Sen den genomgripande ändringen av sjukskrivningsreglerna 2008-07-01 har rehabiliteringen av patienter med långvariga besvär, inte minst ryggont, försvårats betydligt, både för patienterna och dem, som liksom undertecknad, arbetar med sådan rehabilitering. Enligt min erfarenhet är det nu regel snarare än undantag att FKs handläggning av dessa patienter är kontraproduktiv och i praktiken förlänger den sjukskrivnas väg tillbaka till hygglig social funktion. Märkligt nog genomfördes förändringen när sjukskrivningen redan var i snabb minskning (Figur 2).

Tillkortakommandena i det svenska socialförsäkringssystemet nådde en ytterkant 2003, men har nu i restriktivitet kantrat till den andra extremen: från att göra för litet, för sent, till att göra för mycket, för tidigt.

Rehabiliteringen av långvarigt ryggont är som sagt inte bara en fråga för sjukgymnaster, psykologer och läkare. Innan Sverige får ett välavvägt, ”lagom” restriktivt socialförsäkringssystem, utformat efter vetenskapliga och humanistiska principer, kommer denna rehabilitering, enligt min uppfattning, att förbli bristfällig.

REFERENCES

1. Altman DG: **Comparing groups – categorical data.** In *Practical statistics for medical research*. 1st edition. London: Chapman & Hall; 1991:229-276.
2. Altman DG: **Comparing groups – continuous data.** In *Practical statistics for medical research*. 1st edition. London: Chapman & Hall; 1991:179-228.
3. **Analysis and Prognosis. Press message** [In Swedish: Analys och Prognos. Pressmeddelande.] 2010-09-20. Social Insurance Agency [In Swedish; Försäkringskassan].
[<http://www.forsakringskassan.se/irj/servlet/prt/portal/prtroot/docs/guid/e09e302b-cea6-2d10-6c97-abf57c439c4b>].
4. Andersson HI, Ejlertsson G, Leden I, Rosenberg C: **Chronic pain in a geographically defined general population: studies of differences in age, gender, social class, and pain localisation.** *Clin J Pain* 1993;174-182.
5. Andersson HI: **The course of non-malignant chronic pain: a 12-year follow-up of a cohort from the general population.** *Eur J Pain* 2004;8:47-53.
6. Anema JR, Steenstra IA, Bongers PM, de Vet HC, Knol DL, Loisel P, van Mechelen W: **Multidisciplinary rehabilitation for subacute low back pain: graded activity or workplace intervention or both? A randomized controlled trial.** *Spine* 2007;32:291-298.
7. Arrelöv B, Borgquist L, Ljungberg D, Svärdsudd K: **Do GPs sick-list patients to a lesser extent than other physician categories? A population-based study.** *Fam Pract* 2001;18:393-398.
8. Arrelöv B, Borgquist L, Ljungberg D, Svärdsudd K: **The influence of change of legislation concerning sickness absence on physicians' performance as certifiers. A population-based study.** *Health Policy* 2003;63:259-268.
9. Bergendorff S: **Rehabilitation – the history of a long period of suffering** [In Swedish; Rehabilitering – ett långt lidandes historia]. Commission of Inquiry on Social Insurance [In Swedish; Socialförsäkringsutredningen]. *Samtal om socialförsäkring* 2006:10
[http://www.kfs.net/projekt/dokument/SOU_lidande_rehab.pdf].
10. Bergendorff S: **Sickness absence in Sweden – on the way towards European levels?** [In Swedish; Sjukfrånvaron i Sverige – på väg mot europeiska nivåer?]. Summary in English. Social Insurance Report [In Swedish; Socialförsäkringsrapport]. 2009:10 ISSN 1654-8574
[http://www.forsakringskassan.se/irj/go/km/docs/fk_publishing/Dokument/Rapporter/socialforsakringsrapporter/socialforsakringsrapport_2009_10.pdf]
11. Berglund M (chairman). **Summary and conclusions** [In Swedish; Sammanfattning och slutsatser]. In: Treatment of alcohol and drug abuse – an evidence-based review [In Swedish; Behandling av alkohol- och narkotikaproblem – en evidensbaserad kunskapssammanställning]. The Swedish Council on Technology Assessment in Health Care. ISBN: 91-87890-73-9. Stockholm 2001:11-28. Summary in English:
[http://www.sbu.se/upload/Publikationer/Content0/1/alcoholdrug_2001/drugslut.html] A

12. Bergström G, Bodin L, Bertilsson H, Jensen IB: **Risk factors for new episodes of sick leave due to neck or back pain in a working population. A prospective study with an 18-month and a three-year follow-up.** *J Occup Environ Med* 2007;64:279-287.
13. Bergström G, Bodin L, Hagberg J, Lindh T, Aronsson G, Josephson M: **Does sickness presenteeism have an impact on future general health?** *Int Arch Occup Environ Health* 2009;82:1179-1190.
14. Bergström G, Bodin L, Hagberg J, Aronsson G, Josephson M: **Sickness presenteeism today, sickness absenteeism tomorrow? A prospective study on sickness presenteeism and future sickness absenteeism.** *J Occup Environ Med* 2009;51:629-638.
15. Bertilson BC, Grunnesjö M, Strender LE: **Reliability of clinical tests in the assessment of patients with neck/shoulder problems – impact of history.** *Spine* 2003;28:2222-2231.
16. Bertilson B, Grunnesjö M, Johansson SE, Strender LE: **Pain Drawing in the Assessment of Neurogenic Pain and Dysfunction in the Neck/Shoulder Region: Inter-Examiner Reliability and Concordance with Clinical Examination.** *Pain Medicine* 2007;8:134-146.
17. Biering-Sørensen F: **Physical measurements as risk indicators for low-back trouble over a one-year period.** *Spine* 1984;9:106-109.
18. Biering-Sørensen F, Thomsen C: **Medical, social and occupational history as risk indicators for low-back trouble in a general population.** *Spine* 1986;11:720-725.
19. Binkley J, Stratford PW, Gill C: **Interrater reliability of lumbar accessory motion mobility testing.** *Phys Ther* 1995;75:786-795.
20. Borg G: **Psychophysical bases of perceived exertion.** *Med Sci Sports Exerc* 1982;14:377-382.
21. Borg G: **A category scale with ratio properties for inter-modal and inter-individual comparison.** Berlin: VEB Deutscher Verlag der Wissenschaften; 1982:25-34.
22. Borg K, Hensing G, Alexandersson K: **Predictive factors for disability pension – an 11-year follow up of young persons on sick leave due to neck, shoulder or back diagnoses.** *Scand J Public Health* 2001;29:104-112.
23. **Brief history and other relevant background information.** Country Profile of Sweden. United Nations Statistical Division.
[<http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=666#start>]
24. Brown H, Prescott R: **Applied mixed models in medicine.** Edinburgh: John Wiley & Sons LTD; 2001:199-259.
25. Capuano-Pucci D, Rheault W, Aukai J, Bracke M, Day R, Pastrick M: **Intratester and intertester reliability of the cervical range of motion device.** *Arch Phys Med Rehabil* 1991;72:338-340.
26. Carrey TS, Garrett J, Jackman A, McLaughlin C, Fryer J, Smucker DR: **The outcomes and costs of care for acute low back pain among patients seen by primary care practitioners, chiropractors, and orthopedic surgeons: the North Carolina Back Pain Project.** *N Eng J Med* 1995;333:913-917.
27. Chinn S: **Statistics in respiratory medicine. 2. Repeatability and method comparison.** *Thorax* 1991;46:454-456.

28. Dasinger LK, Krause N, Deegan LJ, Brand RJ, Rudolph L: **Physical workplace factors and return to work after compensated low back injury: a disability phase-specific analysis.** *J Occup Environ Med* 2000;42:323-333.
29. DeLeo JA, Winkelstein BA: **Physiology of chronic spinal pain syndromes: from animal models to biomechanics.** *Spine* 2002;27:2526-2537.
30. Demyttenaere K, Bruffaerts R, Lee S, Posada-Villa J, Kovess V, Angermeyer MC, Levinson D, de Girolamo G, Nakane H, Mneimneh Z, Lara C, de Graaf R, Scott KM, Gureje O, Stein DJ, Haro JM, Bromet EJ, Kessler RC, Alonso J, Von Korff M: **Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys.** *Pain* 2007;129:332-342.
31. Dionne CE, Bourbonnais R, Frémont P, Rossignol M, Stock SR, Larocque I: **A clinical return-to-work rule for patients with back pain.** *CMAJ* 2005;172:1559-1567.
32. Du Bois M, Szpalski M, Donceel P: **Patients at risk for long-term sick leave because of low back pain.** *Spine* 2009;9:350-359.
33. **Early and coordinated rehabilitation.** [In Swedish; Tidig och samordnad rehabilitering.] Swedish Government Official Reports. [In Swedish; Statens offentliga utredningar (SOU).] SOU 1988:41. [<http://www.biblioteket.stockholm.se/default.asp?id=8227&extras=242161%2FID>]
34. Eklund M, Fugl-Meyer AR: **Vocational rehabilitation in northern Sweden. II. Some psycho-socio-demographic predictors.** *Scand J Rehab Med* 1991;23:73-82.
35. Englund L, Svärdsudd K: **Sick-listing habits among general practitioners in a Swedish county.** *Scand J Prim Health Care* 2000;18:81-86.
36. Fairbank JC, Couper J, Davies JB, O'Brien JP: **The Oswestry low back pain disability questionnaire.** *Physiotherapy* 1980;66:271-273.
37. Fjellner A, Bexander C, Faleij R, Strenger LE: **Interexaminer reliability in physical examination of the cervical spine.** *J Manipulative Physiol Ther* 1999;22:511-516.
38. Fransen M, Woodward M, Norton R, Coggan C, Dawe M, Sheridan N: **Risk factors associated with the transition from acute to chronic occupational back pain.** *Spine* 2002;27:92-98.
39. Ghaffari M, Alipour A, Farshad AA, Jensen I, Josephson M, Vingard E: **Effect of psychosocial factors on low back pain in industrial workers.** *Occup Med (Lond)* 2008; 58:341-347.
40. **Goals of public health** [In Swedish; Mål för folkhälsan]. *Governmental bill* 2002/03:35 [<http://www.riksdagen.se/webbnav/index.aspx?nid=37&rm=2002/03&bet=35&typ=prop>].
41. Goldberg MS, Scott SC, Mayo NE: **A review of the association between cigarette smoking and the development of nonspecific back pain and related outcomes.** *Spine* 2000;25:995-1014.
42. Goosens M, Evers S: **Cost-effectiveness of treatment for neck and low back pain.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:399-419.

43. Grunnesjö MI, Bogefeldt JP, Svärdsudd KF, Blomberg SI: **A randomized controlled clinical trial of stay-active care versus manual therapy in addition to stay-active care: functional variables and pain.** *J Manipulative Physiol Ther* 2004;27:431-441.
44. Grönwall L: **The rehabilitation responsibility of the employer.** [In Swedish; Arbetsgivarens rehabiliteringsansvar.] Swedish Government Official Reports. [In Swedish; Statens offentliga utredningar (SOU).] SOU 1998:104 [http://www.regeringen.se/content/1/c6/02/42/68/a282e77e.pdf].
45. Gureje O, Von Korff M, Simon GE, Gater R: **Persistent pain and well-being: a World Health Organization study in primary care.** *JAMA* 1998;280:147-151.
46. Gureje O, Simon GE, von Korff M: **A cross-national study of the course of persistent pain in primary care.** *Pain* 2001;92:195-200.
47. Guzman J, Haldeman S, Carroll LJ, Carragee EJ, Hurwitz EL, Peloso P, Nordin M, Cassidy JD, Holm LW, Côté P, van der Velde G, Hogg-Johnson S; Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: **Clinical practice implications of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: from concepts and findings to recommendations.** *Spine* 2008;33(4 Suppl):S199-213.
48. Haas M: **Statistical methodology for reliability studies.** *J Manipulative Physiol Ther* 1991;14:119-132.
49. Hagen KB, Holte HH, Tambs K, Bjerkedal T: **Socioeconomic factors and disability retirement from back pain: a 1983-1993 population-based prospective study in Norway.** *Spine* 2000;25:2480-2487.
50. Haldorsen EM, Grasdal AL, Skouen JS, Risa AE, Kronholm K, Ursin H: **Is there a right treatment for a particular patient group? Comparison of ordinary treatment, light multidisciplinary treatment, and extensive multidisciplinary treatment for long-term sick-listed employees with musculoskeletal pain.** *Pain* 2002;95:49-63.
51. Hansson E, Hansson T, Jonsson R: **Predictors for work ability and disability in men and women with low-back or neck problems.** *Eur Spine J* 2006;15:780-793.
52. Hansson TH, Hansson EK: **The effects of common medical interventions on pain, back function, and work resumption in patients with chronic low back pain: A prospective 2-year cohort study in six countries.** *Spine* 2000;25:3055-3064.
53. Hansson T, Jensen I: **Sickness absence due to back and neck disorders.** In *Sickness absence – causes, consequences, and physicians' sickness-certification practice.* Edited by Alexanderson K and Norlund A. *Scand J Public Health* 2004;32 (Suppl 63):109-151.
54. Harcombe H, McBride D, Derrett S, Gray A: **Physical and psychosocial risk factors for musculoskeletal disorders in New Zealand nurses, postal workers and office workers.** *Inj Prev* 2010;16:96-100.
55. Harkness EF, Macfarlane GJ, Silman AJ, McBeth J: **Is musculoskeletal pain more common now than 40 years ago?: two population-based cross-sectional studies.** *Rheumatology* 2005;44:890-895.
56. Hawk C, Phongphua C, Bleecker J, Swank L, Lopez D, Rubley T: **Preliminary study of the reliability of assessment procedures for indications for**

- chiropractic adjustments of the lumbar spine. *J Manipulative Physiol Ther* 1999;22:382-389.
57. **Health status: life expectancy.** In *Health at a glance 2009. OECD indicators*. December 8 2009. Picture 4
[<http://www.oecd.org/dataoecd/24/8/44231736.ppt#262>].
 58. Heuch I, Hagen K, Heuch I, Nygaard Ø, Zwart JA: **The impact of body mass index on the prevalence of low back pain: the HUNT study.** *Spine* 2010;35:764-768.
 59. Heymans MW, de Vet HC, Bongers PM, Knol DL, Koes BW, van Mechelen W: **The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial.** *Spine* 2006;31:1075-1082.
 60. Hlobil H, Staal JB, Twisk J, Köke A, Ariëns G, Smid T, van Mechelen W: **The effects of a graded activity intervention for low back pain in occupational health on sick leave, functional status and pain: 12-month results of a randomized controlled trial.** *J Occup Rehabil* 2005;15:569-580.
 61. Holmberg SA, Thelin AG: **Primary care consultation, hospital admission, sick leave and disability pension owing to neck and low back pain: a 12-year prospective cohort study in a rural population.** *BMC Musculoskelet Disord* 2006;7:66.
 62. Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM: **Systematic review of psychosocial factors at work and private life as risk factors for back pain.** *Spine* 2000;25:2114-2125.
 63. Horneij E, Holmström E, Hemborg B, Isberg PE, Ekdahl C: **Inter-rater reliability and between-days repeatability of eight physical performance tests.** *Adv Physiother* 2002;4:146-160.
 64. Hosmer DW, Lemeshow S: *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons; 2000.
 65. Hyttiäinen K, Salminen JJ, Suvitie T, Wickström G, Pentti J: **Reproducibility of nine tests to measure spinal mobility and trunk muscle strength.** *Scand J Rehabil Med* 1991;23:3-10.
 66. IJzelenberg W, Burdorf A: **Risk factors for musculoskeletal symptoms and ensuing health care use and sick leave.** *Spine* 2005;30:1550-1556.
 67. Iles RA, Davidson M, Taylor NF: **Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review.** *Occup Environ Med* 2008;65:507-517.
 68. Indahl A, Haldorsen EH, Holm S, Reikeras O, Ursin H: **Five-year follow-up study of a controlled clinical trial using light mobilization and an informative approach to low back pain.** *Spine* 1998;23:2625-2630.
 69. Janke EA, Collins A, Kozak AT: **Overview of the relationship between pain and obesity: What do we know? Where do we go next?** *JRRD* 2007;44:245-262.
 70. Jensen IB, Linton SJ: **Coping strategies questionnaire (CSQ): reliability of the Swedish Version of the CSQ.** *Scand J Behaviour Therapy* 1993;22:139-145.
 71. Jensen IB, Bodin L, Ljungquist T, Bergström G, Nygren Å: **Assessing the needs of patients in pain: a matter of opinion?** *Spine* 2000;25:2816-2823.

72. Jensen IB, Bergstrom G, Ljungquist T, Bodin L: **A 3-year follow-up of a multidisciplinary rehabilitation programme for back and neck pain.** *Pain* 2005;115:273-283.
73. Jensen I, Harms-Ringdahl K: **Strategies for prevention and management of musculoskeletal conditions. Neck pain.** *Best Pract Res Clin Rheumatol* 2007;21:93-108.
74. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, Koes B: **Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults.** *Cochrane Database Syst Rev* 2003 [http://ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=12804428&query_hl=1&itool=pubmed_docsum].
75. Karlsson N, Borg K, Carstensen J, Hensing G, Alexanderson K: **Risk of disability pension in relation to gender and age in a Swedish county; a 12-year population based, prospective cohort study.** *Work* 2006; 27:173-179.
76. Keeley J, Mayer T, Cox R, Gatchel R, Smith V, Mooney V: **Quantification of lumbar function. Part 5: Reliability of range-of-motion measures in the sagittal plane and an in vivo torso rotation measurements technique.** *Spine* 1986;11:31-35.
77. Kelly PJ, Lim LL: **Survival analysis for recurrent event data: an application to childhood infectious diseases.** *Statist Med* 2000;19:13-33.
78. **Key trends and outcomes.** In *Sickness, Disability and Work. Keeping on track in the economic downturn*. OECD, Directorate for employment, labour and social affairs, Background paper: High-Level Forum, Stockholm, 14-15 May 2009 [<http://www.oecd.org/dataoecd/42/15/42699911.pdf>].
79. Kjellström T (chairman). **Summary and conclusions** [In Swedish: Sammanfattning och slutsatser]. In: *Obesity – problems and interventions [In Swedish; Fetma – problem och åtgärder]*. The Swedish Council on Technology Assessment in Health Care. Report number: 160. Stockholm 2002:6-24. Summary in English: [<http://www.sbu.se/en/Published/Yellow/Obesity---problems-and-interventions/>].
80. Krause N, Dasinger LK, Deegan LJ, Rudolph L, Brand RJ: **Psychosocial job factors and return-to-work after compensated low back injury: a disability phase-specific analysis.** *Am J Ind Med* 2001;40:374-392.
81. Lakke SE, Soer R, Takken T, Reneman MF: **Risk and prognostic factors for non-specific musculoskeletal pain: a synthesis of evidence from systematic reviews classified into ICF dimensions.** *Pain* 2009;147:153-164.
82. Lancourt J, Kettelhut M: **Predicting return to work for lower back pain patients receiving worker's compensation.** *Spine* 1992;17:629-640.
83. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159-174.
84. Larsson G: **Rehabilitation back to work, a reform with the individual in the centre** [In Swedish; Rehabilitering till arbete, en reform med individen i centrum]. Swedish Government Official Reports. [In Swedish; Statens offentliga utredningar]. SOU 2000:78 [<http://www.regeringen.se/content/1/c4/28/31/a3196ee9.pdf>].
85. Larsson G: **The cost of the abuse is 150 billion.** Report from the forthcoming Swedish Government Official Reports of Abuse [In Swedish;

- Missbruksutredningen.] Newspaper article. [In Swedish; tidningsartikel.] Svenska Dagbladet, 2010-11-02, p. 5.
86. Leineweber C, Westerlund H, Hagberg J, Svedberg P, Luukkala M, Alexanderson K: **Sickness Presenteeism Among Swedish Police Officers.** *J Occup Rehabil* 2010; Jun 9 [Epub ahead of print]
 87. Leboeuf-Yde C: **Smoking and low back pain. A systematic literature review of 41 journal articles reporting 47 epidemiologic studies.** *Spine* 1999;24:1463-1470.
 88. Leboeuf-Yde C: **Body weight and low back pain. A systematic literature review of 56 journal articles reporting on 65 epidemiologic studies.** *Spine* 2000;25:226-237.
 89. Leboeuf-Yde C, Nielsen J, Kyvik KO, Fejer R, Hartvigsen J: **Pain in the lumbar, thoracic or cervical regions: do age and gender matter? A population-based study of 34,902 Danish twins 20-71 years of age.** *BMC Musculoskelet Disord* 2009;10:39.
 90. Lidwall U: **The Swedish Social Insurance Agency and vocational rehabilitation – active vocational rehabilitation measures and return to work.** [In Swedish; Försäkringskassan och arbetslivsinriktad rehabilitering – aktiva åtgärder och återgång i arbete.] Social Insurance Agency [In Swedish: Försäkringskassan.] Summary in English. 2006:10.
[http://www.forsakringskassan.se/irj/go/km/docs/fk_publishing/Dokument/Rapporter/fk_analyserar/analyserar_2006_10.pdf]
 91. **Life expectancy at birth.** In *The world factbook*. CIA, 2010
[<https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html>]
 92. Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson A: **The effect of graded activity on patients with sub-acute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach.** *Phys Ther* 1992;72:279-293.
 93. Lindström I, Öhlund C, Eek C, Wallin L, Petersen LE, Nachemson A: **Mobility, strength and fitness after a graded activity program for patients with subacute low back pain.** *Spine* 1992;7:641-652.
 94. Linton SJ: **Risk factors for neck and back pain in a working population in Sweden.** *Work & Stress* 1990;4:41-49.
 95. Linton SJ, Halldén K: **Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain.** *Clin J Pain* 1998;14:209-215.
 96. Linton SJ: **Psychological risk factors for neck and back pain.** In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott, Williams & Williams; 2000:57-78.
 97. Linton SJ: **Utility of cognitive-behavioral psychological treatments.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:361-381.
 98. Linton SJ, Ryberg M: **A cognitive-behavioral group intervention as prevention for persistent neck and back pain in a non-patient population: a randomized controlled trial.** *Pain* 2001;90:83-90.

99. Linton SJ, Boersma K, Jansson M, Svard L, Botvalde M: **The effects of cognitive-behavioral and physical therapy preventive interventions on pain-related sick leave: a randomized controlled trial.** *Clin J Pain* 2005;21:109-119.
100. Linton SJ, Nordin E: **A 5-year follow-up evaluation of the health and economic consequences of an early cognitive behavioral intervention for back pain: a randomized, controlled trial.** *Spine* 2006;31:853-858.
101. **Living conditions. Appendix 16. The Swedish Survey of Living Conditions. Design and methods.** Official Statistics of Sweden. Statistics Sweden [http://www.scb.se/statistik/LE/LE0101/_dokument/Appendix_16.PDF].
102. Ljungquist T, Harms-Ringdahl K, Nygren A, Jensen I: **Intra- and inter-rater reliability of an 11-test package for assessing dysfunction due to back or neck pain.** *Physiother Res Int* 1999;4:214-232.
103. Ljungquist T: **Physical performance tests and spinal pain. Assessing impairments and activity limitations.** *PhD thesis.* Division of Physiotherapy, Neurotec Department, and Section for Personal Injury Prevention, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; 2002.
104. Loisel P, Abenhaim L, Durand P, Esdaile JM, Suissa S, Gosselin L, Simard R, Turcotte J, Lemaire J: **A population-based, randomized clinical trial on back pain management.** *Spine* 1997;22:2911-2918
105. Loisel P, Lemaire J, Poitras S, Durand MJ, Champagne F, Stock S, Diallo B, Tremblay C: **Cost-benefit and cost-effectiveness analysis of a disability prevention model for back pain management: a six year follow up study.** *Occup Environ Med* 2002;59:807-815.
106. Macrae IF, Wright V: **Measurement of back movement.** *Ann Rheum Dis* 1969;28:584-589.
107. Magni G, Marchetti M, Moreschi C, Merskey H, Luchini SR: **Chronic musculoskeletal pain and depressive symptoms in the National Health and Nutrition Examination, I: epidemiological follow-up study.** *Pain* 1993;53:163-168.
108. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA: **Comprehensive review of epidemiology, scope, and impact of spinal pain.** *Pain Physician* 2009;2:E35-70.
109. Marklund S: **Risk- och friskfaktorer. Sjukskrivning och rehabilitering i Sverige** [In Swedish; English summary: Sickness absenteeism and vocational rehabilitation in Sweden]. RFV redovisar, 1997:6.
110. Marnetoft SU, Selander J, Bergroth A, Ekholm J: **The unemployed sick-listed and their vocational rehabilitation.** *Int J Rehabil Res* 1997;20:245-53.
111. Marnetoft SU, Selander J, Bergroth A, Ekholm J: **Factors associated with successful vocational rehabilitation in a Swedish rural area.** *J Rehabil Med* 2001;33:71-8.
112. Mastekaasa A: **Parenthood, gender and sickness absence.** *Soc Sci Med* 2000;50:1827-1842.
113. Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ: **Progressiv isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population.** *Spine* 1988;13:998-1002.
114. Mellin G: **Accuracy of measuring lateral flexion of the spine with a tape.** *Clin Biomech* 1986;1:85-89.

115. Melzack R, Wall PD. **Pain mechanism: a new theory.** *Science* 1965;150:971-979.
116. Moreland J, Finch E, Stratford P, Balsor B, Gill C: **Interrater reliability of six tests of trunk muscle function and endurance.** *J Manipulative Physiol Ther* 1997;26:200-208.
117. Nachemson A: **Introduction.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:1-12.
118. Nachemson A, Vingård E: **Influences of individual factors and smoking on neck and low back pain.** In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Philadelphia: Lippincott, Williams & Williams; 2000:79-95.
119. Nachemson A, Waddell G, Norlund A: **Epidemiology of neck and low back pain.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:165-187.
120. **Newly awarded disability pensions. Distribution between counties and diagnoses, 2006** [In Swedish; Nybeviljade sjukersättning/aktivitetsersättningar. Fördelning på län och diagnos, 2006] Social Insurance Agency. Statistics 2007:7 [In Swedish; Försäkringskassan. Statistik 2007:7] [http://statistik.forsakringskassan.se/rfvexcel/statpub/Statistik_2007-02.pdf].
121. Nitschke JE, Nattrass CL, Disler PB, Chou MJ, Ooi KT: **Reliability of the American Medical Association Guides' model for measuring spinal range of motion. Its implication for whole-person impairment rating.** *Spine* 1999;24:262-268.
122. Norberg M, Danielsson M: **Overweight, cardiovascular diseases and diabetes** [In Swedish; Övervikt, hjärt- och kärlsjukdomar och diabetes]. In: *The report of public health 2009* [In Swedish; Folhälsorapporten 2009]. The Swedish National Institute of Public Health. ISBN 978-91-978065-8-9. Stockholm 2009: 201-242 [http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/8495/2009-126-71_200912671.pdf].
123. Nordin M, Hiebert R, Pietrek M, Alexander M, Crane M, Lewis S: **Association of comorbidity and outcome in episodes of nonspecific low back pain in occupational populations.** *J Occup Environ Med* 2002; 44:677-684.
124. Norlund A, Waddell G: **Cost of back pain in some OECD Countries.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:421-425.
125. Oleinick A, Gluck JV, Guire K: **Factors affecting first return to work following a compensable occupational injury.** *Am J Ind Med* 1996;30:540-555.
126. Oleske DM, Lavender SA, Andersson GB, Morrissey MJ, Zold-Kilbourn P, Allen C, Taylor E: **Risk factors for recurrent episodes of work-related low back disorders in an industrial population.** *Spine* 2006;31:789-798.
127. Ostelo RW, de Ver HC, Waddell G, Kerckhoffs MR, Leffers P, van Tulder M: **Rehabilitation following first-time lumbar disc surgery: a systematic review within the framework of the Cochrane collaboration.** *Spine* 2003;28:209-218.
128. Ostelo RW, van Tulder MW, Vlaeyen JW, Linton SJ, Morley SJ, Assendelft WJ: **Behavioural treatment for chronic low-back pain.** *Cochrane Database Syst Rev* 2005

[<http://mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD002014/frame.html>].

129. **Pain terms: a list with definitions and notes on usage. Recommended by the IASP Subcommittee on Taxonomy.** *Pain* 1979;6:249-252.
130. Papageorgiou AC, Croft PR, Ferry S, Jayson MIV, Silman AJ: **Estimating the prevalence of low back pain in the general population. Evidence from the South Manchester back pain survey.** *Spine* 1995;20:1889-1894.
131. Papageorgiou AC, Croft PR, Thomas E, Ferry S, Jayson MI, Silman AJ: **Influence of previous pain experience on the episode incidence of low back pain: results from the South Manchester Back Pain Study.** *Pain* 1996;66:181-185.
132. Papageorgiou AC, Silman AJ, Macfarlane GJ: **Chronic widespread pain in the population: a seven year follow up study.** *Ann Rheum Dis* 2002;61:1071-1074.
133. Platzer W: *Color Atlas and Textbook of Human Anatomy. Volume 1: Locomotor system.* Stuttgart: Georg Thieme Verlag; 1978:141.
134. Price DD, Bushnell MC. **Overview of pain dimensions and their psychological modulation.** In: Price DD, Bushnell MC, editors. *Psychological methods of pain control: basic science and clinical perspectives. Progress in pain research and management.* Seattle: IASP Press; 2004:3-17.
135. **Public and other statistics.** [In Swedish; Officiell statistik och annan statistik.] Social Insurance Agency [In Swedish; Försäkringskassan]. [http://statistik.forsakringskassan.se/portal/page?_pageid=93,233342&_dad=portal&_schema=PORTAL].
136. Rheault W, Albright B, Byers C, Franta M, Johnson A, Skowronek M, Dougherty J: **Intertester reliability of the cervical range of motion device.** *J Orthop Sports Phys Ther* 1992;15:147-150.
137. Rissanen AM, Heliövaara M, Knekt P, Reunanen A, Aromaa A: **Determinants of weight gain and overweight in adult Finns.** *Eur J Clin Nutr* 1991;45:419-430.
138. Romelsjö A, Leifman H, Nyström S: **A comparative study of two methods for the measurement of alcohol consumption in the general population.** *Int J Epidemiol* 1995;24:929-936.
139. Rosner M: **Multisample inference.** In *Fundamental of biostatistics.* 5th edition. Duxbury: Thomson Learning; 2000: 511-576.
140. Rugulies R, Krause N: **Job strain, iso-strain, and the incidence of low back and neck injuries. A 7.5-year prospective study of San Francisco transit operators.** *Soc Sci Med* 2005;61:27-39.
141. Selander J, Marnetoft SU, Bergroth A, Ekholm J: **Vocational rehabilitation and future sick-leave.** *Disabil Rehabil* 1998;20:49-54.
142. Selander J, Marnetoft SU, Bergroth A, Ekholm J: **The process of vocational rehabilitation for employed and unemployed people on sick-leave: employed people vs unemployed people in Stockholm compared with circumstances in rural Jämtland, Sweden.** *Scand J Rehabil Med* 1998;30:55-60.
143. Selander J, Marnetoft SU, Bergroth A, Ekholm J: **Return to work following vocational rehabilitation for neck, back and shoulder problems: risk factors reviewed.** *Disabil Rehabil* 2002;24:704-712.
144. Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E: **The association between smoking and low back pain: a meta-analysis.** *Am J Med* 2010;123:8 7.e7-35.

145. **Sickness benefit.** Information sheet. Social Insurance Office [In Swedish; Sjukpenning. Faktablad. Försäkringskassan] [http://www.forsakringskassan.se/irj/go/km/docs/fk_publishing/Dokument/Publikationer/Faktablad/sj_4086_sjukpenning.pdf].
146. Spitzer WO, Skovron ML, Salmi LR, Cassidy JD, Duranceau J, Suissa S, Zeiss E: **Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining “whiplash” and its management.** *Spine* 1995;20(8 Suppl):1S-73S.
147. Staal J, Hlobil H, van Tulder MW, Köke AJ, Smid T, van Mechelen W: **Return-to-work interventions for low back pain: a descriptive review of contents and concepts of working mechanisms.** *Sports Med* 2002;32:251-267.
148. Staal JB, Hlobil H, Köke AJ, Twisk JW, Smid T, van Mechelen W: **Graded activity for workers with low back pain: who benefits most and how does it work?** *Arthritis Rheum* 2008;59:642-649.
149. StataCorp. 2005. Stata Statistical Software: Release 9. College Station, TX: StataCorp LP.
150. Steenstra IA, Anema JR, Bongers PM, de Vet HC, Knol DL, van Mechelen W: **The effectiveness of graded activity for low back pain in occupational healthcare.** *Occup Environ Med* 2006;63:718-725.
151. **Strategies for the limitation of sick-listing and disability pensioning. Experiences from Denmark, Finland, Iceland, Norway and Sweden 1990–2007.** Social Insurance Report [In Swedish; Strategier för begränsning av sjukfrånvaro och förtidspensionering. Erfarenheter från Danmark, Finland, Island, Norge och Sverige 1990–2007. Socialförsäkringsrapport.] 2008:5 [http://www.forsakringskassan.se/irj/go/km/docs/fk_publishing/Dokument/Rapporter/socialforsakringsrapporter/socialforsakringsrapport_2008_5.pdf]
152. Streder LE, Sjöblom A, Sundell K, Ludwig R, Taube A: **Interexaminer reliability in physical examination of patients with low back pain.** *Spine* 1997;22:814-820.
153. Sullivan M, Karlsson J, Ware JE Jr: **The Swedish SF-36 Health Survey. I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden.** *Soc Sci Med* 1995;41:1349-1358.
154. **Supplement to appendix 16.** Statistics Sweden [<http://www.scb.se/statistik/LE/LE0101/Supplement16.pdf>].
155. The buying of rehabilitation services of the Social Insurance Office. [In Swedish; **Försäkringskassans köp av tjänster för rehabilitering**] **The Swedish National Audit Office (In Swedish: Riksrevisionen) RiR 2004:21** [http://www.riksrevisionen.se/templib/pages/EArchiveItemPage____555.aspx?publicationid=1173&order=ZmFsc2U=]
156. The expenditures of the government budget divided in cost areas, 1995-2001. **The Swedish National Financial Management Authority 2001:Table appendix 3 [In Swedish; Statsbudgetens utgifter fördelade på utgiftsområden 1995-2001. Ekonomistyrningsverket 2001: Tabellbilaga 3]** [http://apps.esv.se/tapir/mm/extra/Bil_3.html]
157. Theorell T, Perski A, Åkerstedt T: **Changes in job strain in relation to changes in psychological state: a longitudinal study.** *Scand J Work Environ Health* 1988;14:189-196.

158. Theorell T, Karasek RA, Eneroth P: **Job strain variations in relation to plasma testosterone fluctuations in working men – a longitudinal study.** *J Intern Med* 1990;227:31-36.
159. Tornqvist EW, Kilbom A, Vingård E, Alfredsson L, Hagberg M, Theorell T, Waldenström M, Wiktorin C, Hogstedt C; MUSIC-Norrköping Study Group: **The influence on seeking care because of neck and shoulder disorders from work-related exposures.** *Epidemiology* 2001;12:537-545.
160. Trief P, Stein M: **Pending litigation and rehabilitation outcome of chronic low back pain.** *Arch Phys Med Rehabil* 1985;66:95-99.
161. Turk DC, Rudy TE: **Neglected topics in the treatment of chronic pain patients: Relapse, noncompliance, and adherence enhancement.** *Pain* 1991;44:5-28.
162. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, Simmons A, Williams G: **Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation.** *Ann Rheum Dis* 1998;57:649-655.
163. Valentine RE, Lewis JS: **Intraobserver reliability of 4 physiologic movements of the shoulder in subjects with and without symptom.** *Arch Phys Med Rehabil* 2006;87:1242-1249.
164. Van der Giezen AM, Bouter LM, Nijhuis FJ: **Prediction of return-to-work of low back pain patients sicklisted for 3-4 months.** *Pain* 2000;87:285-294.
165. Van der Weide WE, Verbeek JH, Sallé HJ, van Dijk FJ: **Prognostic factors for chronic disability from acute low-back pain in occupational health care.** *Scand J Work Environ Health* 1999;25:50-56.
166. Van Tulder MW, Assendelft WJ, Willem JJ, Koes BW, Bouter LM: **Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders.** *Spine* 1997;22:2323-2330.
167. Vingård E, Nachemson A: **Work-related influences on neck and low back pain.** In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Philadelphia: Lippincott, Williams & Williams; 2000:97-126.
168. Von Korff M, Dworkin SF, LeResche L, Kruger A: **An epidemiologic comparison of pain complaints.** *Pain* 1988;32:173-183.
169. Waddell G, Main CJ: **Assessment of severity in low back disorders.** *Spine* 1984;9:204-208.
170. Waddell G: **Low back pain: a twentieth century health care enigma.** *Spine* 1996;21:2820-2825.
171. Waddell G, Waddell H: **A review of social influences on neck and back pain and disability.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment.* Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:13-55.
172. Waddell G, Burton K: **Occupational health guidelines for the management of low back pain at work: evidence review.** *Occup Med* 2001;51:124-135.
173. Waddell G: **Diagnostic triage.** In *Back Pain Revolution.* 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:9-26.
174. Waddell G: **Pain and disability.** In *Back Pain Revolution.* 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:27-45.
175. Waddell G: **The epidemiology of back pain.** In *Back Pain Revolution.* 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:71-89.

176. Waddell G: **The clinical course of back pain.** In *Back Pain Revolution*. 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:115-136.
177. Waddell G, van Tulder M: **Clinical guidelines.** In *Back Pain Revolution*. 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:283-322.
178. Waddell G, Watson PJ: **Rehabilitation.** In *Back Pain Revolution*. 2nd edition. Edited by Waddell G. Edinburgh: Churchill Livingstone; 2004:371-399.
179. Webb R, Brammah T, Lunt M, Urwin M, Allison T, Symmons D: **Prevalence and predictors of intense, chronic, and disabling neck and back pain in the UK general population.** *Spine* 2003;28:1195-1202.
180. Whiplashkommissionen.se [homepage on internet]. **What are whiplash-related injuries?** In: The whiplash commission final report 2005:35-49. [http://www.whiplashkommissionen.se/pdf/WK_finalreport.pdf]
181. Willis BH: **Spectrum bias – why clinicians need to be cautious when applying diagnostic test studies.** *Family Practice* 2008;25:390-396.
182. Åstrand NE, Isacsson SO: **Back pain, back abnormalities, and competing medical, psychological, and social factors as predictors of sick leave, early retirement, unemployment, labour turnover and mortality: a 22 year follow up of male employees in a Swedish pulp and paper company.** *Br J Ind Med* 1988;45:387-395.
183. Öhlund C, Eek C, Palmblad S, Areskoug B, Nachemson A: **Quantified pain drawing in subacute low back pain. Validation in a nonselected outpatient industrial sample.** *Spine* 1996;21:1021-30.

I

Living conditions, including life style, in primary-care patients with non-acute, non-specific spinal pain compared with a population-based sample: a cross-sectional study

Odd Lindell^{1*}, Sven-Erik Johansson¹, Lars-Erik Strender¹

¹Center for Family and Community Medicine, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Alfred Nobels allé 12, SE-141 83 Huddinge, Sweden; Phone +46-8-524 800 00, fax +46-8-524 887 07.

* Corresponding author; e-mail address: odd.lindell@ki.se

Abstract

Background: Non-specific spinal pain (NSP), comprising back and/or neck pain, is one of the leading disorders behind long-term sick-listing, including disability pensions. Early interventions to prevent long-term sick-listing require the identification of patients at risk. The aim of this study was to compare living conditions associated with long-term sick-listing for NSP in patients with non-acute NSP, with a non-patient population-based sample. Non-acute NSP is pain that leads to full-time sick-listing >3 weeks.

Methods: One hundred and twenty-five patients with non-acute NSP, 2000–2004, were included in a randomised controlled trial in Stockholm County with the object of comparing cognitive-behavioural rehabilitation with traditional primary care. For these patients, a cross-sectional study was carried out with baseline data. Living conditions were compared between the patients and 338 non-patients by logistic regression. The conditions from univariate analyses were included in a multivariate analysis. The non-significant variables were excluded sequentially to yield a model comprising only the significant factors (p -value <0.05). The results are shown as odds ratios (OR) with 95% confidence intervals.

Results: In the univariate analyses, 13 of the 18 living conditions had higher odds for the patients with a dominance of physical work strains and *Indication of alcohol over-consumption*, OR 14.8 [3.2–67.6]. Five conditions qualified for the multivariate model: *High physical workload*, OR 13.7 [5.9–32.2]; *Hectic work tempo*, OR 8.4 [2.5–28.3]; *Blue-collar job*, OR 4.5 [1.8–11.4]; *Obesity*, OR 3.5 [1.2–10.2]; and *Low education*, OR 2.7 [1.1–6.8].

Conclusions: As most of the living conditions have previously been insufficiently studied, our findings might contribute a wider knowledge of risk factors for long-term sick-listing for NSP. As the cross-sectional design makes causal conclusions impossible, our study should be complemented by prospective research.

Keywords: non-specific spinal pain, back pain, neck pain, long-term sick-listing, population-based sample, cross-sectional study.

Introduction

Since the late 1990s, the industrial world, particularly Sweden, has seen a substantial growth of sick-listing, especially on a long-term basis, including disability pensions. In 2007, despite a slight decrease since 2004, 11% of Swedes of working age was sick-listed versus 6% in comparable countries.¹ Up to and including 2004, musculoskeletal disorders, dominated by spinal pain, comprising back and/or neck pain, formed the largest diagnostic group behind disability pensions in Sweden. Following international trends, it was outflanked from 2005 by mental disorders.² Nevertheless, despite this relative decrease, recent data indicate a continued increase in the total cost to society of spinal pain.³ The vast majority of cases concern non-specific spinal pain (NSP) and present a task for primary care.⁴

Clinical guidelines emphasize the necessity of early intervention to prevent long-term sick-listing caused by NSP,⁴ requiring the identification of patients at risk. Socio-economic and medical factors are associated both with the onset of acute NSP and the progression to non-acute NSP.^{5–8} Acute and non-acute NSP is pain that leads to full-time sick-listing for ≤3 weeks and >3 weeks respectively.⁹ However, research within the area has been seriously limited with, e.g., an under-representation of women.¹⁰

Sweden has a unique tradition of keeping population statistics, going back as far as 1749.¹¹ Since 1975 extensive annual surveys of living conditions, including life style, have been conducted on large random samples representing Sweden as a whole as well as local districts.¹² This provides an

exceptional opportunity for epidemiological research. However, we have found no previous study where primary-care patients with non-acute NSP were compared with a population-based sample.

The aim of this study was to compare living conditions associated with long-term sick-listing for NSP in patients with non-acute NSP, with a non-patient population-based sample.

Methods

The study was approved by the local ethics committee at Karolinska University Hospital, Huddinge, Sweden.

Setting and source population

The study area was the Southern part of Stockholm County, including 5 urban districts (Enskede-Årsta-Vantör, Farsta, Älvsjö, Skarpnäck and Hägersten-Liljeholmen) and 4 semi-urban districts (Huddinge, Nynäshamn, Tyresö and Haninge). The number of inhabitants (31 December 2001) in the county totalled about 1,830,000, of whom 1,100,100 were of the same age as the patients studied (18–59 years). The study area had about 467,000 inhabitants, of whom 281,000 were aged 18–59 years and constituted the source population. A detailed description of the distribution of the inhabitants between the districts is shown in Table 1.

Table 1. Distribution of inhabitants and patients between the study districts. Ranking by the number of patients.

Districts (inhabitants; total 467,298 ^a)	Inhabitants aged 18–59 years (%)		Patients (%)	
	Frequency	Cumulative frequency	Frequency	Cumulative frequency
Huddinge (85,700)	50,430 (18.0)	50,430 (18.0)	37 (29.6)	37 (29.6)
Nynäshamn (24,332)	13,523 (4.8)	63,953 (22.8)	36 (28.8)	73 (58.4)
Tyresö (39,434)	22,454 (8.0)	86,407 (30.8)	26 (20.8)	99 (79.2)
Enskede-Årsta-Vantör (80,984)	49,562 (17.7) ^b	135,969 (48.5)	11 (8.8)	110 (88.0)
Haninge (70,432)	42,487 (15.1)	178,456 (63.6)	5 (4.0)	115 (92.0)
Farsta (45,597)	26,211 (9.3) ^b	204,667 (72.9)	3 (2.4)	118 (94.4)
Älvsjö (20,786)	11,861 (4.2) ^b	216,528 (77.2)	3 (2.4)	121 (96.8)
Skarpnäck (40,060)	24,979 (8.9) ^b	241,507 (86.1)	3 (2.4)	124 (99.2)
Hägersten-Liljeholmen (59,973)	39,118 (13.9) ^b	280,625 (100.0)	1 (0.8)	125 (100.0)

^a 31 December 2001; ^b Concerns age group 20–64 years (data for 18–59 years were not available)

Patients

One hundred and twenty-five patients with non-acute NSP, between August 2000 and January 2004, were included in a randomized controlled trial, which was described in detail in a previous study.¹³ The patients were allocated to a multi-disciplinary, cognitive-behavioural programme at a rehabilitation center or continued with traditional primary care. The rehabilitation center opened in 1990 and was situated in Haninge, geographically near the middle of the study area.

The criteria for inclusion: 1. Vocationally active, up to and including 59 years of age. 2. Sick-listed full-time for spinal pain for at least 6 weeks (42 days) and for at most 2 years (730 days). 3. Able to fill in forms. *The criteria for exclusion:* 1. Temporary disability pension, or disability pension being paid or in preparation. 2. A primary need for action by a hospital specialist (e.g., operation for intra-vertebral slipped disc). 3. Pregnancy and diseases (other than spinal pain) that would probably make rehabilitation impracticable (e.g., advanced pulmonary disease). 4. Whiplash associated disorders as a

primary obstacle to working. 5. Previous rehabilitation at the rehabilitation center. 6. Other multi-disciplinary rehabilitation ongoing or planned.

The patients living in the study area were recruited by 41 family doctors at 13 primary-care health centers. Twelve of the centers engaged >1 family doctor, and 1 center was a 1-doctor clinic. To ensure that all the study patients, including those who were allocated to continued primary care, received a high minimum level of treatment, only permanently employed or long-term substitute doctors were engaged. The rehabilitation center was well known to the family doctors, as they had been referring patients to it for several years. The recruitment of the patients was non-systematic, i.e., dependent on the motivation and available time of the family doctor. Before randomization, the study patients met a research assistant in the health center and completed a questionnaire of baseline characteristics. A detailed description of the distribution of included patients between the family doctors is shown in Table 2, and of the distribution of the patients between the districts in Table 1. One of the 125 patients failed to complete the questionnaire and was excluded. The remaining 124 patients were included in this study.

Table 2. Distribution of patients (n=125) between the recruiting family doctors (n=41). Ranking by the number of patients.

Family doctors (%)		Patients (%)	
Frequency	Cumulative frequency	Frequency	Cumulative frequency
1 (2.4)	1 (2.4)	17 (13.6)	17 (13.6)
1 (2.4)	2 (4.9)	16 (12.8)	33 (26.4)
1 (2.4)	3 (7.3)	10 (8.0)	43 (34.4)
1 (2.4)	4 (9.8)	8 (6.4)	51 (40.8)
1 (2.4)	5 (12.2)	7 (5.6)	58 (46.4)
1 (2.4)	6 (14.6)	5 (4.0)	63 (50.4)
4 (9.8)	10 (24.4)	4 (3.2)	79 (63.2)
5 (12.2)	15 (36.6)	3 (2.4)	94 (75.2)
5 (12.2)	20 (48.8)	2 (1.6)	104 (83.2)
21 (51.2)	41 (100.0)	1 (0.8)	125 (100.0)

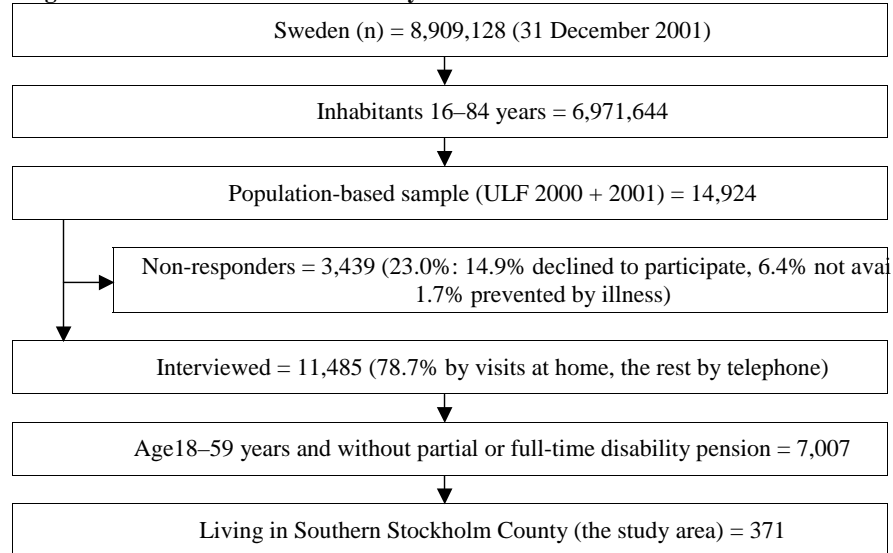
Non-patients

From a nationwide sample, a simple, random, local sample of 338 non-patients was selected as a comparison group to the patients:

Statistics Sweden, a governmental authority, conducts The Survey of Living Conditions annually (In Swedish: Undersökningarna av levnadsförhållanden (ULF)).^{12,14} To reach an acceptable power, two years of ULF data, 2000 + 2001, were combined. Most of the patients (81/124) were recruited during that period. A flowchart of ULF is shown in Figure 1.

ULF 2000+2001 was a simple, random sample of 7,465 and 7,459 individuals respectively, aged 16–84 years. They were invited to participate in an interview in their homes. Non-responders and those who declared that they did not want to be visited were offered a telephone interview. From the interviewed individuals we selected subjects of the same age as the patients except for those with partial or total disability pensions. This resulted in a nationwide sample, of which 371 individuals were living in the home districts of the patients. By exclusion of the vocationally inactive and the full-time sick-listed subjects, a comparison group of 338 non-patients was achieved.

Figure 1. Flowchart of the ULF surveys 2000 + 2001.



Living conditions associated with long-term sick-listing for NSP

The cross-sectional design made conclusions about causes and effects impossible. For example, anxiety, depression and low physical activity could be both explanatory and responding variables for non-acute NSP.^{6,7} We therefore limited our analyses to living conditions that could reasonably be supposed to have existed before the start of the current sick-listing and excluded comparisons of, e.g., mental distress, pain and exercise habits.

For a majority (10 out of 18) of the living conditions, the questions in the patient questionnaire and the ULF questionnaire were identical or nearly identical. As regards 8 living conditions, we made modifications so they were reasonably comparable. The non-identical questions in the study and ULF, and our modifications of them, are shown in Table 3.

Questions concerning alcohol consumption were put only to the ULF subjects of 2001, of whom 169 belonged to the non-patients. Questions regarding work conditions were put exclusively to the 325 non-patients in employment. The questions concerning the other living conditions were put to all non-patients.

The 18 living conditions associated with long-term sick-listing for NSP are shown in Table 4. The rationale of the choice of conditions is shown as references in the table.

Outcome measure

As the outcome variable of logistic regression, being either a patient or a non-patient.

Statistics

The patients were compared with the non-patients by applying logistic regression. Stata, version 10.1 was used to analyze the data.¹⁵

We first estimated the distribution of the living conditions for the patients and the non-patients. The results are shown as proportions (means) with 95% confidence intervals. Differences between the patients and the non-patients were evaluated by univariate-logistic regression, adjusted for gender and age.¹⁶ Two age classes were defined: *Old age* ≥ 45 years and *Young age* ≤ 44 years. The outcome (dependent) variable was the sample class, i.e., patient or non-patient. The predictive (independent) variable was the living condition. The results are presented with odds ratios (OR), 95% confidence intervals and *p*-values.

Table 3. Non-identical questions in the questionnaires of the randomized controlled trial and the ULF surveys.

Living condition	The wording of the questions is shown in italics.	
	The randomized controlled trial	ULF ^a
<i>High physical workload</i>	<i>State work conditions that you regularly (not occasionally) have been or are exposed to. For each alternative, Yes/No:</i> <i>Lifting heavy things or greater muscular efforts?</i>	<i>Does your work require lifting heavy things? Yes/No. (If Yes:) Are lifting heavy things required: Daily – Some time every week – More seldom (Question 124). Specification: We considered “daily” as equivalent to “regularly”.</i>
<i>Monotonous work moments</i>	<i>Monotonous work movements?</i>	For each alternative, Yes/No: <i>Does your work include very frequent and monotonous movements?</i>
<i>Difficult work postures</i>	<i>Difficult work postures (bent, twisted, locked, etc.)?</i>	<i>In your work, are you forced to be bent, twisted or in other ways to adopt unsuitable working postures?</i>
<i>Vibrations in work</i>	<i>Vibrations? ⁴⁷</i>	<i>Are you exposed to powerful shakings or vibrations in your work? (Question 123)</i>
<i>Hectic work tempo</i>	For each alternative, Yes, often – Yes, sometimes – No, seldom – No, practically never, 1–4: <i>Does your work require that you work very fast?</i> Specification: Hectic work tempo ≤ 3.	<i>Is your work hectic? Yes/No.</i> For each alternative, No possibilities – Very many possibilities, 0–10: <i>What possibilities do you think you have to... ...decide how your daily work should be performed? ...influence decisions of the general direction of your work? (Question 128 b+d) Specification: Low decision latitude = <5 in both questions.</i>
<i>Low decision latitude</i>	<i>Do you have the freedom to decide... ... how your work should be performed? ... what to be done in your work? ²²</i> Specification: Low decision latitude ≥ 2 in both questions.	
<i>Indication of alcohol over-consumption</i>	<i>How often do you on one and the same occasion drink half a bottle of strong spirits (bottle = 75 cl) or 2 bottles of wine or 6 tins of strong beer (= 8 bottles of 33 cl) or 12 bottles of medium-strong beer? Almost every day (at least 5 days weekly) – 3-4 times weekly – 1-2 times weekly – 2-3 times monthly – Once monthly – 1-6 times yearly – Never, 1–7 ³⁵</i> Specification: 1–4 = increased tolerance, which indicates alcohol over-consumption. This cut-off point, i.e., a frequency of binge drinking of at least 2-3 times monthly, is based on a personal communication (Anders Romelsjö, 27 August 2007).	<i>Roughly, how often during the last 12 months have you drunk any alcoholic drinks, i.e. wine, strong beer or strong spirits? Daily or almost daily (at least 5 days weekly) – 2-4 times weekly – Once weekly – 2-3 times monthly – Once monthly – 6-11 times yearly – More seldom – Never, 1–8 (Question 64 e).</i> <i>Roughly, how many glasses do you usually drink at those occasions? One glass could be 1 glass of wine, 1 bottle or tin of strong beer, 1 snaps or drink: number of glasses (Question 64 f).</i> Specification: 1–4 in question 64 e + >8 glasses in question 64 f indicate increased tolerance.
<i>Comorbidity</i>	<i>Except your back/neck/shoulder pain – do you have any other, current diseases? Yes/No. (If Yes:) What disease/s?:</i>	<i>Do you have any prolonged disease, trouble after an accident, any handicap or other frailty? Yes/No. (If Yes:) Note every trouble or disease as precisely as possible:</i> (Question 42–43)

^a The complete ULF questionnaire: http://www.scb.se/statistik/LE/LE0101/_dokument/ULF_2001.pdf; ^b Snaps is Swedish for a little glass (often 4–6 centiliters) of pure liquor, e.g., vodka.

Table 4. Living conditions. Univariate analyses. One hundred and twenty-four patients with non-acute NSP compared with 338 non-patients by logistic regression, adjusted for gender and age. If not otherwise stated, results are shown as number (in case of missing data, the total number is also shown) with percentage in parenthesis; 95% confidence intervals within brackets.

	Patients (n=124)	Non-patients (n=338)	Odds ratio	p-val
<i>Woman</i> [52,49,75,32]	68 (54.8 [46.0–63.7])	161 (47.6 [42.3–53.0])	1.3 [0.9–2.0]	NS
<i>Older age</i> (= ≥45 years) [80,28]	57 (46.0 [37.1–54.9])	107 (31.7 [26.7–3.6])	1.8 [1.2–2.8]	0.006
<i>Immigrant</i> (= born outside Sweden) [22]	34 (27.4 [19.5–35.4])	43 (12.7 [9.2–16.3])	2.6 [1.6–4.4]	<0.001
<i>Single life</i> (= living alone without children) [112]	22 (17.7 [10.9–24.6])	101 (29.9 [25.0–34.8])	0.5 [0.3–0.9]	0.02
<i>Living with children at home</i> [112]	69 (55.7 [46.8–64.5])	167 (49.4 [44.1–54.8])	1.3 [0.9–2.0]	NS
<i>Low education</i> (= at most junior high school) [182]	44 (35.5 [26.9–44.0])	41 (12.1 [8.6–15.6])	3.8 [2.3–6.3]	<0.001
<i>Unemployed</i> [143]	29 (23.4 [15.8–30.9])	13 (3.9 [1.8–5.9])	8.2 [4.0–16.5]	<0.001
<i>Blue-collar job</i> ^{a,b} [171,94]	83 (87.4 [80.6–94.2])	108 (33.2 [28.1–38.4])	15.0 [7.7–29.1]	<0.001
Physical work strains ^a : [94]				
<i>High physical workload</i> [52]	79 (83.2 [75.5–90.8])	51/325 (15.7 [11.7–19.7])	30.4 [15.9–58.3]	<0.001
<i>Monoton. work moments</i> [52]	61 (64.2 [54.4–74.0])	134/324 (41.4 [36.0–46.7])	2.7 [1.7–4.3]	<0.001
<i>Difficult work postures</i> [52]	76 (80.0 [71.8–88.2])	107/324 (33.0 [27.9–38.2])	9.0 [5.1–15.9]	<0.001
<i>Vibrations in work</i> [159]	35 (36.8 [27.0–46.7])	15/324 (4.6 [2.3–6.9])	18.6 [8.7–39.9]	<0.001
Psychosocial work strains ^a : [157]				
<i>Hectic work tempo</i> [62]	88 (92.6 [87.3–98.0])	239/324 (73.8 [68.9–78.6])	4.5 [2.0–10.1]	<0.001
<i>Low decision latitude</i> [52]	30 (31.6 [22.1–41.1])	42/321 (13.1 [9.4–16.8])	3.2 [1.8–5.5]	<0.001
<i>Smoking</i> (daily + not daily) [80]	49 (39.5 [30.8–48.2])	118/336 (35.1 [30.0–40.2])	1.2 [0.8–1.8]	NS
<i>Indication of alcohol over-consumption</i> ^c [30]	17 (13.7 [7.6–19.8])	2/164 (1.2 [-0.0–2.9])	14.8 [3.2–67.6]	0.001
<i>Obesity</i> (= BMI ≥ 30 [79]) [58]	30 (24.2 [16.6–31.8])	23/332 (6.9 [4.2–9.7])	4.3 [2.3–7.7]	<0.001
<i>Comorbidity</i> ^d [123]	45 (36.3 [27.7–44.9])	105 (31.1 [26.1–36.0])	1.1 [0.7–1.7]	NS

^a Concerning the subjects in employment: 95/124 patients and 325/338 non-patients; ^b According to Socio-Economic Classification (In Swedish “Socioekonomisk indelning (SEI)”) [http://www.scb.se/statistik/LE/LE0101/_dokument/SEIstandard.pdf]. **Modification:** the subjects in the group “Entrepreneur” were considered *Blue-collar job* starting from their probable level of education; ^c The alcohol questions were put to 169/338 non-patients; ^d = Any other prolonged disease except NSP and obesity.

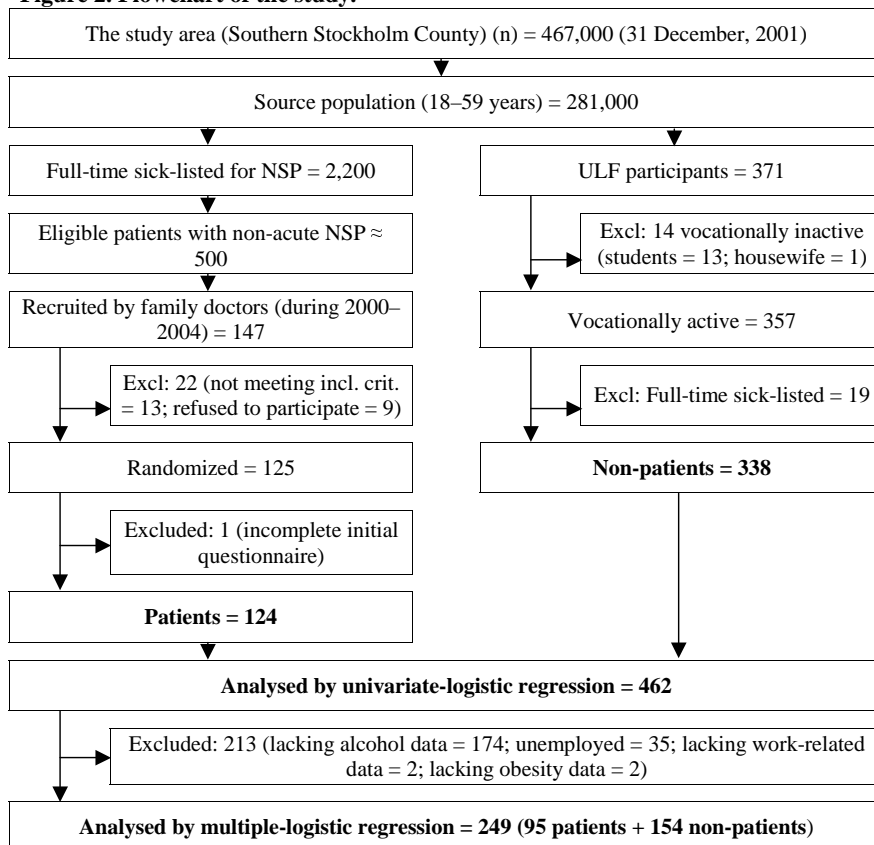
Several of the living conditions could be expected to intercorrelate, e.g., *Immigrant* and *Low education*, and *Blue-collar job* and *High physical workload*. To find the most discriminative living conditions we used multiple-logistic regression, adjusted for gender and age, with the sample class as the outcome variable and the living conditions as the explanatory variables. A prerequisite for multiple-logistic regression is the same number of respondents for the different variables,¹⁶ so subjects with missing data were excluded from the multivariate analysis (Figure 2). This left 249 subjects (95 patients and 154 non-patients) for multiple-logistic regression analysis. We first explored univariate analyses. The variables with a *p*-value of at most 0.10 are presented with OR, *p*-values and 95% confidence intervals. They were included in a multiple model, from which the variables with *p*-values

of 0.05 or higher were excluded stepwise to yield a model comprising only variables with p -values <0.05 . The final multivariate model is presented with OR, p -values, 95% confidence intervals, a goodness-of-fit test by Hosmer-Lemeshow, the percentage of correctly predicted patients, and the area under the ROC-curve.¹⁶

Results

A flowchart of the study is shown in Figure 2.

Figure 2. Flowchart of the study.



Eligible subjects in the source population

From ULF data, we estimated a point prevalence of individuals with full-time sick-listing for NSP to 0.8% or 2,200 subjects. As these data included both short and long-term sick-listing, we had to estimate the proportion of non-acute NSP, i.e., the individuals with sick-listing >3 weeks. Previous research indicates an initial high recovery speed: starting from full-time sick-listing for NSP, ~90% of the individuals have returned to work after 12 weeks, and the rate clearly levels off thereafter.⁹ We estimated the point prevalence in the source population of non-acute NSP to be ~0.2% or ~500 individuals. We have no data for the prevalence over time.

Patients

A majority of the patients were recruited by a minority of the doctors: 15 doctors (36.6%) recruited in all 94 patients (75.2%). Twenty-one doctors recruited only 1 patient each (Table 2). Ninety-nine

patients (79.2%) were living in 3 of the 9 districts, this number of inhabitants corresponding to 30.8% of the total number of inhabitants in the study area (Table 1).

The mean age of the 124 patients was 42.6 (range 18–59) years. The proportion of *Old age* was significantly higher than among the non-patients (Table 4). Females predominated slightly. The current sick-listing period at baseline was *m* 170.9 (range 43–721) days.

Non-patients

The mean age of the 338 non-patients was 39.3 (range 19–59) years. Males predominated slightly. However, the difference in gender distribution versus the patients was non-significant (Table 4).

Outcome

In the univariate analyses, 13 of the 18 conditions had higher odds for the patients with a dominance of physical and psychosocial work strains, and *Indication of alcohol over-consumption* (OR 14.8); only 1 condition, *Single life* (OR 0.5), had lower odds (Table 4).

Five conditions qualified for the final multivariate model: *High physical workload* (OR 13.7), *Hectic work tempo* (OR 8.4), *Blue-collar job* (OR 4.5), *Obesity* (OR 3.5) and *Low education* (OR 2.7) (Table 5). The proportion of correctly classified subjects was high (85.5%) and the area under ROC-curve was large (0.92; the maximum would be 1.0).

Discussion

Living conditions associated with long-term sick-listing in primary-care patients with non-acute NSP were compared with a local sample of non-patients. In the univariate analyses, the patients had higher odds for 13 of the 18 conditions. In the multivariate analysis, 5 conditions qualified, indicating work strains, lower social class and life-style.

Work strains

High physical workload and *Hectic work tempo* were the two outstanding living conditions in the model. The association of *High physical workload* with NSP has been pinpointed in many studies.^{18–21} Job strain, i.e., high demands, including among other items a high work tempo, and low control,²² has been associated with disabling NSP in several studies.^{23–26} *Hectic work tempo* as a single risk factor, however, is far less clear. In a review of risk factors for NSP, insufficient evidence was found for high work pace.²⁷ Despite occasional studies that indicate a relationship between high work tempo and a longer time to return-to-work,²⁸ a recent review of psychosocial predictors of failure to return to work in NSP showed strong evidence for the recovery expectations of the patients, while stress/psychological strain were non-predictive.²⁹ This is also in line with our newly-published prospective study, where *High self prediction* qualified as a predictor of stable return-to-work, while work-related variables did not.¹⁷

Indicators of lower social class

Blue collar job and *Low education* are closely associated and might be looked upon as different aspects of belonging to a lower social class.⁵ *Low education* limits the chances of getting a white-collar job, which explains the great dominance of work strains in the model and the fairly low degree of variance for *Blue collar job* and *Low education* in themselves. There is conflicting evidence in previous research of a relationship between NSP and lower social class. A possible association is probably a matter of social disadvantage, although it is not clear which aspects of the disadvantage are important.⁵ In our prediction study, there were indications that *Low education*, though a non-predictor, may have qualified as a predictor with a longer follow-up than the 2 years of that study.¹⁷ *Blue-collar job*, however, was a clear non-predictor. With a prevalence of 87.4% of the patients versus 33.2% of the non-patients, it is logical that such a great difference qualifies for a multivariate analysis with the sample class as outcome variable. The prediction study, however, exclusively involved patients with return-to-work/non-return-to-work as the outcome. A variable of such overwhelming frequency might be non-discriminative, although it has a powerful effect on sick-listing. Concerning the possible association between sick-listing for NSP and social class, and according to a large review from 2004,

there is a lack of conclusive studies.¹⁰ Our research might contribute to the elucidation of this complex issue.

Life style

While the prevalence of *Smoking* was non-significantly higher in the patients, the prevalences of *Indication of alcohol over-consumption* and *Obesity* were remarkably higher.

Smoking as a non-predictor of disabling NSP was indicated in a cross-national, prospective study from 2000, including about 2,000 subjects.²⁰ However, associations between smoking and NSP have been found in several other studies. A review of 1999 indicated smoking as a weak risk indicator but not a cause of NSP and signs of causality were evident only in the study with the largest sample, >30,000 subjects.³⁰ In a review from 2000, a possible association between NSP and cigarette smoking was suggested, but the lack of prospective studies was emphasized.³¹ In a recent meta-analysis of both cross-sectional and prospective studies, current as well as former smoking was associated with NSP, though the association was fairly modest.³² The non-significance of *Smoking* in this study and in our prediction study might therefore be due to the small sample size.

More or less hidden alcohol abuse constitutes one of the greatest public health problems, with substantial social and clinical implications. Large population studies have shown that 10% to 15% of all men and approximately 5% of all women suffer from chronic alcohol dependency³³ and quite recent primary-care research indicates a continued increase of those proportions.³⁴ Among the subjects with chronic alcohol dependency about one quarter are in a phase of active abuse.³³ This should correspond to around 2–2.5% of the non-patients of our study, which was approximately confirmed by the ULF data. The patients had a substantially higher prevalence, and we have found nothing equivalent in any other study of NSP. The reason could be our use of one single binge-drinking question (Table 3),³⁵ which might decrease the risk for under-estimation of alcohol abuse in questionnaires that ask for total intake. We have found no previous study of NSP where this question has been used. However, in the multivariate context, the alcohol issue was eliminated by other closely-correlated variables. For example, 15 of the 16 subjects in the multivariate analysis with *Indication of alcohol over-consumption* had a *Blue-collar job*. One study showed that alcohol over-consumption was not a risk factor for long-term sick-listing for NSP,³⁶ but this was contradicted by an other study.³⁷ In our prediction study, *Indication of alcohol over-consumption* did not predict sick-listing during a 2-year follow-up.¹⁷ Though these conflicting results motivate further research, this cross-sectional study might contribute in pinpointing the comparatively higher prevalence of abuse problems among those patients.

During recent decades the prevalence of obesity has increased remarkably but with a certain international variation. For example, while the prevalence in the USA has increased to a full 20%,³⁸ it doubled in Sweden from the years 1980/81 to 2004/5 from 5% to 10% in both women and men.³⁹ Comorbidities with obesity include diabetes, cardiovascular diseases, pain in general and NSP in particular.⁴⁰ In our study, the prevalence of *Obesity* in the non-patients during the years 2000/01 corresponded well with the 7% in 1996/97 concerning all Swedes 16-84 years,³⁹ while among the patients it was more than 3-fold higher. *Obesity* remained significant in the multivariate model, though with a decreased OR, probably influenced by *Low education*, which is a risk factor for obesity.⁴¹ According to a review from 2000, obesity should be considered a possible weak risk indicator, but with insufficient data to assess whether it causes spinal pain.⁴² In a prospective study from 2002, obesity was a risk factor for the transition from acute to non-acute NSP, though with low OR (1.7).²¹ However, in a large review from 2004 concerning predictors for non-acute NSP, there was insufficient evidence for obesity as a risk factor.¹⁰ A quite recent, very large, cross-sectional population-based study from Norway indicated associations between obesity and NSP and commented that further studies were needed to determine whether the association was causal.⁴³ Our prediction study, however, found no impact of *Obesity* on sick-listing.¹⁷ *Obesity* was found in 24.2% of the patients versus 6.9% of the non-patients. In line with the paragraph above, such a difference might qualify for a model with the sample class as outcome variable, but be eliminated in an analysis with return-to-work/non-return-to-work as outcome. It therefore remains unclear whether, how and why obesity and NSP are

correlated.⁴⁰ Furthermore, the clinical relevance of that association, if any, is obscure. Recently, however, a reduction of musculoskeletal pain was reported in a study of a weight-reduction program, at least on a short-term basis, which might be of future clinical interest in the treatment of disabling NSP.⁴⁴

To sum up: the patients were distinguished by higher odds of obesity, higher odds of indication of alcohol abuse that vanished in the multivariate analysis, and non-significant differences concerning smoking. Prospective research, including our prediction study, has yielded conflicting results. Therefore, the causal associations between smoking, alcohol abuse and obesity and sick-listing for NSP, if any, are small.

Strengths of the study

One of the strengths of our study was the good representation of women.

As in the ULF surveys, the patient questionnaires were completed under the supervision of an assistant during an interview with the patient in the recruiting health center. This might have contributed to the high quality of the patient data, and increased the comparability between the patients and the non-patients.

The design of the nationwide ULF also allows local comparisons to be made. The responding rates of the ULF in 2000 and 2001 were practically 80%. These high-quality data concerning the comparison group were a strength.

Another strength of the study was the excellent model fit. The number of variables in the multivariate model was by a wide margin within the upper limit, which is suggested in previous research.⁴⁵

Limitations of the study

The sample of 124 patients was a very low proportion of the eligible subjects and the inclusion procedure was prolonged and non-systematic. These limitations are discussed in detail in our predictor study.¹⁷ A closely-related limitation was the geographical imbalance in the recruitment; however, the greatest number of recruited patients were living in the district with the greatest number of inhabitants (Huddinge) (Table 1).

A limitation was the non-prospective design. However, this study might contribute to a more detailed cross-sectional picture of the patients with non-acute NSP, which is also of value in the planning of prospective research, e.g., our predictor study.¹⁷

External validity

To what extent might the results be generalized beyond the samples of patients and non-patients studied and be applied to other subjects (population validity) or settings (ecological validity)? As the rehabilitation center and the family doctors engaged were very well established, the 124 patients might be reasonably representative of the everyday primary care in the study area, comprising a comparably large part of Stockholm County. The 338 non-patients in the study were generally comparable with non-patients in the nation. The only significant ($p < 0.05$) differences from the national sample of 7,007 subjects were a higher prevalence of *Immigrant* (12.7 vs. 10.5%), *Unemployed* (3.9 vs. 8.4%), *Blue-collar job* (33.3 vs. 39.2%), *Heavy physical workload* (15.7 vs. 35.3%) and *Vibrations in work* (4.6 vs. 8.6%). According to a large cross-national study, including primary care in 14 countries in 5 continents, the dominating pain problem was non-acute spinal pain; and despite certain variations, the cross-national manifestations of spinal pain were surprisingly equivalent.⁴⁶ Therefore, given that the study samples are reasonably representative of Swedish primary care, the external validity might also be satisfactory from a non-Swedish perspective.

Clinical implications

Standing alone, the cross-sectional design of this study limits its clinical implications. However, together with prospective studies, it might increase the knowledge of what distinguishes patients with non-acute, non-specific spinal pain. Though this knowledge in no way includes unambiguous

management options, it might help family doctors, supervisors in the work place, handling officers of the Social Insurance Agency, etc, to identify subjects at risk.

Conclusions

The living conditions associated with long-term sick-listing of 124 patients with non-acute non-specific spinal pain were compared with 338 non-patients by applying logistic regression. In the univariate analyses, 13 of the 18 conditions had higher odds for the patients with a dominance of physical work strains and *Indication of alcohol over-consumption* (OR 14.8). Five conditions qualified for the multivariate analysis: *High physical workload* (OR 13.7), *Hectic work tempo* (OR 8.4), *Blue-collar job* (OR 4.5), *Obesity* (OR 3.5), and *Low education* (OR 2.7). As most of those living conditions have hitherto been insufficiently studied, our findings might help extend our knowledge of what distinguishes the individuals at risk for long-term sick-listing due to NSP. As the cross-sectional design makes causal conclusions impossible, our study should be complemented by prospective research.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

OL was the main investigator and carried out the study, performed the analysis and drafted the manuscript. SEJ contributed to the statistical analysis. LES, as supervisor of OL, participated in all phases of the study. All authors read and approved the final manuscript.

Acknowledgement

This study was supported by grants from the Stockholm County Social Insurance Agency, Stockholm County Council, Ministry of Health and Social Affairs, Vårdal Foundation, Cardionics and Pharmacia (now part of Pfizer).

References

1. Oecd.org [homepage on the Internet]. Key trends and outcomes. In: *Sickness, Disability and Work. Keeping on track in the economic downturn*. OECD, Directorate for employment, labour and social affairs. Background paper: High-Level Forum, Stockholm, 14-15 May 2009. p. 11-16. Available from: <http://www.oecd.org/dataoecd/42/15/42699911.pdf> Accessed October 8, 2010.
2. Forsakringskassan.se [homepage on the Internet]. Newly awarded disability pensions. Distribution between counties and diagnoses, 2006 [In Swedish; Nybeviljade sjukersättningar/aktivitetsersättningar. Fördelning på län och diagnos, 2006] Social Insurance in Sweden. Statistics 2007:7 [In Swedish; Försäkringskassan. Statistik 2007:7]. Available from: http://statistik.forsakringskassan.se/rfvexcel/statpub/Statistik_2007-02.pdf Accessed October 8, 2010.
3. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician*. 2009 July;12(4): [about 36 p.]. <http://www.painphysicianjournal.com/2009/july/2009;12;E35-E70.pdf> Accessed October 8, 2010.
4. Waddell G, van Tulder M. Clinical guidelines. In: Waddell G, editor. *Back Pain Revolution*. 2nd ed. Edinburgh: Churchill Livingstone; 2004:283-322.
5. Waddell G, Waddell H. A review of social influences on neck and back pain and disability. In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott, Williams & Williams; 2000:13-55.
6. Linton SJ. Psychological risk factors for neck and back pain. In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott, Williams & Williams; 2000:57-78.
7. Nachemson A, Vingård E. Influences of individual factors and smoking on neck and low back pain. In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott, Williams & Williams; 2000:79-95.

8. Vingård E, Nachemson A. Work-related influences on neck and low back pain. In: Nachemson A, Jonson E, editors. *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott, Williams & Williams; 2000:97-126.
9. Waddell G. The biopsychosocial model. In: Waddell G, editor. *Back Pain Revolution*. 2nd ed. Edinburgh: Churchill Livingstone; 2004:265-282.
10. Hansson T, Jensen I. Sickness absence due to back and neck disorders. In: Alexanderson K, Norlund A, editors. *Sickness absence – causes, consequences, and physicians' sickness-certification practice*. *Scand J Public Health*. 2004;63 Suppl:S109-S151.
11. Un.org [homepage on the Internet]. Brief history and other relevant background information. Country Profile of Sweden. United Nations Statistical Division. [<http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=666#start>] Accessed October 8, 2010.
12. Scb.se [homepage on the Internet]. Living conditions. Appendix 16. The Swedish Survey of Living Conditions. Design and methods. Official Statistics of Sweden. Statistics Sweden. [http://www.scb.se/statistik/LE/LE0101/_dokument/Appendix_16.PDF] Accessed October 8, 2010.
13. Lindell O, Johansson SE, Strender LE. Subacute and chronic, non-specific back and neck pain: cognitive-behavioural rehabilitation versus primary care. A randomized controlled trial. *BMC Musculoskeletal Disord*. 2008 Dec;9(172):[about 17 p.]. [<http://www.ncbi.nlm.nih.gov/pubmed/19116007>] Accessed October 8, 2010.
14. Scb.se [homepage on the Internet]. Supplement to appendix 16. Statistics Sweden. [<http://www.scb.se/statistik/LE/LE0101/Supplement16.pdf>] Accessed October 8, 2010.
15. StataCorp. 2005. Stata Statistical Software: Release 9. College Station, TX: StataCorp LP.
16. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons; 2000.
17. Lindell O, Johansson SE, Strender LE. Predictors of stable return-to-work in non-acute, non-specific spinal pain: low total prior sick-listing, high self prediction and young age. A two-year prospective cohort study. *BMC Fam Pract*. 2010 July;11(53):[about 12 p.]. [<http://www.ncbi.nlm.nih.gov/pubmed/20646286>] Accessed October 8, 2010.
18. Krause N, Dasinger LK, Deegan LJ, Rudolph L, Brand RJ. Psychosocial job factors and return-to-work after compensated low back injury: a disability phase-specific analysis. *Am J Ind Med*. 2001;40(4):374-392.
19. Dasinger LK, Krause N, Deegan LJ, Brand RJ, Rudolph L. Physical workplace factors and return to work after compensated low back injury: a disability phase-specific analysis. *J Occup Environ Med*. 2000;42(3):323-333.
20. Hansson TH, Hansson EK. The effects of common medical interventions on pain, back function, and work resumption in patients with chronic low back pain: A prospective 2-year cohort study in six countries. *Spine*. 2000;25(23):3055-3064.
21. Fransen M, Woodward M, Norton R, Coggan C, Dawe M, Sheridan N. Risk factors associated with the transition from acute to chronic occupational back pain. *Spine*. 2002;27(1):92-98.
22. Theorell T, Karasek RA, Eneroth P. Job strain variations in relation to plasma testosterone fluctuations in working men - a longitudinal study. *J Intern Med*. 1990;227(1):31-36.
23. Rugulies R, Krause N. Job strain, iso-strain, and the incidence of low back and neck injuries. A 7.5-year prospective study of San Francisco transit operators. *Soc Sci Med*. 2005;61(1):27-39.
24. IJzelenberg W, Burdorf A. Risk factors for musculoskeletal symptoms and ensuing health care use and sick leave. *Spine*. 2005;30(13):1550-1556.
25. Ghaffari M, Alipour A, Farshad AA, Jensen I, Josephson M, Vingard E. Effect of psychosocial factors on low back pain in industrial workers. *Occup Med (Lond)*. 2008;58(5):341-347.
26. Harcombe H, McBride D, Derrett S, Gray A. Physical and psychosocial risk factors for musculoskeletal disorders in New Zealand nurses, postal workers and office workers. *Inj Prev*. 2010;16(2):96-100.
27. Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM. Systematic review of psychosocial factors at work and private life as risk factors for back pain. *Spine*. 2000;25(16):2114-2125.
28. van der Weide WE, Verbeek JH, Sallé HJ, van Dijk FJ. Prognostic factors for chronic disability from acute low-back pain in occupational health care. *Scand J Work Environ Health*. 1999;25(1):50-56.
29. Iles RA, Davidson M, Taylor NF. Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review. *Occup Environ Med*. 2008;65(8):507-517.
30. Leboeuf-Yde C. Smoking and low back pain. A systematic literature review of 41 journal articles reporting 47 epidemiologic studies. *Spine*. 1999;24(14):1463-1470.
31. Goldberg MS, Scott SC, Mayo NE. A review of the association between cigarette smoking and the development of nonspecific back pain and related outcomes. *Spine*. 2000; 25(8):995-1014.

32. Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E. The association between smoking and low back pain: a meta-analysis. *Am J Med*. 2010 July;123(1):[about 29 p]. [http://www.sciencedirect.com/science?_ob=MIimg&_imagekey=B6TDC-4Y0T775-N-1&_cdi=5195&_user=3285060&_pii=S000293430900713X&_origin=search&_coverDate=01%2F31%2F2010&_sk=998769998&view=c&wchp=dGLzVtb-zSkzk&md5=7582f0241d1b0042e2b3e6972ac94462&ie=/sdarticle.pdf] Accessed October 8, 2010.
33. Berglund M (chairman). Summary and conclusions [In Swedish; Sammanfattning och slutsatser]. In: *Treatment of alcohol and drug abuse - an evidence-based review [In Swedish; Behandling av alkohol- och narkotikaproblem - en evidensbaserad kunskapssammanställning]*. The Swedish Council on Technology Assessment in Health Care. ISBN: 91-87890-73-9. Stockholm 2001:11-28. Summary in English: [http://www.sbu.se/upload/Publikationer/Content0/1/alcoholdrug_2001/drugslut.html] Accessed 8 October, 2010.
34. Zivin K, Pfeiffer PN, Szymanski BR, et al. Initiation of Primary Care-Mental Health Integration programs in the VA Health System: associations with psychiatric diagnoses in primary care. *Med Care*. 2010;48(9):843-851.
35. Romelsjö A, Leifman H, Nyström S. A comparative study of two methods for the measurement of alcohol consumption in the general population. *Int J Epidemiol*. 1995;24(5):929-936.
36. Holmberg SA, Thelin AG. Primary care consultation, hospital admission, sick leave and disability pension owing to neck and low back pain: a 12-year prospective cohort study in a rural population. *BMC Musculoskelet Disord*. 2006 Aug;7(66):[about 8 p.]. [<http://www.biomedcentral.com/content/pdf/1471-2474-7-66.pdf>] Accessed October 8, 2010.
37. Demyttenaere K, Bruffaerts R, Lee S, et al. Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys. *Pain*. 2007;129(3):332-342.
38. Kjellström T (chairman). Summary and conclusions [In Swedish: Sammanfattning och slutsatser]. In: *Obesity - problems and interventions [In Swedish; Fetma – problem och åtgärder]*. The Swedish Council on Technology Assessment in Health Care. Report number: 160. Stockholm 2002:6-24. Summary in English: [<http://www.sbu.se/en/Published/Yellow/Obesity---problems-and-interventions/>] Accessed October 8, 2010.
39. Norberg M, Danielsson M. Overweight, cardiovascular diseases and diabetes [In Swedish; Övervikt, hjärt- och kärlsjukdomar och diabetes]. In: *The report of public health 2009 [In Swedish; Folhälsorapporten 2009]*. The Swedish National Institute of Public Health. ISBN 978-91-978065-8-9. Stockholm 2009:201-242 [http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/8495/2009-126-71_200912671.pdf] Accessed October 8, 2010.
40. Janke EA, Collins A, Kozak AT. Overview of the relationship between pain and obesity: What do we know? Where do we go next? *JRRD*. 2007;44(2):245-262.
41. Rissanen AM, Heliövaara M, Knekt P, Reunanen A, Aromaa A. Determinants of weight gain and overweight in adult Finns. *Eur J Clin Nutr*. 1991;45(9):419-430.
42. Leboeuf-Yde C. Body weight and low back pain. A systematic literature review of 56 journal articles reporting on 65 epidemiologic studies. *Spine*. 2000;25(2):226-237.
43. Heuch I, Hagen K, Heuch I, Nygaard Ø, Zwart JA. The impact of body mass index on the prevalence of low back pain: the HUNT study. *Spine*. 2010;35(7):764-768.
44. Kotowski SE, Davis KG. Influence of weight loss on musculoskeletal pain: Potential short-term relevance. *Work*. 2010;36(3):295-304.
45. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373-1379.
46. Gureje O, Simon GE, von Korff M. A cross-national study of the course of persistent pain in primary care. *Pain*. 2001;92(1-2):195-200.
47. Linton SJ. Risk factors for neck and back pain in a working population in Sweden. *Work & Stress*. 1990;4(1):41-49.
48. Hagen KB, Holte HH, Tambs K, Bjerkedal T. Socioeconomic factors and disability retirement from back pain: a 1983-1993 population-based prospective study in Norway. *Spine*. 2000;25(19):2480-2487.
49. Karlsson N, Borg K, Carstensen J, Hensing G, Alexanderson K. Risk of disability pension in relation to gender and age in a Swedish county; a 12-year population based, prospective cohort study. *Work*. 2006;27(2):173-179.
50. Du Bois M, Szpalski M, Donceel P. Patients at risk for long-term sick leave because of low back pain. *Spine*. 2009;9(5):350-359.

51. Borg K, Hensing G, Alexandersson K. Predictive factors for disability pension - an 11-year follow up of young persons on sick leave due to neck, shoulder or back diagnoses. *Scand J Public Health*. 2001;29(2):104-112.
52. Mastekaasa A. Parenthood, gender and sickness absence. *Soc Sci Med*. 2000;50(12):1827-1842.
53. Åstrand NE, Isacson SO. Back pain, back abnormalities, and competing medical, psychological, and social factors as predictors of sick leave, early retirement, unemployment, labour turnover and mortality: a 22 year follow up of male employees in a Swedish pulp and paper company. *Br J Ind Med*. 1988;45(6):387-395.
54. Selander J, Marnetoft SU, Bergroth A, Ekholm J. Return to work following vocational rehabilitation for neck, back and shoulder problems: risk factors reviewed. *Disabil Rehabil*. 2002;24(14):704-712.
55. Tornqvist EW, Kilbom A, Vingård E, et al. The influence on seeking care because of neck and shoulder disorders from work-related exposures. *Epidemiology*. 2001;12(5):537-545.
56. Theorell T, Perski A, Åkerstedt T. Changes in job strain in relation to changes in psychological state: a longitudinal study. *Scand J Work Environ Health*. 1988;14(3):189-196.
57. Nordin M, Hiebert R, Pietrek M, Alexander M, Crane M, Lewis S. Association of comorbidity and outcome in episodes of nonspecific low back pain in occupational populations. *J Occup Environ Med*. 2002;44(7):677-684.

II

Research article

Open Access

The reliability of a 10-test package for patients with prolonged back and neck pain: could an examiner without formal medical education be used without loss of quality? A methodological study

Odd Lindell^{*1,2}, Lars Eriksson³ and Lars-Erik Strender¹

Address: ¹Center for Family Medicine, Karolinska Institutet, Alfred Nobels allé 12, SE-141 83 Huddinge, Sweden, ²Rehab Station Stockholm, Frösundaviks allé 13, SE-169 89 Solna, Sweden and ³Haninge FysioCenter, Rörvägen 4, SE-136 50 Haninge, Sweden

Email: Odd Lindell^{*} - odd.lindell@ki.se; Lars Eriksson - lars@hfc.se; Lars-Erik Strender - lars-erik.strender@ki.se

^{*} Corresponding author

Published: 3 April 2007

Received: 19 May 2006

BMC Musculoskeletal Disorders 2007, 8:31 doi:10.1186/1471-2474-8-31

Accepted: 3 April 2007

This article is available from: <http://www.biomedcentral.com/1471-2474/8/31>

© 2007 Lindell et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: In the rehabilitation of patients with prolonged back and neck pain, the physical impairment should be assessed. Previous research has exclusively engaged medically educated examiners, mostly physiotherapists. However, less biased evaluations of efforts at rehabilitation might be achieved by personnel standing outside the treatment work itself. Therefore, if medically untrained examiners could be used without cost to the quality, this might produce a better evaluation at defensible cost and could also be useful in a research context. The aim of this study was to answer the question: given a 10-test package for patients with prolonged back and neck pain, could an examiner without formal medical education be used without loss of quality? Five of the ten tests required the examiner to keep a firm hold against the foundation of those parts of the participant's body that were not supposed to move during the test.

Methods: Examination by an experienced physiotherapist (A) in performing the package was compared with that by a research assistant (B) without formal medical education. The reliability, including inter- and intra-rater reliability, was assessed. In the inter-rater reliability study, 50 participants (30 patients + 20 healthy subjects) were tested once each by A and B. In the intra-rater reliability study, the 20 healthy subjects were tested twice by A or B. One-way ANOVA intra-class-correlation coefficient (ICC) was calculated and its possible systematic error was determined using a t-test.

Results: All five tests that required no manual fixation had acceptable reliability (ICC > .60 and no indication of systematic error). Only one of the five tests that required fixation had acceptable reliability. The difference (five vs. one) was significant ($p = .01$).

Conclusion: In a 10-test package for patients with prolonged back and neck pain, an examiner without formal medical education could be used without loss of quality, at least for the five tests requiring no manual fixation. To make our results more generalizable and their implications more searching, a similar study should be conducted with two or more examiners with and without formal medical education, and the intra-rater reliability study should also include patients and involve more participants.

Background

In the industrial world, back and neck pain, i.e. pain in the lumbar, thoracic and/or cervical spine, constitutes the largest diagnostic group underlying sick-listing, including disability pensions [1]. In the rehabilitation of patients with prolonged back and neck pain, it is necessary to assess the physical impairment, i.e. the pathological, anatomical or physiological abnormality of structure or function leading to loss of normal ability [2]. The vast majority (around 95%) of these patients suffer from non-specific back and neck pain and require no specific surgical, rheumatological or neurological treatment. Therefore, the focus of assessment of prolonged back and neck pain is on abnormality of function [3]. Acceptable reliability of an assessment method includes acceptable inter- and intra-rater reliability, i.e. it requires that the measurements are comparable when performed (a) on the same subject by numerous examiners and (b) on several occasions by the same examiner [4]. Besides reliability, validity, i.e. the degree to which a useful interpretation can be inferred from a measurement [5], is an important aspect of an assessment method. For example, if in a lifting test the patient is able to lift 10 kg as maximum, how is the clinical meaning for that individual interpreted? However, the concept of validity is outside the framework of this study.

Forward bending, when it is measured as the distance of the fingertip to the floor and by the modified Schober test, had acceptable reliability [6], as did lateral bending measured as the distance moved by the hand down the outside of the thigh [7]. Trunk rotation and active-straight-leg raise have been examined by goniometers, but those tests were not validated [8]. Cervical bending and rotation as investigated by the CROM instrument demonstrated acceptable reliability [9,10]. Isometric endurance of the abdominal muscles as examined in the form of a partial sit-up had acceptable reliability [11]. Moreau et al. [12] found that the Biering-Sørensen test was the most useful of the isometric back-extension endurance tests. In an 11-test package, six of the tests had acceptable reliability [4]; in an 8-test package, only one test had acceptable reliability [13].

Patients with prolonged back and neck pain were offered rehabilitation at a Swedish primary-care centre. The physiotherapists at the centre used a 10-test package. Most of the tests in this package had been validated in previous studies by comparing the results obtained by medically trained examiners. From August 2000 to January 2006 a randomized controlled trial was running, in which rehabilitation at the centre was compared with traditional primary care. At the time of inclusion and one year later, each patient in the randomized controlled trial met a research assistant at that patient's health centre. Among other items, the patients performed the 10-test package. For

practical and economic reasons it was appropriate for the person who administrated the study and visited the different health centres also to execute the test package. Although the research assistant had no formal medical education, this seemed reasonable, since the tests were standardized and easy to perform. In some reliability studies, chiropractors [14], naprapaths [15] or physicians [6,15-18] have been represented. The vast majority of reliability studies, however, have been performed with physiotherapists as examiners [4,9-11,13,19-21]. We have found no study of reliability in which examiners without formal medical education were engaged. However, the evaluation of rehabilitation efforts might be less biased if performed by personnel standing outside the treatment work itself. It seems economically unrealistic for ordinary clinics to keep medically-trained personnel only for assessment tasks. Therefore, if medically untrained examiners could be used without decreased quality, this might produce a better assessment of outcome at defensible cost and could also be useful in a research context.

The aim of this study was to answer the question: given a 10-test package for patients with prolonged back and neck pain, could an examiner without formal medical education be used without loss of quality?

Methods

Settings

The study was performed in Haninge, a rural district 25 km south-east of Stockholm, at a primary-care rehabilitation centre and a physiotherapy centre situated next door.

Examiners

In appraising the assessment work of a medically untrained examiner it seemed logical to use an experienced physiotherapist as the gold standard.

Examiner A (LE) had the highest Swedish degree in orthopaedic manual therapy and had been working as a physiotherapist for ten years. Examiner B had a B.A. (Bachelor of Arts) in psychology but no formal medical education. She had been working as a research assistant with purely administrative tasks for 2 1/2 years and had no previous vocational experience of manual contact with patients. B was prepared for this reliability study by (a) four hours' training in the performance of the 10-test package and (b) practising the package during the autumn of 2000 on barely 40 patients who were included in the above-mentioned randomized controlled trial.

Subjects

Fifty participants were included and gave their consent to participate in the study: 30 patients with prolonged back and/or neck pain, and 20 healthy subjects.

Patients

From March until September 2001, a total of 30 patients were recruited at the physiotherapy centre. Seventeen were females (mean (*m*) 41.5, range (*r*) 28–60, years) and 13 males (*m* 42.4, *r* 20–63, years). They were supplied with both verbal and written information.

Inclusion criteria

1. Back and/or neck pain for more than four weeks. 2. The patient was considered able to execute the whole 10-test package.

Exclusion criteria

1. Such severe pain or dysfunction that it might be harmful for the patient to participate. 2. Whiplash-associated disorders. 3. Inability to read the written information.

Thirty-one consecutive patients fulfilling the criteria were asked to participate in the study. All but one agreed.

Healthy subjects

From February until September 2001, 20 healthy subjects were recruited among the staff at the rehabilitation centre and the physiotherapy centre. Fourteen were females (*m* 36.2, *r* 22–55, years) and six males (*m* 40.2, *r* 28–53, years). Twenty staff members (physiotherapists, physicians and receptionists) were asked consecutively and all of them agreed to participate.

The 10-test package

Four tests included motion in one direction only. Four comprised motion to the right and to the left, and one involved motion forward and backward. A lifting test included a lumbar and a cervical sub-test. This resulted in ten tests composed of 16 sub-tests.

Five of the ten tests required that the examiner kept a firm hold against the foundation of those parts of the participant's body that were not supposed to move during the test. This manual fixation was done to eliminate misleading co-movements from those parts.

The package followed the protocol of previous studies, with some modifications. We used the widely-adopted modification of the Schober test by Macrae and Wright [22]. To save examination time, we simplified the procedures for another two original tests, the Biering-Sørensen test and the PILE test (see below). The total examination time of the package was approximately 30 minutes. A detailed description is given below.

1. Forward bending

The participant (P) stood barefoot with the heels together. P bent forward, keeping the knees straight and with the arms straightened out downwards the floor. When P had

bent maximally, the examiner (E) measured the distance between the middle-finger tip and the floor, to within 1 cm, with a wooden stick. If the floor was reached, the distance was noted as 0 cm [6].

2. Modified Schober

P stood with the feet together. Three dots were marked: dot a between the lowest lumbar spinal process and sacrum, dot b 10 cm above and dot c 5 cm beneath a. P bent forward, keeping the knees straight. The distance b-c when P was bent maximally forward was measured with a tape to within 1 cm. The difference of b-c when maximally bent forward and standing was noted. Normally, b-c increases by at least 5 cm [22].

3. Lateral bending (right/left)

P stood with 20 cm between the feet and with the back, neck, back of the head and shoulders against a wall and the arms loosely against the sides of the body. The middle-finger tip positions on the outside of the thighs were marked with dot a. P bent to the right side, keeping the knees straight and without losing contact between the shoulders and the wall. In the maximally bent position, the middle-finger tip position on the right thigh was marked by dot b. The same procedure was performed on the left side. The distances a-b on the right and left thighs were measured with a tape to within 1 cm [7].

4. Trunk rotation (right/left)

P sat on a stool with the knees together holding a rod horizontally in the frontal plane across the upper sternum and the front of the deltoid muscles. From the ends of the rod, a line with a plumb weight hung down pointing at a semicircular protractor lying on the floor under and in front of P. In the initial position, the base line of the protractor was in the same frontal plane as the rod and the middle of the base line was directly below the middle of the rod. E stood behind P holding the lower part of P's body still by firmly pressing the iliac crests down towards the seat of the stool. P rotated the trunk maximally to the right. The maximally rotated position was read, to within 5 degrees, where the plumb weight pointed at the protractor. The same procedure was performed on the left side [8].

5. Active-straight-leg raise (right/left)

P was lying supine on a couch with the knees straight. An MIE meter was placed on the lower part of the right leg at the tuberositas tibiae. While the left leg was held in its initial position by E, P raised the right leg, keeping the knee straight. When the leg was maximally raised, the angle between the leg and the horizontal plane was read to within 1 degree. The same procedure was performed with the right leg fixed to the couch and the left leg raised [8].

6. Cervical bending (forward/backward)

P sat on a chair with the head in a neutral position. A CROM meter was placed on the head. E held P's thoracic and lumbar spine fixed to the back support of the chair. P bent the head forward and then backward. In the maximally bent positions, the angle between the head and the vertical line was read to within 1 degree [9].

7. Cervical rotation (right/left)

The same procedure as in test 6, except that P rotated the head to the right and then to the left. The angle between the head in neutral and in maximally rotated position was read to within 1 degree [9].

8. Abdominal endurance

P was lying supine on a couch with the knees bent at 90°, the soles of the feet on the couch and the palms resting on the front of the thighs. P performed a sit-up, with the fingertips touching the upper part of the patellae, and sustained this position as long as possible. The maximal sit-up time, until the fingers lost contact with the patellae, was measured with a stop-watch to within 1 second [11].

9. Modified Biering-Sørensen

P was lying prone with the lower part of the body, from the upper part of the iliac crest downwards, placed on a couch. The upper part of the body hung down from the short side of the couch, resting on the seat of a chair 2 dm beneath the level of the couch. E held P's feet fixed to the couch. P lifted the upper body from the seat and held it straight out from the edge of the couch, with the arms folded across the chest. The maximal time for which P was able to keep the unsupported upper body horizontal was measured with a stop-watch to within 1 second.

Modifications

In the original Biering-Sørensen, the buttocks and legs are fixed by three canvas straps and there is an upper time limit of 240 seconds [6].

10. Modified PILE (lumbar/cervical)

PILE = Progressive Iso-inertial Lifting Evaluation.

Modified PILE lumbar

P lifted a tray with weights (plastic bottles filled with sand) from the floor to a 75-cm-high table and back again to the floor. The table was placed 90° to the left of P, which added a twisting factor. An electronic pulse-counter was attached to P's thorax. The starting weight was 4 kg. E added 2 kg after each successful attempt. Each attempt had to be carried out within 20 seconds. The weight managed during the last lifting moment was recorded as the test result. The test was discontinued if the heart rate reached 85% of the estimated maximal heart rate or if the load reached 55% of the body weight.

Modified PILE cervical

This sub-test was carried out as described above, except that P stood in front of the table and lifted the tray from the table up to a 50-cm-high platform (i.e. 125 cm above the floor). The platform was placed on the left side of the table, which added a twisting factor.

Modifications

In the original PILE, the table is 76 cm high, the platform is 137 cm above the floor, men and women have different weights at the start (3.6 vs. 5.9 kg) and different weights are added to men and women (2.25 vs. 4.5 kg), and the result is adjusted for the body weight [16]. Our modifications are in line with Lindström et al. ([8]; Lindström, personal communication, 2000).

Examination procedure

The test package was performed at different times of day. Along with the agreement to participate, the participants received identical instructions, both verbally and in written form, from a manual produced for this study. They were to wear training clothes or underclothes, not to do any warming up, and to perform the tests to their maximum capacity within the limits of exertion and pain; they could discontinue whenever they wanted. The participants were also informed that the examiners were a physiotherapist and a research assistant. The patients were not informed about which of the two examiners they were seeing. The healthy subjects could not be blinded to the examiner because they were co-workers of one or both of the examiners. Whether A or B would conduct the first examination was randomized by envelopes, which were prepared by an independent statistician and opened immediately before the first test. Close to the start of the examination the participant was once again verbally instructed to perform the tests to his or her maximum capacity within the limits of exertion and pain, and was reminded that the tests could be discontinued whenever he or she wanted. The test package was then conducted straight through without a break and without further verbal communication, except for purely technical instructions on how to perform the test. Before the first and after the last test of the package, the participants were asked to estimate their exertion on Borg's 20-point scale [23] and their level of pain on Borg's 10-point scale [24].

The participants and the examiners were given no results on any occasion until all the tests were completed. The participants were asked not to tell the second examiner their experiences at the first examination.

The study was approved by the local ethics committee at Karolinska University Hospital, Huddinge, Sweden.

Inter-rater reliability study

The 30 patients and 20 healthy subjects were first tested by one of the examiners (examination 1). After a break for 30 minutes, they were re-tested by the other examiner (examination 2).

Intra-rater reliability study

The 20 healthy subjects participated. Examiners A and B tested ten healthy subjects each. After examination 2, the subjects rested for another 30 minutes and were then re-tested (examination 3) by the same examiner as at examination 1.

The reason for including only healthy subjects in the intra-rater reliability study was that we considered three consecutive examinations too much of a strain for the patients to be ethically defensible; it would also have made the results of the third examination difficult to interpret.

In total, the patients and the healthy subjects were occupied in the study for approximately 1 1/2 and 2 1/2 hours respectively.

Statistics

Although the intra-class correlation coefficient (ICC) is questioned by some authors [25], it is the basic measure in most reliability studies involving continuous data (degrees, centimetres, etc.) [10,13,17,20,26,27]. The ICC increases with the degree of reliability up to a maximum of 1.00 for identical ratings [28]. We calculated the one-way ANOVA (analysis of variance) ICC, random-effects model, and its 95% confidence interval (CI) as described by Haas [28]. We also calculated the standard error of measurement (SEM) of the ICC [29]. The 95% CI is a band of values that, with 95% confidence, contains the true reliability. A narrow CI suggests a more precise estimate of reliability. The SEM enables the reliability of a measurement expressed in the units of the measurement of interest, such as degrees or centimetres, to be assessed. As such, it is valuable for the clinician because it provides guidance on whether the measured change is due to measurement error or to real change [27].

There is a lack of consensus concerning the cut-off values for ICC. For example, Rheault et al. [10] considered ICC > .80 to indicate high reliability and ICC > .60 up to and including .80 to represent moderate reliability. Horneij et al. [13] defined an ICC > .75 as excellent reliability and .40-.75 as fair to good reliability. We chose to consider an ICC > .60 to indicate acceptable reliability and an ICC ≤ .60 to indicate poor reliability, which is a modification of Landis and Koch [30] and in line with the recommendation of Chinn [31].

For each sub-test, the mean difference between the measurements and its 95% CI were calculated. The possible systematic error of the ICC was calculated, using a *t*-test to evaluate the mean difference [17]. We considered a sub-test to have acceptable inter- or intra-rater reliability when ICC was > .60 and there was no significant, systematic error. A test was considered to have acceptable reliability when it had (1) acceptable inter-rater reliability for the 50 participants, (2) acceptable intra-rater reliability for both examiners A and B and (3), for tests comprising two sub-tests, when both sub-tests had acceptable inter- and intra-rater reliability. The proportions of tests that showed acceptable inter-rater reliability were calculated for the patients and for the healthy subjects, and for the five tests that required manual fixation and the five that did not. The proportions of tests with acceptable intra-rater reliability were calculated for A and B and for the tests that did and did not require manual fixation. The proportions of tests with acceptable reliability were calculated for the tests that did and did not require manual fixation. The mutual proportions were then compared by a *z*-test [32].

For each sub-test, scatter plots were used to visualize the agreement. The plots were constructed from the difference between the measurements and the mean difference, and the limits of agreement were indicated by the 95% CI of the mean difference [33].

The exertion and pain before and after each examination were analysed. The difference between examinations 1 and 2 of the 50 participants was compared by the Wilcoxon sign-rank test. The differences between examinations 1 and 3 and the differences between the healthy subjects of examiners A and B were compared by the Wilcoxon rank-sum test [34].

A *p*-value < .05 was considered statistically significant. The statistical calculations were performed and the figures constructed using STATA, version 9.1.

Results

All 50 participants completed all the tests.

Inter-rater reliability

Seven of the ten tests had acceptable inter-rater reliability (Table 1). Three tests had poor inter-rater reliability: active-straight-leg raise, cervical bending and modified Biering-Sørensen.

For the patients and the healthy subjects, seven and four of the ten tests respectively had acceptable inter-rater reliability (not significant (NS)).

All five tests that required no manual fixation by the examiner had acceptable inter-rater reliability, compared with two of the

Table 1: Inter-rater reliability. Fifty participants tested by A (the physiotherapist) and B (the research assistant). The five tests that required manual fixation are italicized. ICC in bold text indicates acceptable ICC (> .60). The mean difference between the measurements by A and B is compared, p-value in bold text indicates a significant difference (p < .05). + indicates acceptable, - indicates poor inter-rater reliability.

10-test package (including 16 sub-tests):	Forward bending (cm)	Modified Schober (cm)	Lateral bending (cm)		Trunk rotation (°)		Active-straight-leg raise (°)		Cervical bending (°)		Cervical rotation (°)		Abdom. endurance (seconds)	Mod. Biering- Sorensen (sec.)	Modified PILE (kg)
			Right	Left	Right	Left	Right	Left	Forward	Backward	Right	Left			
All of the 50 participants															
ICC	.99	.79	.93	.95	.82	.85	.94	.90	.61	.84	.70	.69	.92	.91	.97
95% CI of ICC	.98–1.00	.67–.88	.89–.96	.91–.97	.70–.89	.75–.91	.91–.97	.86–.95	.45–.78	.78–.92	.54–.83	.51–.81	.87–.96	.85–.95	.94–.98
SE of measurement	1.2	.7	1.3	1.1	6	6	4	6	7	5	6	6	8	16	2.2
Mean	6.4	6.8	17.9	18.1	48	47	68	70	52	65	65	68	32	79	27.8
Mean difference	-.1	2	.3	.4	1	-1	3	4	4	3	2	1	-2	-8	.5
95% CI of mean diff.	-6–.4	-1–.5	-2–.8	-1–.9	-1–.3	-2.8–1.8	1.2–4.6	1.6–6.0	1.2–6.7	1.3–5.1	-4–4	-1.0–3.9	-5.4–1.4	-14.3–1.1	-4–1.3
p-value	NS	NS	NS	NS	NS	NS	.002	.001	.006	.001	NS	NS	NS	.02	NS
Inter-rater reliability	+	+	+	+	+	+	-	-	-	-	+	+	+	-	+
30 patients															
ICC	.99	.94	.98	.97	.85	.88	.96	.96	.52	.81	.64	.68	.90	.96	.98
95% CI of ICC	.98–1.00	.90–.97	.93–.98	.95–.98	.74–.91	.81–.93	.95–.98	.94–.98	.36–.74	.69–.89	.44–.78	.49–.80	.85–.95	.92–.98	.96–.99
SE of measurement	1.4	.4	1.0	.9	6	5	4	4	8	5	6	7	6	10	2.1
Mean	9.2	6.6	16.4	16.8	46	43	64	65	48	60	61	66	16	54	24.6
Mean difference	.0	2	.1	-.2	1	2	2	2	5	4	2	-1	-3	-2	.3
95% CI of mean diff.	-8–.8	-1–.4	-5–.6	-7–.3	-1.6–4.3	-.9–4.3	1.3–3.9	2.4–2	8–8.9	9–6.3	-1.7–4.9	-4.1–3.2	-6.0–.2	-7.3–3.5	-8–1.4
p-value	NS	NS	NS	NS	NS	NS	.04	.04	.02	.01	NS	NS	.04	NS	NS
Inter-rater reliability	+	+	+	+	+	+	-	-	-	-	+	+	-	+	+
20 healthy subjects															
ICC	.95	.22	.79	.85	.75	.75	.84	.70	.59	.86	.66	.63	.86	.69	.95
95% CI of ICC	.92–.97	.07–.46	.68–.89	.84–.95	.59–.85	.64–.87	.78–.92	.62–.86	.40–.76	.80–.93	.49–.80	.58–.84	.76–.92	.59–.85	.92–.97
SE of measurement	.9	1.0	1.5	1.1	6	6	5	7	6	4	5	4	12	22	2.3
Mean	2.2	7.1	20.1	20.2	50	52	75	77	58	72	70	72	55	116	32.5
Mean difference	-.3	.3	.8	1.4	2	-.4	4	6	3	3	2	4	0	-16	.7
95% CI of mean diff.	-8–.3	-.4–.9	-.3–1.8	.6–2.1	-.2–4.5	-.7–8.3	.8–7.6	1.5–10.8	-1.2–6.3	.0–5.3	-1.0–5.2	1.6–7.0	-8.0–7.3	-30.7–2.1	-8–2.2
p-value	NS	NS	NS	.001	NS	NS	.02	.01	NS	.047	NS	.004	NS	.03	NS
Inter-rater reliability	+	-	+	+	+	+	-	-	-	-	+	-	+	-	+

ICC = Intra-class-correlation coefficient. NS = Not significant. SE = Standard error

five tests that required such fixation. The difference in proportion (five vs. two out of five tests) was significant ($p = .04$).

Examples of scatter plots showing acceptable and poor agreement, respectively, are shown in Figure 1a–b.

The exertion and the pain before and after examination 1 did not differ significantly from those before and after examination 2 (data not shown).

Intra-rater reliability

For examiner A (the physiotherapist), all ten tests had acceptable intra-rater reliability (Table 2). For examiner B (the research assistant), eight tests, i.e. all but the trunk rotation and the modified Biering-Sørensen, had acceptable intra-rater reliability (NS).

All the tests requiring no manual fixation had acceptable intra-rater reliability for both A and B. Of the five tests that required manual fixation, five and three tests had acceptable intra-rater reliability for A and B, respectively (NS).

Examples of scatter plots showing acceptable and poor agreement, respectively, are shown in Figure 2a–b.

The exertion and the pain before and after examinations 1 and 3 did not differ significantly between the ten healthy subjects of A and B (data not shown).

Reliability

All five tests requiring no manual fixation had acceptable reliability, i.e. acceptable inter-rater reliability, acceptable intra-rater reliability for both A and B and, if the test was composed of two sub-tests, acceptable inter- and intra-rater reliability for both sub-tests (Tables 1 and 2). Those tests were forward bending, modified Schober, lateral bending, abdominal endurance and modified PILE.

The five tests that required manual fixation – trunk rotation, active-straight-leg raise, cervical bending, cervical rotation and modified Biering-Sørensen – all had poor reliability except cervical rotation. The difference in proportion (five vs. one out of five tests) was significant ($p = .01$).

Discussion

The aim of this study was to answer the question: given a 10-test package for patients with prolonged back and neck pain, could an examiner without formal medical education be used without loss of quality?

Was the composition of the 10-test package suitable for answering this question? From our knowledge, there has been no previous reliability study involving a medically untrained examiner. However, numerous studies have

elucidated the problem of achieving agreement between medically skilled examiners, including both the choice of tests and the circumstances during which the examinations are performed. Some reliability studies include tests of inter-segmental mobility, i.e. passive mobility between two vertebrae levels [20]. Strender et al. [18] demonstrated the acceptable inter-rater reliability of such tests, provided that the examination situation is ideal. An ideal situation implies that the examiners have been able to standardize their techniques by working together for a sufficiently long period. In non-ideal conditions, Fjellner et al. [21] obtained acceptable inter-rater reliability in several tests of general motion but in few tests of inter-segmental mobility. As the everyday clinical situation is seldom ideal, we chose motion tests for our test package that exclusively concerned general mobility. The comparatively high proportion of tests with acceptable inter-rater reliability in our study (seven out of ten tests) supported this despite the non-ideal conditions. Notwithstanding the absence of previous references, it seems reasonable to predict that an examiner without medical education and practice will experience even greater difficulties in performing a standardized technique of manual fixation than an examiner with such skills. In support of this, the tests in our package that required fixation tended to have a higher proportion of acceptable intra-rater reliability for the physiotherapist than for the research assistant (five vs. three tests), though the difference was not significant. As a matter of fact, all the technically least advanced of our tests, i.e. the five that required no manual fixation by the examiner, had acceptable inter-rater reliability (five out of five tests). The proportion was significantly lower for the five tests requiring manual fixation (two out of five tests). This is consistent with the study of Bertilsson et al. [15], in which a simple sensitivity test had acceptable inter-rater reliability while several more sophisticated tests had not. The abdominal endurance had acceptable reliability, as against the study of Moreland et al. [26], in which the hands of the participant were held on the cheeks. In our study, as in the studies of Hyytiäinen et al. [11] and Lindström et al. [8], the hands were stretched out towards the patellae. The test package was inexpensive and easy to perform. Our study indicates, however, that Biering-Sørensen, when it is simplified as we described, has poor reliability. We found that the modified PILE had acceptable reliability, which complements the study of Lindström et al. [8]. They found this modification to have good validity, i.e. that the lifting capacity, when measured as described, correlated significantly with the rate of return to work, but their study included no test of reliability. Without exception, the five tests requiring no manual fixation had acceptable reliability. Five of the tests required such fixation, including the modified Biering-Sørensen and the previously unvalidated tests of trunk rotation and active-straight-leg raise. Only one of them (cervical rota-

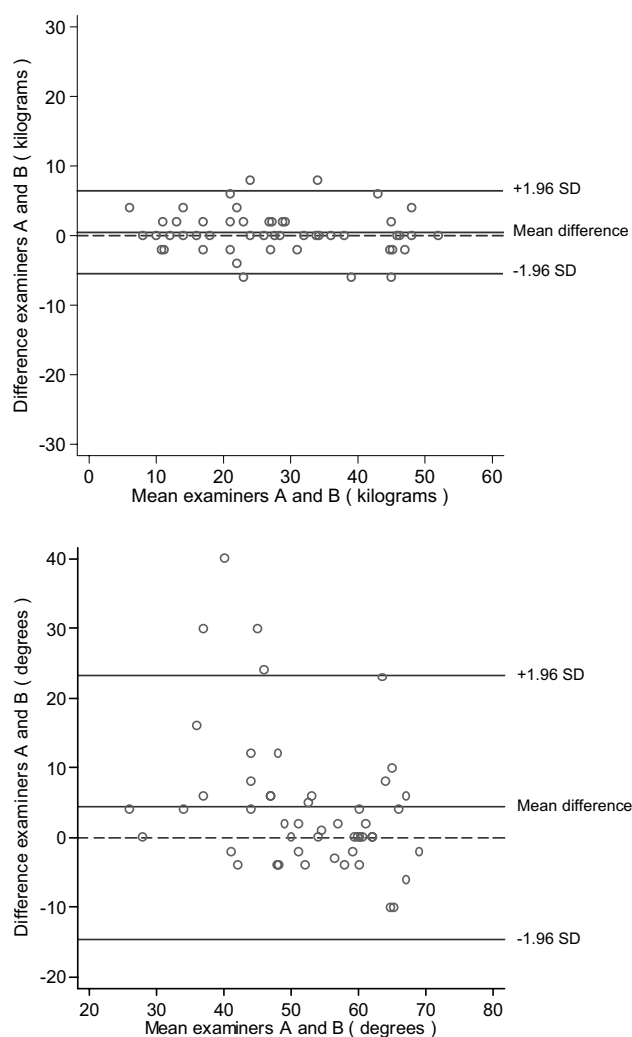


Figure 1

a-b. Inter-rater reliability. Fifty participants tested by A (the physiotherapist) and B (the research assistant). The difference between the measurements by A and B against the mean of the measurements by A and B with 95% limits of agreement (= the mean difference of the measurements with 95% CI). **I a. Modified PILE lumbar.** Acceptable agreement. The mean difference is close to the zero line, which indicates a small systematic error. The limits of agreement are narrow, which indicates a small random error. **I b. Cervical bending forward.** Poor agreement. The mean difference is fairly far from the zero line and the limits of agreement are wide, which indicates high systematic and random error.

Table 2: Intra-rater reliability. Twenty healthy subjects tested twice by A or B. Further explanations in Table 1.

I-o-test package (including 16 sub- tests):	Forward bending (cm)	Modified Schober (cm)	Lateral bending (cm)		Trunk rotation (°)		Active-straight-leg raise (°)		Cervical bending (°)		Cervical rotation (°)		Abdom. endurance (sec)	Med. Biering- Sorensen (sec)	Modified PILE (kg)	Lumbar	Cervical
			Right	Left	Right	Left	Right	Left	Forward d	Backward d	Right	Left					
Examiner A																	
ICC	.95	.87	.99	.94	.92	.96	.99	.97	.86	.98	.94	.86	.90	.92	.93		.95
95% CI of ICC	.89-.99	.68-.96	.95-1.00	.82-.98	.76-.97	.87-.99	.96-1.00	.92-.99	.67-.96	.95-.99	.82-.98	.63-.95	.75-.97	.75-.97	.80-.98		.86-.98
SE of measurement	.9	.3	.5	1.0	3	3	2	3	2	2	2	3	9	16	2.3		1.5
Mean	2.5	7.1	21.2	21.0	55	53	75	78	58	75	72	74	66	117	31.8		20.8
Mean difference	-.7	.2	-.1	.1	1	-.1	0	1	2	1	1	-.1	7	4	.8		.4
95% CI of mean diff.	-1.6-.1	-.1-1.5	-.5-.5	-.9-1.1	-.26- 3.6	-.38- 1.8	-1.6-1.4	-1.7-3.3	-.8-4.0	-.5-2.5	-.14- 3.0	-.38- 1.4	-2.7-15.7	-12.3-19.3	-1.6-3.2		-1.1-1.9
p-value	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS
Intra-rater reliability	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+
Examiner B																	
ICC	.95	.79	.73	.86	.46	.83	.97	.93	.62	.80	.82	.82	.65	.20	.97		.94
95% CI of ICC	.86-.98	.46-.93	.37-.91	.61-.95	.13-.85	.54-.94	.90-.99	.78-.97	.12-.85	.53-.94	.53-.94	.52-.94	.18-.86	.14-.85	.89-.99		.83-.98
SE of measurement	.9	.7	1.6	1.4	7	5	3	4	6	5	4	4	17	17	2.4		2.9
Mean	1.8	7.2	19.8	19.7	48	51	70	76	57	67	68	67	46	104	32.8		23.5
Mean difference	.4	.2	-.7	-.2	-.8	-.1	0	-.1	-.1	3	-.1	-.1	3	33	.0		-.1.8
95% CI of mean diff.	-.5-1.3	-.5-.9	-.23-.9	-.16- 1.2	-.14.7- 3	-.63- 4.3	-.2.8-3.0	-.4.9-3.5	-.6.7- 5.6	-.2.0-7.6	-.5.4- 2.6	-.4.7- 2.7	-.14.3-20.5	15.3-50.5	-.2.4-2.4		-.4.7-1.1
p-value	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	.002	NS		NS
Intra-rater reliability	+	+	+	+	-	+	+	+	+	+	+	+	+	-	+		+

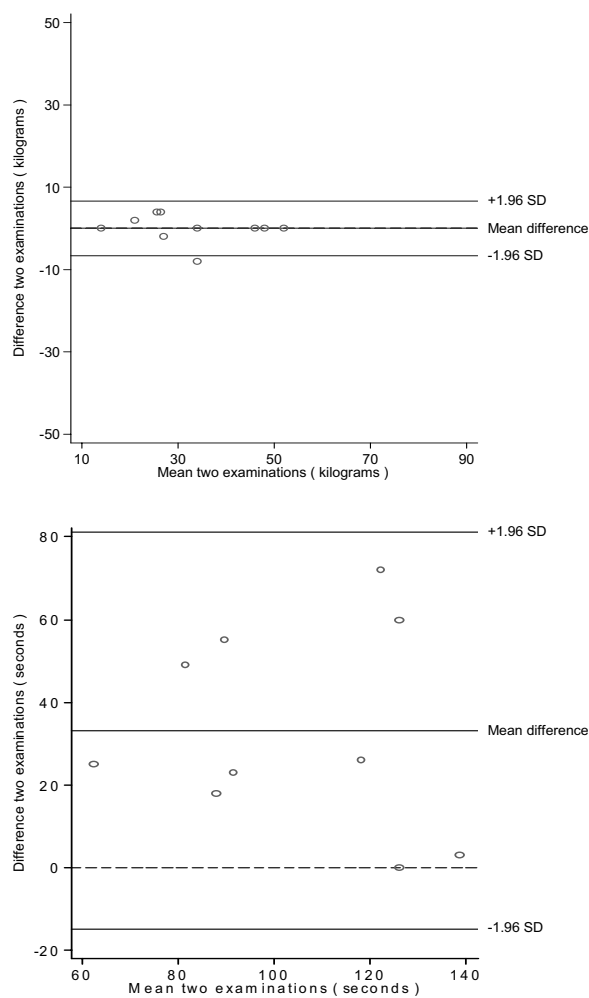


Figure 2

a-b. Intra-rater reliability. Ten healthy subjects tested twice by B. The difference between the two examinations against the mean of the two examinations with 95% limits of agreement. Further explanations in Figure 1 a-b. **2 a. Modified PILE lumbar.** Acceptable agreement. The mean difference is identical to the zero line, which indicates a very small systematic error. The limits of agreement are narrow, which indicates a small random error. **2 b. Modified Biering-Sørensen.** Poor agreement. The mean difference is far from the zero line and the limits of agreement are very wide, which indicates high systematic and random error.

tion) had acceptable reliability. This difference (five vs. one) was significant ($p = .01$). All tests requiring no manual fixation had acceptable intra-rater reliability for both A and B. Concerning the composition of our test package, it seemed right to include motion tests exclusively concerned with general mobility, but we underestimated the technical difficulties of manual fixation. Thus, the composition of the 10-test package proved to be fairly suitable for answering the question of this study, indicating *inter alia* that an examiner without formal medical education should not perform tests that require manual fixation, with the possible exception of cervical rotation. Abdominal endurance should be tested in the same way as in our study; the Biering-Sørensen test with our modification should not be used; and the modified PILE used in this study could be recommended.

Although the difference was not significant, the proportion of tests with acceptable inter-rater reliability tended to be higher for the patients than for the healthy subjects (seven vs. four tests). That is in line with previous research [19,21]. The intra-rater reliability of the package tended to be greater than the inter-rater reliability, which also corresponds with other studies [19,35]

The study has several limitations, which diminish the generalizability of the results. One weakness was that the gold standard consisted of one single physiotherapist. For example, the active-straight-leg-raise and cervical bending showed an acceptable intra-reliability for both the physiotherapist and the research assistant, while the inter-reliability for those tests was poor (see Table 1 and 2). The reason for that could, hypothetically, be that the research assistant, not the physiotherapist, performed those tests more reliably. However, the substantially narrower 95% CI and lower SEM of the physiotherapist (see Table 2) indicate the opposite. Also, the use of only one examiner without medical education is a limitation. The total lack of previous references concerning the use of examiners without medical education makes it difficult to evaluate the representativeness of the medically untrained examiner of our study. Another weakness was that the intra-reliability study only included a comparatively small number of healthy subjects. A way to overcome the ethical and methodological difficulties of using patients for as many as three examinations is to spread them out over several days, as in the studies of Ljungquist et al. [4] and Horneij et al. [13]. This option, however, was beyond the limits of the resources of our study. The intra-rater reliability study was limited to ten participants for each examiner. Ljungquist et al. [4] used as few as 11 healthy subjects in one of the two samples for studying the intra-rater reliability of an 11-test package. They all performed all the tests on every test occasion, which made a valuable contribution to the comprehensive assessment of the package. In

the other sample used by Ljungquist et al., 24 patients with back or neck pain were engaged. Although the examinations were distributed over three different days, only 16 of them performed all 11 tests each time, mainly because of pain. This illustrates the problems involved in engaging patients in numerous examinations.

Notwithstanding its limitations, this study indicates that even an examiner with no formal medical education could be used without loss of quality, at least for tests that require no manual fixation. This might produce a better assessment of outcome at defensible cost and might also be useful in a research context. To make our results more generalizable and their implications more searching, a similar study should be conducted with two or more examiners with and without formal medical education, and the intra-rater reliability study should also include patients and involve more participants.

When the complete data of the randomized controlled trial (see Background) are available, the measurement results of the tests with poor reliability should be interpreted with caution.

Conclusion

Given a 10-test package for patients with prolonged back and neck pain, an examiner without formal medical education could be used without loss of quality, at least for the five tests that require no manual fixation. This might produce a better assessment of outcome at defensible cost and might also be useful in a research context. To make our results more generalizable and their implications more searching, a similar study should be conducted with two or more examiners with and without formal medical education, and the intra-rater reliability study should also include patients and involve more participants.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

OL and LE participated in the design of the study and coordinated and monitored the performance of the tests and data collection. OL performed the statistical analyses and prepared drafts of the manuscript. LES, as supervisor for OL and LE, participated in all phases of the study. All authors read and approved the final manuscript.

References

1. Alexanderson K, Norlund A, (editors): **SBU summary and conclusions. Sickness absence – causes, consequences, and physicians' sickness-certification practice. A systematic literature review by the Swedish Council on Technology Assessment in Health Care. Scand J Public Health** 2004:6-11.
2. Waddell G, Main CJ: **Assessment of severity in low back disorders. Spine** 1984, **9**:204-208.
3. Waddell G: **Diagnostic triage. In Back Pain Revolution** 2nd edition. Edited by: Waddell G. Edinburgh: Churchill Livingstone; 2004:9-26.

4. Ljungquist T, Harms-Ringdahl K, Nygren A, Jensen I: **Intra- and inter-rater reliability of an I I-test package for assessing dysfunction due to back or neck pain.** *Physiother Res Int* 1999, 4:214-232.
5. **Task force on standards for measurement in physical therapy. Standards for tests and measurements in physical therapy practice.** *Phys Ther* 1991, 71:589-622.
6. Biering-Sørensen F: **Physical measurements as risk indicators for low-back trouble over a one-year period.** *Spine* 1984, 9:106-109.
7. Mellin G: **Accuracy of measuring lateral flexion of the spine with a tape.** *Clin Biomech* 1986, 1:85-89.
8. Lindström I, Ohlund C, Eek C, Wallin L, Petersen LE, Nachemson A: **Mobility, strength and fitness after a graded activity program for patients with subacute low back pain.** *Spine* 1992, 17:641-652.
9. Capuano-Pucci D, Rheault W, Aukai J, Bracke M, Day R, Pastrick M: **Intratester and intertester reliability of the cervical range of motion device.** *Arch Phys Med Rehabil* 1991, 72:338-340.
10. Rheault W, Albright B, Byers C, Franta M, Johnson A, Skowronek M, Dougherty J: **Intertester reliability of the cervical range of motion device.** *J Orthop Sports Phys Ther* 1992, 15:147-150.
11. Hyytiäinen K, Salminen JJ, Suviö T, Wickström G, Pentti J: **Reproducibility of nine tests to measure spinal mobility and trunk muscle strength.** *Scand J Rehabil Med* 1991, 23:3-10.
12. Moreau CE, Green BN, Johnson CD, Moreau SR: **Isometric back extension endurance tests: A review of the literature.** *J Manipulative Physiol Ther* 2001, 24:110-122.
13. Horneij E, Holmström E, Hemborg B, Isberg PE, Ekdahl C: **Inter-rater reliability and between-days repeatability of eight physical performance tests.** *Adv Physiother* 2002, 4:146-160.
14. Hawk C, Phongphua C, Bleecker J, Swank L, Lopez D, Rubley T: **Preliminary study of the reliability of assessment procedures for indications for chiropractic adjustments of the lumbar spine.** *J Manipulative Physiol Ther* 1999, 22:382-389.
15. Bertelson BC, Grunnesjö M, Strender LE: **Reliability of clinical tests in the assessment of patients with neck/shoulder problems – impact of history.** *Spine* 2003, 28:2222-2231.
16. Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ: **Progressiv isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population.** *Spine* 1988, 13:998-1002.
17. Nitschke JE, Nattrass CL, Disler PB, Chou MJ, Ooi KT: **Reliability of the American Medical Association Guides' model for measuring spinal range of motion. Its implication for whole-person impairment rating.** *Spine* 1999, 24:262-268.
18. Strender LE, Sjöblom A, Sundell K, Ludwig R, Taube A: **Interexaminer reliability in physical examination of patients with low back pain.** *Spine* 1997, 22:814-820.
19. Keeley J, Mayer T, Cox R, Gatchel R, Smith V, Mooney V: **Quantification of lumbar function. Part 5: Reliability of range-of-motion measures in the sagittal plane and an in vivo torso rotation measurements technique.** *Spine* 1986, 11:31-35.
20. Binkley J, Stratford PW, Gill C: **Interrater reliability of lumbar accessory motion mobility testing.** *Phys Ther* 1995, 75:786-795.
21. Fjellner A, Bexander C, Faleij R, Strender LE: **Interexaminer reliability in physical examination of the cervical spine.** *J Manipulative Physiol Ther* 1999, 22:511-516.
22. Macrae IF, Wright V: **Measurement of back movement.** *Ann Rheum Dis* 1969, 28:584-589.
23. Borg G: **Psychophysical bases of perceived exertion.** *Med Sci Sports Exerc* 1982, 14:377-382.
24. Borg G: **A category scale with ratio properties for inter-modal and inter-individual comparison.** Berlin: VEB Deutscher Verlag der Wissenschaften; 1982:25-34.
25. Bland JM, Altman DG: **Comparing two methods of clinical measurement: A personal history.** *Int J Epidemiol* 1995:7-14.
26. Moreland J, Finch E, Stratford P, Balsor B, Gill C: **Interrater reliability of six tests of trunk muscle function and endurance.** *J Orthop Sports Phys Ther* 1997, 26:200-208.
27. Valentine RE, Lewis JS: **Intraobserver reliability of 4 physiologic movements of the shoulder in subjects with and without symptom.** *Arch Phys Med Rehabil* 2006, 87:1242-1249.
28. Haas M: **Statistical methodology for reliability studies.** *J Manipulative Physiol Ther* 1991, 14:119-132.
29. Rosner M: **Multisample inference.** In *Fundamental of biostatistics* 5th edition. Duxbury: Thomson Learning; 2000:511-576.
30. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, 33:159-174.
31. Chinn S: **Statistics in respiratory medicine. 2. Repeatability and method comparison.** *Thorax* 1991, 46:454-456.
32. Altman DG: **Comparing groups – categorical data.** In *Practical statistics for medical research* 1st edition. London: Chapman & Hall; 1991:229-276.
33. Bland JM, Altman DG: **Measuring agreement in method comparison studies.** *Stat Methods Med Res* 1999, 8:135-160.
34. Altman DG: **Comparing groups – continuous data.** In *Practical statistics for medical research* 1st edition. London: Chapman & Hall; 1991:179-228.
35. Hendriks EJM, Brandsma JW, Heerkens YF, Oostendorp RAB, Nelson RM: **Intraobserver and interobserver reliability of assessments of impairments and disabilities.** *Phys Ther* 1997, 77:1097-1106.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2474/8/31/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



III

Research article

Open Access

Subacute and chronic, non-specific back and neck pain: cognitive-behavioural rehabilitation versus primary care. A randomized controlled trial

Odd Lindell*, Sven-Erik Johansson and Lars-Erik Strender

Address: Center for Family and Community Medicine, Karolinska Institutet, Alfred Nobels allé 12, SE-141 83 Huddinge, Sweden

Email: Odd Lindell* - odd.lindell@ki.se; Sven-Erik Johansson - sven-erik.johansson@ki.se; Lars-Erik Strender - lars-erik.strender@ki.se

* Corresponding author

Published: 30 December 2008

Received: 23 May 2008

BMC Musculoskeletal Disorders 2008, 9:172 doi:10.1186/1471-2474-9-172

Accepted: 30 December 2008

This article is available from: <http://www.biomedcentral.com/1471-2474/9/172>

© 2008 Lindell et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**Abstract**

Background: In the industrial world, non-specific back and neck pain (BNP) is the largest diagnostic group underlying sick-listing. For patients with subacute and chronic (= full-time sick-listed for 43 – 84 and 85 – 730 days, respectively) BNP, cognitive-behavioural rehabilitation was compared with primary care. The specific aim was to answer the question: within an 18-month follow-up, will the outcomes differ in respect of sick-listing and number of health-care visits?

Methods: After stratification by age ($\leq 44/\geq 45$ years) and subacute/chronic BNP, 125 Swedish primary-care patients were randomly allocated to cognitive-behavioural rehabilitation (rehabilitation group) or continued primary care (primary-care group). Outcome measures were *Return-to-work share* (percentage) and *Return-to-work chance* (hazard ratios) over 18 months, *Net days* (crude sick-listing days \times degree), and the number of *Visits* (to physicians, physiotherapists etc.) over 18 months and the three component six-month periods. Descriptive statistics, Cox regression and mixed-linear models were used.

Results: All patients: *Return-to-work share* and *Return-to-work chance* were equivalent between the groups. *Net days* and *Visits* were equivalent over 18 months but decreased significantly more rapidly for the rehabilitation group over the six-month periods ($p < .05$). Subacute patients: *Return-to-work share* was equivalent. *Return-to-work chance* was significantly greater for the rehabilitation group (hazard ratio 3.5 [95%CI 1.001 – 12.2]). *Net days* were equivalent over 18 months but decreased significantly more rapidly for the rehabilitation group over the six-month periods and there were 31 days fewer in the third period. *Visits* showed similar though non-significant differences and there were half as many in the third period. Chronic patients: *Return-to-work share*, *Return-to-work chance* and *Net days* were equivalent. *Visits* were equivalent over 18 months but tended to decrease more rapidly for the rehabilitation group and there were half as many in the third period (non-significant).

Conclusion: The results were equivalent over 18 months. However, there were indications that cognitive-behavioural rehabilitation in the longer run might be superior to primary care. For subacute BNP, it might be superior in terms of sick-listing and health-care visits; for chronic BNP, in terms of health-care visits only. More conclusive results concerning this possible long-term effect might require a longer follow-up.

Trial registration: NCT00488735.

Background

In Sweden, as all over the industrial world, back and neck pain is the largest diagnostic group underlying sick-listing [1]. The vast majority consists of non-specific back and neck pain (BNP) that requires no specific surgical, rheumatological or neurological treatment [2].

As 93% of the societal costs of back and neck pain are connected with sick-listing [3], return-to-work is crucial [4]. However, there is a lack of consistency and comprehensiveness in return-to-work measurements [5]. While earlier studies compared the return-to-work share at a specific time point, for example one year after baseline [6], later research has evaluated the time of return-to-work in survival analyses [7,8]. Another important issue is the health-care utilization needed to achieve certain treatment results. In that respect, a frequently-used outcome measure is the number of health-care visits [9,10].

Concerning treatment of BNP, the 1990s saw a breakthrough for the biopsychosocial model, which pinpoints time off work as an important disabling factor. Acute, subacute and chronic BNP are defined as BNP with full-time sick-listing for 0 – 21 days, 22 – 84 days and more than 12 weeks, respectively [11]. Acute BNP is managed by continuing ordinary activities as normally as possible, and manipulation if necessary. In cases of subacute and chronic BNP, multidisciplinary rehabilitation should be considered [12]. Multidisciplinary treatment includes a physician's consultation in addition to psychological, social or vocational intervention or a combination of these [13]. The three key components of successful multidisciplinary rehabilitation programmes for BNP are: reactivation and progressive increase in activity levels, addressing dysfunctional beliefs and behaviour by a cognitive-behavioural therapeutic approach, and occupational interventions [4]. Concerning back pain, programmes including these items have shown good results in several studies [7,14-17]. Randomized controlled trials have concerned patients with subacute back pain only [7-9,14,15,17,18], mixed groups with subacute or chronic back pain [16,19] or patients with chronic back pain only [20]. There is a serious lack of evidence concerning the rehabilitation of neck pain [13]. We have found no randomized controlled trial in which the same programme was offered to patients who were stratified by subacute and chronic BNP.

The high frequency of relapses after rehabilitation of BNP is associated with inadequate follow-ups. A short program might fail to achieve long-standing behavioural changes [21]. In the 1990s the vast majority of rehabilitation programs in Sweden were comparatively short, with a fixed duration averaging six weeks [22].

Primary care is the appropriate source of treatment for BNP [12]. In Sweden, however, notwithstanding clinical guidelines, only a small minority of individuals with subacute and chronic BNP receive multidisciplinary rehabilitation [23]. One reason might be the relative lack of family doctors. While the total number of Swedish physicians meets international standards, there are proportionately fewer physicians within primary care: the density of family doctors is .5 per 1000 population, compared with an OECD (Organisation for Economic Co-operation and Development) average of .8 [24].

Our project started in 2000 with the aim of comparing a multidisciplinary programme of cognitive-behavioural rehabilitation for subacute and chronic BNP with primary care. The specific aim of this study was to answer the question: within an 18-month follow-up, will the outcomes differ in terms of sick-listing and number of health-care visits?

Methods

Sick-listing in Sweden

In Sweden, publicly provided, tax-financed social insurance compensates loss of income due to illness. The ultimate decisions about sick-listing benefits, including sickness benefit, rehabilitation benefit, temporary disability pension and disability pension, are made by the Social Insurance Agency. For sick-listing exceeding seven calendar days, a physician's certificate is required. The certificate comprises a detailed description of symptoms and signs and a recommendation of the degree (.25, .50, .75 or 1.00 (= full-time)) and duration of sick-listing.

Participants

The rehabilitation centre of this study was situated at Hänge, a municipality 25 kilometres south-east of Stockholm city. As the centre was well known to the local residents, the study participants were recruited within the primary care of the adjoining municipalities. One-hundred-and-twenty five patients were recruited by 42 family doctors at 12 health centres.

The criteria for inclusion: 1. Working age up to and including 59 years. 2. Sick-listed full-time for BNP at least six weeks (42 days) and at most two years (730 days). 3. Able to fill in forms. *The criteria for exclusion:* 1. Temporary disability pension or disability pension being paid or in preparation. 2. A primary need for a hospital specialist (for example, operation for slipped disc). 3. Pregnancy and diseases (other than BNP) that might make rehabilitation impracticable (for example, advanced pulmonary disease). 4. Whiplash-associated disorders as a primary obstacle to working. 5. Previous rehabilitation at the rehabilitation centre. 6. Other multidisciplinary rehabilitation current or planned.

Interventions

One treatment group was allocated to cognitive-behavioural rehabilitation at the rehabilitation centre (rehabilitation group). The other treatment group was allocated to continued primary care (primary-care group).

Cognitive-behavioural rehabilitation

The rehabilitation centre was opened in 1991 within Stockholm County Council. From 2002 it operated as a private company and the number of rehabilitation teams was decreased from four to one, comprising four team members: a physician (OL), a physiotherapist trained in manual therapy, a psychologist or a social worker trained in cognitive-behavioural therapy and a health-care adviser. Manual therapy includes manipulation, mobilisation and stabilizing training [25]. The centre used a cognitive-behavioural programme with the aim of achieving the maximal degree of work ability lasting for at least 30 consecutive days. Work ability was inversely proportional to sick-listing, which is the definition used by the Social Insurance Agency. Work abilities of 1.00 (= full-time), .75, .50 and .25 corresponded to sick-listings of 0, .25, .50 and .75, respectively. Zero work ability equalled full-time sick-listing. Possible relapses were met by individual and, when needed, long rehabilitation periods. The program is described in Table 1.

Participation in the rehabilitation group did not exclude the patient from seeking other care, including primary care, during the follow-up period.

Primary care

The hubs of Swedish primary care are the health centres. They serve the local population and cater to its needs, with no restrictions as to illness, age or patient category, for basic medical treatment, nursing, preventive work or rehabilitation that does not require the medical and technical resources of hospitals or other special competences [26]. Most primary care in Sweden is publicly provided. Only a quarter is privately conducted [27]. Overall medical responsibility belongs to the family doctor. The 12 health centres in this study were situated in the municipalities of Tyresö, Huddinge, Stockholm and Nynäshamn. Ten of the centres were publicly provided, two were private. In total, they engaged 84 family doctors and served a population of 148,000 individuals, equivalent to barely .6 family doctors per 1000 population. Besides family doctors, their staff consisted of physiotherapists, nurses, assistant nurses, occupational therapists and social workers. Besides management at the health centre, primary care could include referral to consultation by, for example, an orthopedist or a neurologist.

Participation in the primary-care group excluded the patient from turning to the rehabilitation centre during

the follow-up period but not from any other health-care, including multidisciplinary rehabilitation at units other than the rehabilitation centre.

Outcome measures

Return-to-work share

The percentage of patients who regained any degree of work ability for at least 30 days in succession over 18 months. This was the primary outcome measure. Secondary outcome measures were:

Return-to-work chance

The chance, as expressed in hazard ratios, of achieving any degree of work ability over 18 months, irrespective of the duration of that work ability.

Net days

Sick-listing, expressed in whole days, over 18 months and the three component six-month periods. *Net days* = crude days \times degree [28].

Visits

The total number of health-care visits over 18 months and over the three component six-month periods. *Visits* comprised consultations at the rehabilitation centre, within primary care and other care, including alternative-care providers, but excluded consultations relating to multidisciplinary rehabilitation at units other than the rehabilitation centre.

Analyses and statistics

Except for descriptive statistics [29,30], Cox regression and mixed-linear models were used.

Return-to-work chance was compared by a Cox regression analysis for recurrent events with event dependence and a time interaction with the exposure variable (i.e. rehabilitation group or primary-care group) and is presented as hazard ratios with 95% confidence intervals [31]. It was analysed at six, 12 and 18 months.

Net days and *Visits* in the first, second and third six-month periods were outcome variables in two separate mixed-linear models. In the models, the main effects of three explanatory variables and two interaction terms were compared using a random intercept model of the unstructured covariance type on the group level and time as repeated factor [32]. The explanatory variables were time (i.e. six-month period 1, 2 or 3), rehabilitation group or primary-care group, and subacute or chronic patient. The interaction terms were time \times rehabilitation group or primary-care group and time \times rehabilitation group or primary-care group \times sub-acute or chronic. The models were also adjusted for possible baseline characteristics with significant differences between the groups. The analyses were

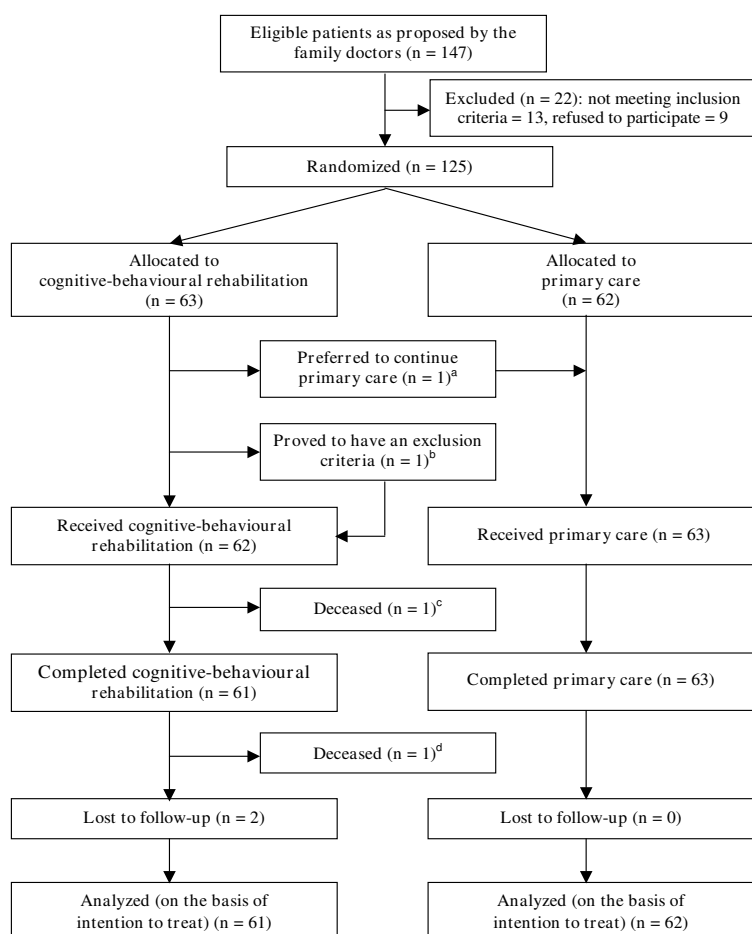
Table 1: Cognitive-behavioural rehabilitation.

Staff category	Investigation and treatment phase, 2 – 8 weeks	Frequency
Physician	Mapping out of medical obstacles to working. Handling of the sick-listing. If needed, prescription of drugs (antidepressants, analgesics etc.) and injections of cortisone (in shoulder- or hip-muscle attachments etc.) [25].	1 – 2 (consultations)/week.
Physiotherapist	Mapping out of biomechanical obstacles to working including a visit to the work place [14]. Start of graded activity: the patient first carried out an activity measurable in minutes, metres, etc., for example a walk, until the pain increased. The starting level was about 25% below that. A gradual increase of the activity was decided on check-ups, the final aim being to manage the load in a job, for the unemployed an imaginary one [14]. If needed, manual therapy [25].	2 – 3 consultations. 1/week. 1/week.
Psychologist or social worker	Mapping out of psychosocial obstacles to working. Cognitive-behavioural therapy focussed on anxiety and depression [46].	1/week.
Health-care adviser	Start of education in applied relaxation [46].	1/week for 6 – 8 w.
Action phase, 2 – 8 months		
Team	Conference that produced a written rehabilitation plan with: 1. <i>Final aim</i> = the optimal degree of work ability that could be achieved and maintained for at least 30 consecutive days. 2. <i>Partial aims</i> concerning functioning only (for example, increase of vocational training by five hours/week); symptom aims, for example, pain reduction, were excluded [14]. 3. <i>Means of reaching the aims</i> (for example, increase of vocational training 1/2 hour/day week 1, 1 hour/day w. 2 etc.).	At the start of the action phase.
Team	Check-up conferences produced fresh partial aims.	1/3 – 4 weeks.
Team member (usually the physiotherapist)	Vocational conferences with the employer and a clerk from the Social Insurance Agency or, for unemployed patients, the Employment Office.	
Physician	Handling of the sick-listing.	1/3 – 4 weeks.
Physiotherapist	Completion of graded activity. Check-ups less frequent.	1/3 – 4 weeks.
Health-care adviser	Completion of education in applied relaxation.	1/week (f. 6 – 8 w.)
Psychologist or social worker	If needed: cognitive-behavioural therapy as support during the re-training process.	1/week.
	When the final aim was reached, or when it was obvious that return-to-work would not be achieved.	The end of rehabilitation.

performed using PROC MIXED in SAS, version 9.1, and the results are presented as separate graphs for the subacute and chronic patients and as means with 95% confidence intervals and *p*-values, adjusted for all parameters (main effect and interactions).

The two patients who died (Figure 1) were excluded from the outcome analyses except from the Cox regression [31].

Visits at 18 months were analysed for those patients who had completed all the follow-up forms, while the mixed-linear model also included incomplete responders. To evaluate their possible influence on the treatment results, we also analysed the days of hospital care, the use of surgery for musculoskeletal disorders and multidisciplinary rehabilitation at units other than the rehabilitation centre.



^a Woman, 58 (years), randomized to cognitive-behavioural rehabilitation, but preferred to continue at the health centre.

^b Male, 45, incorrectly included: except BNP, he suffered from whiplash-associated disorders that during the initial mapping out (Table 1) showed to be a primary obstacle to working.

^c Male, 55, died 12 months after inclusion from lung cancer.

^d Male, 53, died 11 months after inclusion of a reason which was unknown to us.

All these four patients had chronic BNP.

Figure 1
Flowchart.

The analyses were performed on an intention-to-treat basis. The primary outcome measure was also subjected to a per-protocol analysis [33]. The total percentage of withdrawals and drop-outs was calculated. This sum should not exceed 30% [34]. Baseline characteristics of responders and non-responders were compared. A p -value $< .05$ or, concerning the Cox regression, a 95% confidence interval not including 1.00, was considered statistically significant. Except for the mixed linear models, analyses were performed using Stata, 9.1.

Blinding

The analyst of the sick-listing data was blind to the intervention alternative. Blinding was not possible for the other outcomes. For example, which of the two interventions was offered could not be concealed from either the care providers or the patients.

Data collection

The sick-listing data were provided by the Stockholm County Social Insurance Agency. Data concerning the rehabilitation centre were collected from the medical records of the centre. Primary care and other health-care data were obtained from follow-up forms. Although these self-report measures have been used successfully in previous research, their reliability and validity have not been established. However, because the patients were free to seek treatment at any other facility, the only comprehensive sources of health-care data were self-ratings [9]. The data were fed into a specially designed database using Access version 2000.

Power calculation

To calculate the power, a preliminary study was performed. In this retrospective study, 172 consecutive patients with subacute and chronic BNP, who completed rehabilitation at the centre during the period 1996 – 2000, were included. The mean rehabilitation period was 266 (SD \pm 170) days. The *Return-to-work share* was 76%; for subacute and chronic BNP 89% and 73%, respectively ($p < .05$). The power calculation was based on this preliminary study and a forecast of the probability of return-to-work after traditional care for BNP [35]. The forecast probability for the patients in the preliminary study was calculated from their current sick-listing at baseline. It proved to be 49%, i.e. 27 percentage units less than the actual rate of 76%. Including an uncertainty about the application of this forecast to our patient sample, we expected to reach a difference between the rehabilitation group and the primary-care group of at least 22 percentage units. With an alpha of .05 and a power of 80%, this should require the inclusion of 154 patients; or, to allow a reasonable drop-out rate, 170 patients.

Inclusion procedure

For the patients who fulfilled the criteria, the family doctor gave verbal and written information about the project. Each patient who gave his or her oral consent to participate to the family doctor was interviewed by telephone by a research assistant within two days. The patients who still qualified for the study saw the assistant at the health centre within five days. At the appointment, the patient signed an informed consent to participate and went through an initial form including, among other items, the baseline characteristics in Table 2. Then the assistant carried out, among other tests, a lift test [36]. The reliability of that test procedure was confirmed in a separate study [37]. After stratification by age ($\leq 44/\geq 45$ years) and subacute/chronic BNP, the assistant performed the randomization. The two treatment alternatives were distributed in opaque envelopes by a computerized block-randomization procedure produced by an independent statistician. The assistant opened the remaining envelope with the lowest random number and presented the content to the patient.

Ethical approval

Approval for the study was given by The Research Ethics Committee, Karolinska University Hospital, Huddinge.

Premature cessation of recruitment

The recruitment of participants started in August 2000 and was discontinued in January 2004, when 125 patients were included. The reason was the opening in April 2004 of a large back-rehabilitation centre in a neighbouring municipality (Nacka) on the initiative of the Stockholm County Social Insurance Agency and Stockholm County Council. We presumed that many future study patients who would be randomized to the primary-care group would be referred to that centre and would contaminate the primary-care branch of our study.

Follow-up

Six, 12 and 18 months after inclusion, the patients completed forms concerning, among other items, health-care utilization. If necessary, a postal reminder was sent after two weeks and a telephone reminder after another two weeks. If the forms were not returned despite these measures, the data were considered missing. The patient who was last to be included completed the 18-month follow-up period in July 2005.

Results

Response rate and missing data

Data for the baseline characteristics, sick-listing and care at the rehabilitation centre were complete. For other health-care data, the response rates for the six-, 12- and 18-month forms in the rehabilitation group ($n = 61$) were 57 (93%), 56 (92%) and 55 (90%) respectively and all

Table 2: Baseline characteristics.

	Rehabilitation group (n = 63)	Primary-care group (n = 62)	p-value
Women	33 (52 [40 – 65]%)	35 (56 [44 – 69]%)	NS
Age (years)	42.2 [39.8 – 44.6]	43.0 [40.4 – 45.7]	NS
Neck-pain domination	17 (27 [16 – 38]%)	21 (34 [22 – 46]%)	NS
Widespread (= back + neck) pain	55 (87 [79 – 96]%)	45 (73 [61 – 84]%)	.04
Pain score (VAS, 0 – 100; median (IQR))[48]:			
"Just now"	61 (30)	53 (30)	NS
"Worst last week"	77 (29)	73 (26)	NS
Health-related quality of life (EQ-5D)[49]			
(median (IQR))	.489 (.332)	.497 (.332)	NS
Immigrants (= born outside Sweden)	19 (30 [19 – 42]%)	15 (24 [13 – 35]%)	NS
Single life	19 (30 [19 – 42]%)	21 (34 [22 – 46]%)	NS
Low education (= at most junior high school)	37 (60 [47 – 72]%)	35 (56 [44 – 69]%)	NS
Blue-collar work (of the non-unemployed)	41 (87 [77 – 97]%)	47 (87 [77 – 97]%)	NS
Unemployed	14 (22 [12 – 33]%)	15 (24 [13 – 35]%)	NS
Previous sick-listing (days)*	223 [189 – 257]	222 [188 – 256]	NS
Lifting capacity (kg; mean):			
PILE lumbar [36]	12.3 [10.4 – 14.2]	12.4 [10.3 – 14.6]	NS
PILE cervical [36]	11.5 [9.7 – 13.3]	11.6 [9.6 – 13.6]	NS

Descriptive statistics. The 95% confidence intervals are shown within brackets. Bold figures indicate a significant difference.

NS = Non-significant; IQR = Inter-quartile-range.

* = Net days over the 18 months preceding baseline.

forms were answered by 51 patients (84%). The corresponding rates for the primary-care group (n = 62) were 50 (81%), 48 (77%), 50 (81%) and 42 (68%). Non-responders and responders are compared in Table 3.

Baseline characteristics and participant flow

Except for a higher prevalence of widespread pain in the rehabilitation group, there were no significant differences (Table 2). When analyzed separately (data not shown), the subacute rehabilitation-group patients were equal to the subacute primary-care-group patients while the chronic rehabilitation-group patients had a much higher prevalence of widespread pain: 93 [85 – 100]% versus 68

[54 – 82]% for the chronic primary-care-group patients ($p = .004$).

Patients who were allocated to the rehabilitation group started the programme within one week. Patients who were allocated to the primary-care group continued care at their health-centres. Sixty-one patients in the rehabilitation group completed cognitive-behavioural rehabilitation; all primary-care-group patients completed primary care (Figure 1). The two deceased rehabilitation-group patients had passed the "red-flags" examinations [12] at the start without remark.

Table 3: Missing data.

Follow-up	Six months	p-value	12 months	p-value	18 months*	All forms	p-value
Rehabilitation group (n = 61)							
Previous sick-listing (days)**	397 vs. 215	.008	371 vs. 214	.01	-	-	
Current sick-listing at baseline (days)	367 vs. 158	< .001	346 vs. 156	< .001	-	275 vs. 151	.003
Unemployment (%)	-	-	60 vs. 18	.03	-	-	
Primary-care group (n = 62)							
Age (years)	35.8 vs. 44.8	.006	-	-	-	38.3 vs. 45.3	.01
Single (%)	58 vs. 28	.046	-	-	-	-	
EQ-5D [49]	-	-	.357 vs. .562	.046	-	-	

Non-responders versus responders. Significant differences at baseline. Descriptive statistics.

*At 18 months there were no significant differences.

** = Net days over the 18 months preceding baseline.

Outcome measures

Return-to-work share

There were no significant differences between the rehabilitation group and the primary-care group, or between the subacute and chronic patients considered separately (Table 4). In both the rehabilitation group and the primary-care group, most of the patients who regained any degree of work ability returned to full-time work: 20/35 (57%) and 25/35 (71%) respectively (non-significant). The mean degrees of work ability at return to work were .77 [.67 – .87] and .85 [.76 – .94] respectively (non-significant).

Return-to-work chance

The hazard ratio for the rehabilitation group increased over the three six-month periods in comparison to the primary-care group, but the difference did not reach significance (Table 5). The subacute rehabilitation-group patients showed a substantial increase over these periods and achieved a significantly higher hazard ratio at 18 months than the subacute primary-care-group patients. There were no differences for the chronic patients.

Net days

At 18 months there were no significant differences between the treatment groups, or between the subacute and chronic patients considered separately (Table 4). Over the three six-month periods, the decrease was significantly more rapid for the whole rehabilitation group and for the subacute rehabilitation-group patients considered separately (bottom of Figure 2a–b). In the first six-month period, there were 50 more *Net days* for the subacute rehabilitation-group patients; in the third period there were 31 days fewer (Figure 2a). There were no differences for the chronic patients (Figure 2b). Adjustment for widespread pain showed no changes.

Visits

At 18 months there were no significant differences between the treatment groups or between the subacute and chronic patients considered separately (Table 4). Over the three six-month periods, the decrease was significantly more rapid for the whole rehabilitation group (bottom of Figure 3a–b). For the subacute patients, the rehabilitation group showed a continuously decreasing

Table 4: Return-to-work share, Net days and Visits.

	Patients	Rehabilitation group	Primary-care group
Return-to-work share (%)	All	35/61 (57 [45 – 70])	35/62 (57 [44 – 69])
	Subacute	18/20 (90 [76 – 104])	15/18 (83 [64 – 102])
	Chronic	17/41 (42 [26 – 57])	20/44 (46 [30 – 61])
Net days	All	397 [354 – 440]	391 [345 – 436]
	Subacute	327 [261 – 392]	292 [194 – 391]
	Chronic	431 [377 – 486]	431 [383 – 478]
Visits	All	55.7 [49.3 – 62.2]	52.0 [38.1 – 66.0]
	Subacute	48.3 [38.5 – 58.1]	40.6 [23.1 – 58.1]
	Chronic	60.1 [51.6 – 68.7]	56.6 [38.1 – 75.2]

Point estimates at 18 months. Descriptive statistics.

Table 5: Return-to-work chance.

Rehabilitation group	Six months	12 months	18 months
All patients (n = 61)	.9 [.6 – 1.4]	1.2 [.7 – 2.0]	1.6 [.7 – 3.6]
Subacute patients (n = 20)	.9 [.5 – 1.6]	1.8 [.8 – 3.9]	3.5 [1.001 – 12.2]
Chronic patients (n = 41)	.9 [.5 – 1.6]	.9 [.4 – 2.1]	1.0 [.3 – 3.9]

Cox regression for recurrent events. Hazard ratios for the rehabilitation group as compared with the primary-care group with 95% confidence intervals. Significant differences in bold figures.

trend while the primary-care group showed a substantial decrease between the first and second six-month periods but no further reduction (Figure 3a). For the chronic patients, the rehabilitation group showed a continuous decrease while the primary-care group showed no reduction (Figure 3b). *Visits* were substantially more numerous for both the subacute and chronic rehabilitation-group patients during the first period, but there were around half as many in the third period. However, there was no significant difference in the rate of decrease between the subacute and chronic patients considered separately (bottom of Figure 3a–b). Adjustment for widespread pain gave no changes.

Interventions

Cognitive-behavioural rehabilitation

Cognitive-behavioural rehabilitation over 18 months included 45.1 [39.2 – 50.9] consultations. Most of the consultations took place in the first six-month period, followed by a rapid reduction (Figure 4). Totalling 0 – 18 months, the most and second most frequent consultations were with a physician (16.6 [14.4 – 18.7]) and a physiotherapist (12.3 [10.5 – 14.1]). A detailed description of the rehabilitation programme is shown in Table 6.

Primary care

For the rehabilitation group, primary care over 18 months comprised 11.7 [6.7 – 16.7] consultations. After a slight increase from the first to the second six-month period, there was stagnation (Figure 4). During the first six-month period most of the rehabilitation-group patients (41/57 (72%)) had no primary-care consultations at all.

For the primary-care group, primary care over 18 months included 50.9 [37.5 – 64.3] consultations. After a slight decrease from the first to the second six-month period there was no further reduction (Figure 5). Totalling 0 – 18 months, the most and second most frequent consultations were with a physiotherapist (28.9 [19.4 – 38.4]) and a physician (12.4 [10.2 – 14.7]).

Other treatment efforts

Hospital care was received by the rehabilitation group and the primary-care group for 1.2 [–.2 – 2.6] days and .8 [–.1 – 1.6] days respectively, surgery for musculoskeletal disorders

by 1/51 (2 [–2 – 6]%) and 3/43 (7 [–1 – 15]%) respectively, and multidisciplinary rehabilitation at other units than the rehabilitation centre by 1/50 (2 [–2 – 6]%) and 4/43 (9 [0 – 18]%) respectively. The differences were non-significant.

Per-protocol analysis

When the incorrectly included rehabilitation-group patient (Figure 1, footnote b) was excluded from the analyses and the rehabilitation-group patient who preferred to continue primary care (Figure 1) was counted with the primary-care group, the *Return-to-work share* increased to 44 [28 – 59]% for the chronic rehabilitation-group patients, and decreased to 44 [30 – 59]% for the chronic primary-care-group patients. This differed only marginally from the intention-to-treat analyses.

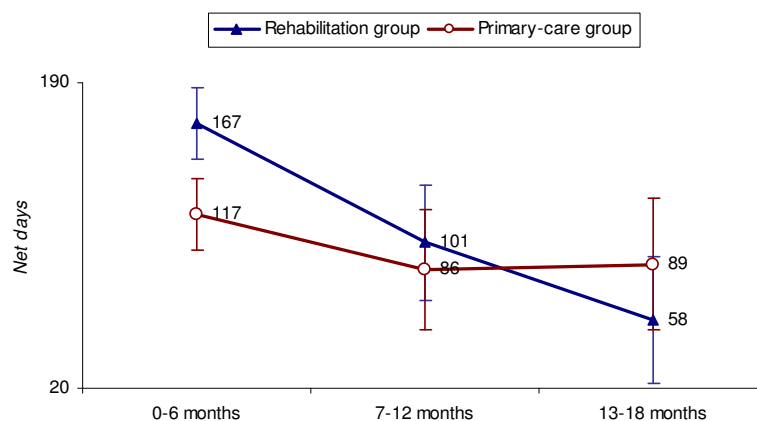
Discussion

This randomized controlled trial concerned primary-care patients with subacute and chronic BNP. A programme of cognitive-behavioural rehabilitation was compared with continued primary care. The results were equivalent over 18 months. However, analyses of the three component six-month periods indicated that the rehabilitation programme might be superior to primary care in the longer run, especially for subacute patients.

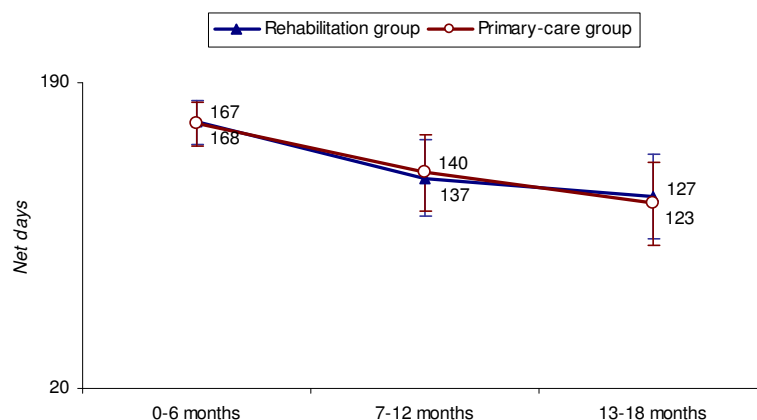
Sick-listing

Why was the *Return-to-work share* substantially lower than expected for the rehabilitation group and higher than expected for the primary-care group? According to Englund et al. [38], sick-listing in Swedish primary care might depend more on the patient's wishes than on guidelines: even when the family doctor did not recommend sick-listing, a certificate was issued in 87% of cases. In view of this, what explains the substantial underestimation of the *Return-to-work share* for the primary-care group (49% vs. the actual share of 57%)? One explanation might be a project that was initiated by the Swedish government in 2002 to halve the extent of sick-listing by 2008 [39]. The focus has been on applying more restrictions in the social insurance system, including failing an increasing number of sick-listing certificates, while the resources for multidisciplinary rehabilitation have been even scarcer than before [40,41]. Anyhow, the low *Return-to-work share* in

a. Subacute patients.



b. Chronic patients.



Time (first, second or third six-month period): $p < .001$.

Rehabilitation group or primary-care group: NS.

Subacute or chronic: $p < .001$.

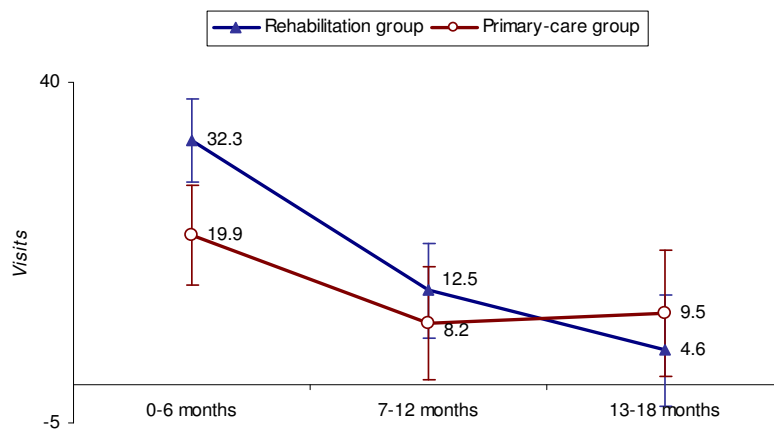
Time x rehabilitation group or primary-care group: $p = .008$.

Time x rehabilitation group or primary-care group x subacute or chronic: $p < .001$.

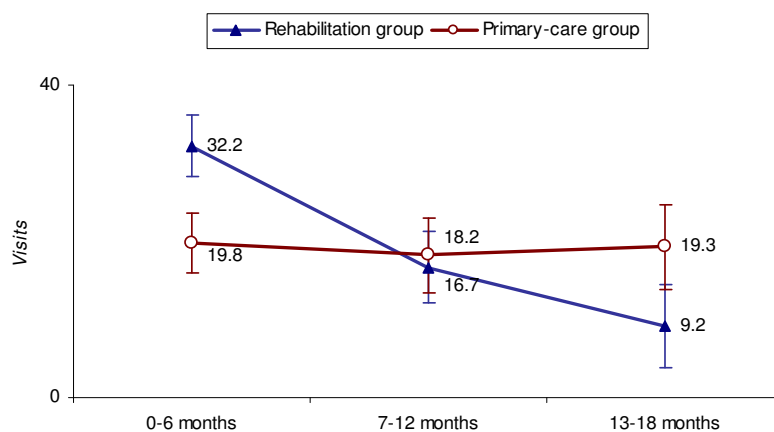
Figure 2

a – b. Net days. Mixed linear model. In the diagrams, 95% confidence intervals are included. At the bottom the explanatory variables and their p -values are shown. Bold figures indicate a significant difference. NS = non-significant.

a. Subacute patients.



b. Chronic patients.



Time (first, second or third six-month period): $p < .001$.

Rehabilitation group or primary-care group: NS.

Subacute or chronic: NS.

Time x rehabilitation group or primary-care group: $p < .001$.

Time x rehabilitation group or primary-care group x subacute or chronic: NS.

Figure 3

a – b. Visits. Mixed linear model. Further explanations in Figure 2a–b.

Table 6: Cognitive-behavioural rehabilitation.

Rehabilitation period (days)	Total period 328 (± 195); median 283 (IQR215)	Investigation and treatment phase 42 (± 18); median 40 (IQR22)	Action phase 287 (± 193); median 249 (IQR232)	
	One-to-one	Treatment measure	Consultations At conferences	In total
Physician	7.3 (± 5.2)	Administration of sick-listing 61/61 (100%) Prescription of drugs 53/61 (87%) Cortisone injections 9/61 (15%)	10.6 (± 6.8)	17.9 (± 11.0)
Physiotherapist	7.8 (± 4.9)	Graded activity 61/61 (100%) Orthopaedic manual therapy 15/61 (25%)	4.6 (± 3.4)	12.4 (± 7.1)
Psychologist or social worker	4.8 (± 5.2)	Cognitive-behavioural therapy 58/61 (95%)	3.4 (± 3.0)	8.2 (± 7.8)
Health-care adviser	6.2 (± 4.8)	Applied relaxation 48/61 (79%)	.3 (± .8)	6.6 (± 5.3)
Conferences: Team conferences Vocational conf. (incl. workpl. visits)	8.6 (± 5.7) 2.4 (± 2.4)	Vocational training 32/61 (52%)		
Sum of treatment occasions	37.1 (± 19.2)		Sum of consultations	45.1 (± 22.8)
Physical activity (days/week):				
Exercise programme	5.5 (± 2.2)			
Gym training	1.0 (± 1.3)			

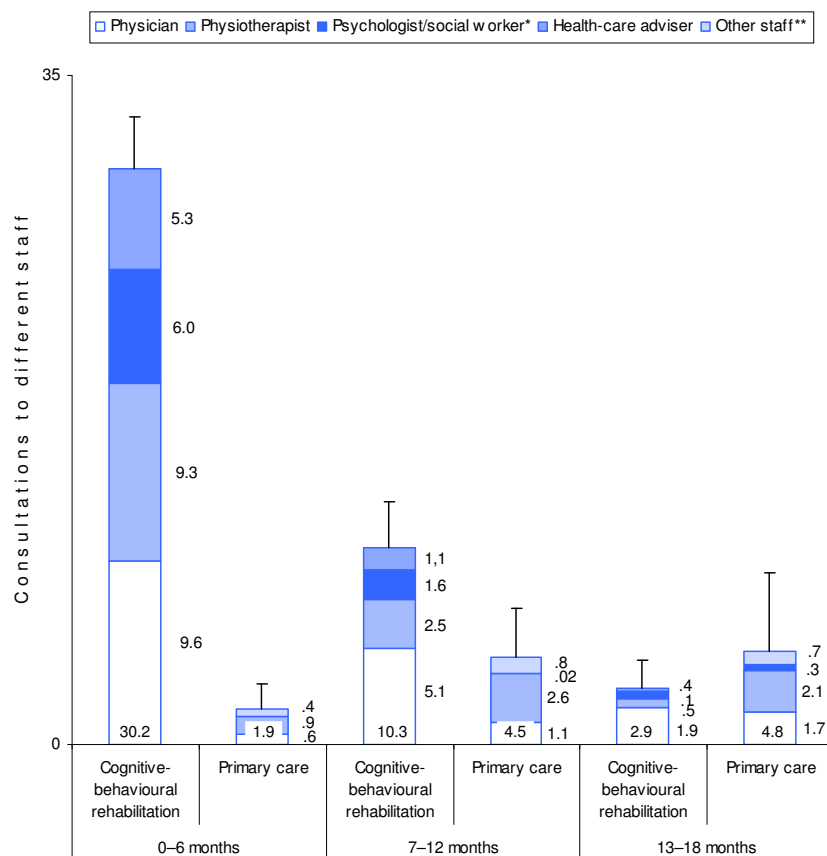
Specification of measures. Number of consultations (mean (SD)) unless otherwise stated.
SD = Standard deviation; IQR = Inter-quartile range.

the rehabilitation group was disappointing, and even if the primary-care group had shown as low a *Return-to-work share* as predicted, the difference between the groups would have remained non-significant.

However, when subacute and chronic patients were analysed separately, a different picture emerged: the *Return-to-work share* for the subacute rehabilitation-group patients was as expected, but the share for the chronic rehabilitation-group patients was far lower. The significantly better *Return-to-work chance* at 18 months and the more rapid reduction in *Net days* among the subacute rehabilitation-group patients highlighted this. Previous research supports the view that cognitive-behavioural interventions at an early stage of disabling BNP can prevent long-term disability [9,10,14,42], while the effect on sick-listing is more doubtful for chronic back pain. Schonstein et al. [43] concluded that physical conditioning programs with a cognitive-behavioural and work-related approach reduced sick-

listing, whereas another Cochrane review revealed that behavioural-rehabilitation programmes had no better effect on sick-listing for chronic back pain than active conservative treatment [20].

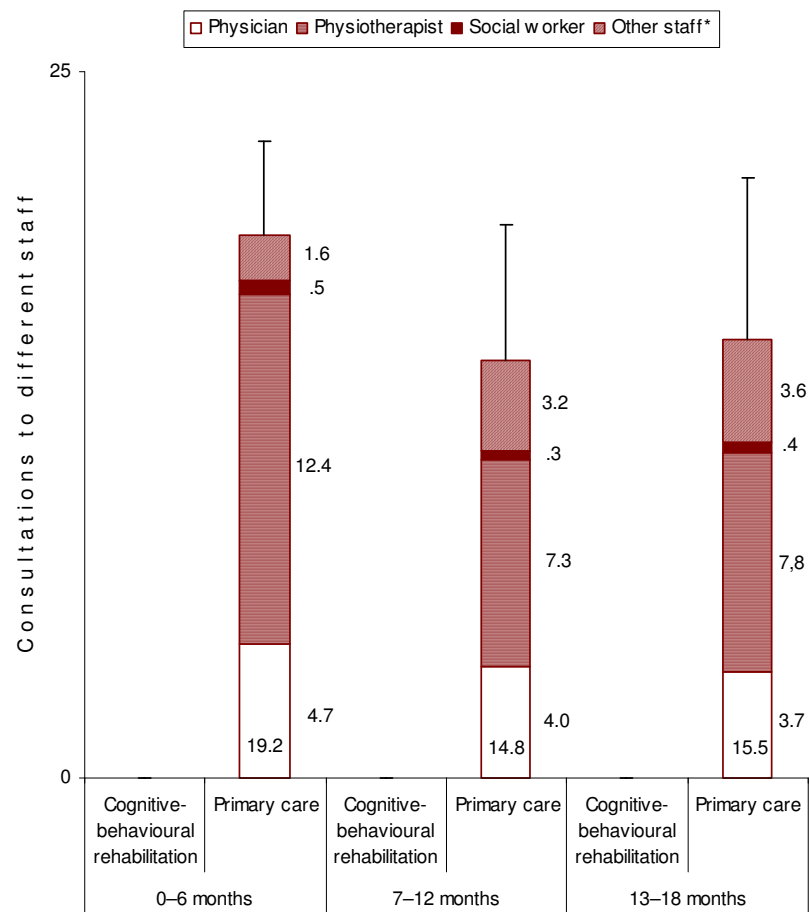
What components of our programme could explain its possible superiority in the long run for subacute patients? Previous research on graded activity had an occupational-care setting and concerned subacute patients only [7,8,14,18,44,45]. Two earlier studies [7,14] found that graded activity in multidisciplinary contexts decreases sick-listing. Two later studies [18,44] contradicted that. Steenstra et al. [18] found that workplace interventions alone reduced sick-listing, while graded activity alone or in combination with workplace interventions did not. One explanation might be that the earlier studies were performed in specialised in-company clinics by a limited number of physiotherapists, including some of the researchers, while the study by Steenstra et al. also



*Concerning primary care, social worker was the only option.

**= Occupational therapist, nurse, X-ray/MRI staff, laboratory personnel, and complementary-medicine staff (for example, masseur and "Chinese doctor").

Figure 4
Consultations to different care staff for the rehabilitation group. For the total number (presented at the bottom of the staples), 95% confidence intervals (upper part) are shown.



*= Occupational therapist, nurse, X-ray/MRI staff, laboratory personnel, and complementary-medicine staff (for example, masseur and “Chinese doctor”)

Figure 5
Consultations to different care staff for the primary-care group. Further explanations in Figure 4.

included out-company clinics with many physiotherapists who had received additional training [18]. These six-month results were confirmed at a 12-month follow-up [8]. Heymans et al. [44] found that standard care plus a low-intensity back school of eight hours was superior to standard care alone, while standard care plus a high-intensity graded-activity-like back school tended to be inferior. The follow-up period of those later studies did not exceed 12 months. In our study, however, the better sick-listing trend for the subacute rehabilitation-group patients was not obvious until after 12 months. Thus, the possibility that a longer period of graded activity has a positive effect on sick-listing for subacute patients in a primary-care setting could not be excluded from those later studies. As to the rest of our specific cognitive-behavioural elements (therapy by a psychologist or a social worker and training in applied relaxation), earlier conclusive studies are lacking [46].

Unlike previous research on graded activity, we also included chronic BNP. Most of the rehabilitation-group patients (43/63 (68%)) had a current sick-listing exceeding 12 weeks at baseline. Our programme did not reduce their sick-listing. Why? One reason could be its comparatively limited extent. Haldorsen et al. [16] showed that, for return-to-work, light multidisciplinary treatment was adequate for moderately-disabled but not for highly-disabled patients. For the latter group, extensive multidisciplinary treatment totalling 120 hours was required; the light programme was no better than standard care. Jensen et al. [19] showed that an extensive behavioural-rehabilitation programme (fully 120 hours) for long-term BNP in female patients reduced sick-listing while more limited efforts did not. Males, however, achieved no better results from the full-time programme than from a light programme or standard care. Quite recently, Staal et al. [45] found that moderately disabled subjects benefited more from graded activity than those with higher disability scores. These studies indicate that return-to-work for patients with chronic BNP, if it is ever possible, requires a more extensive concept than our programme.

Another reason could be methodological defects. Graded activity *by the book* includes: two sessions/week over a maximum of 3–6 months until lasting full-time return-to-work, an early agreement with the patient on a return-to-work date regardless of the actual pain on that particular day, and a hands-off approach [7,18]. As our patients were comparatively more disabled, we found it realistic to apply less frequent sessions to increase the likelihood of positive changes at the next session (there was also a lack of resources for more frequent sessions), no upper time limit (which is also in accordance with the original concept [14]), the possibility of part-time return-to-work, an individual agreement about the return date (early in the

rehabilitation period for some patients, later for others) and, when needed, manual therapy and cortisone injections early in the rehabilitation period (however, the hands-off approach was applied to most (46/61 (75%)) of our patients). Notwithstanding the logical reasons for most of our modifications, they might have contributed to the failure to decrease the sick-listing of the chronic patients. These discrepancies might also explain why the positive effect on the subacute rehabilitation-group patients was not seen until the third six-month period, while those patients had substantially more *Net days* during the first period. It has recently been pointed out that suboptimal rehabilitation items in the pre-phase of return-to-work entail the risk of a counterproductive effect [18].

Health-care visits

In total, the rehabilitation group had more consultations by a physician, which is more costly than other staff categories. However, the resources spent on the rehabilitation group in the first six-month period were balanced by fewer consultations in primary care and a trend towards fewer *Visits* in the long run. Also, although the differences were not significant, the rehabilitation group tended to experience less surgery and other multidisciplinary rehabilitation. For patients with subacute BNP, this agrees with Linton et al. [9], whose cognitive-behavioural interventions were followed by a decrease in health-care utilization. For patients with chronic BNP, our findings are consistent with a large review showing that cognitive-behavioural programs have a substantial positive impact on psychological and medical function but only a small impact on sick-listing [46].

Strengths of the study

The design of our study, a randomized controlled trial, is the gold standard for evaluating treatment methods for back and neck pain [2].

The sick-listing data were complete. We also consider the health-care data to be acceptably representative. The response rate was higher than 80% except at 12 months, when it was nearly 80% for the primary-care group. Even when the missing data for the two deceased patients were included, the rehabilitation group met drop-out criteria [34]. For the primary-care group, *Visits* over 18 months should be interpreted with some caution as 32% were non-responders, but in other respects the follow-up rate of the primary-care group was also satisfactory. The non-responders in the rehabilitation group had characteristics that may have increased health-care use (longer sick-listing periods and higher unemployment). In the primary-care group the non-responders were younger, which could have decreased utilization, whereas the lower health-related quality of life could possibly increase utilization.

However, for the great majority, there were no significant differences at baseline between the non-responders and responders.

Limitations of the study

The inclusion plan was not fulfilled. A possible consequence may have been that some differences between the groups could not be demonstrated. However, certain differences in favour of the rehabilitation group were clear with the number of patients actually included.

Comparison of health-care visits gives only a limited idea of cost effectiveness. A complete health-economic evaluation is planned in a future study, including a cost-benefit analysis in which the direct costs (mainly of the interventions themselves), the indirect costs (mainly of the sick-listing), and the health-related quality of life are compared [47].

The primary outcome measure showed no difference. Notwithstanding the positive trends in favour of the rehabilitation group, especially for the subacute patients, *Net days* and *Visits* were also equivalent over 18 months. As differences in the results of various interventions tend to even out after 12 – 18 months [19], more conclusive results might require a longer follow-up period than in this study.

Conclusion

For patients with subacute and chronic BNP, cognitive-behavioural rehabilitation was compared with primary care. The results were equivalent over 18 months. However, there were indications that cognitive-behavioural rehabilitation in the longer run might be superior. For subacute BNP, it might be superior in terms of both sick-listing and health-care visits; for chronic BNP, in terms of health-care visits only. More conclusive results concerning this possible long-term effect might require a longer follow-up.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

OL was the main investigator, carried out the study, performed the analysis, and drafted the manuscript. SEJ contributed to the statistical analysis. LES, as supervisor for OL, participated in all phases of the study. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by grants from the Stockholm County Social Insurance Agency, Stockholm County Council, Ministry of Health and Social Affairs, Vårdal Foundation, Cardionics and Pharmacia (now part of Pfizer).

References

- Hansson T, Jensen I: **Sickness absence due to back and neck disorders. Sickness absence – causes, consequences, and physicians' sickness-certification practice.** *Scand J Public Health* 2004, **32**(Suppl 63):109-151.
- Nachemson A: **Introduction.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment* Edited by: Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:1-12.
- Norlund A, Waddell G: **Cost of back pain in some OECD Countries.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment* Edited by: Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:421-425.
- Waddell G, Watson PJ: **Rehabilitation.** In *Back Pain Revolution* 2nd edition. Edited by: Waddell G. Edinburgh: Churchill Livingstone; 2004:371-399.
- Wasiak R, Young AE, Roessler RT, McPherson KM, van Poppel MN, Anema JR: **Measuring return to work.** *J Occup Rehabil* 2007, **17**:766-781.
- Hazard RG, Fenwick JW, Kalisch SM, Redmond J, Reeves V, Reid S, Frymoyer JW: **Functional restoration with behavioural support. A one-year prospective study of patients with chronic low-back pain.** *Spine* 1989, **14**:157-161.
- Hlobil H, Staal JB, Twisk J, Koke A, Ariens G, Smid T, van Mechelen W: **The effects of a graded activity intervention for low back pain in occupational health on sick leave, functional status and pain: 12-month results of a randomized controlled trial.** *J Occup Rehabil* 2005, **15**:569-580.
- Anema JR, Steenstra IA, Bongers PM, de Vet HC, Knol DL, Loisel P, van Mechelen W: **Multidisciplinary rehabilitation for subacute low back pain: graded activity or workplace intervention or both? A randomized controlled trial.** *Spine* 2007, **32**:291-298.
- Linton SJ, Nordin E: **A 5-year follow-up evaluation of the health and economic consequences of an early cognitive behavioral intervention for back pain: a randomized, controlled trial.** *Spine* 2006, **31**:853-858.
- Linton SJ, Ryberg M: **A cognitive-behavioral group intervention as prevention for persistent neck and back pain in a non-patient population: a randomized controlled trial.** *Pain* 2001, **90**:83-90.
- Waddell G: **The biopsychosocial model.** In *Back Pain Revolution* 2nd edition. Edited by: Waddell G. Edinburgh: Churchill Livingstone; 2004:265-282.
- Waddell G, van Tulder M: **Clinical guidelines.** In *Back Pain Revolution* 2nd edition. Edited by: Waddell G. Edinburgh: Churchill Livingstone; 2004:283-322.
- Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, Koes B: **Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults.** *Cochrane Database Syst Rev* 2003:CD002194.
- Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson A: **The effect of graded activity on patients with sub-acute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach.** *Phys Ther* 1992, **72**:279-293.
- Loisel P, Lemaire J, Poitras S, Durand MJ, Champagne F, Stock S, Diallo B, Tremblay C: **Cost-benefit and cost-effectiveness analysis of a disability prevention model for back pain management: a six year follow up study.** *Occup Environ Med* 2002, **59**:807-815.
- Haldorsen EM, Grasdahl AL, Skouen JS, Risa AE, Kronholm K, Ursin H: **Is there a right treatment for a particular patient group? Comparison of ordinary treatment, light multidisciplinary treatment, and extensive multidisciplinary treatment for long-term sick-listed employees with musculoskeletal pain.** *Pain* 2002, **95**:49-63.
- Indahl A, Haldorsen EH, Holm S, Reikeras O, Ursin H: **Five-year follow-up study of a controlled clinical trial using light mobilization and an informative approach to low back pain.** *Spine* 1998, **23**:2625-2630.
- Steenstra IA, Anema JR, Bongers PM, de Vet HC, Knol DL, van Mechelen W: **The effectiveness of graded activity for low back pain in occupational healthcare.** *Occup Environ Med* 2006, **63**:718-725.
- Jensen IB, Bergstrom G, Ljungquist T, Bodin L: **A 3-year follow-up of a multidisciplinary rehabilitation programme for back and neck pain.** *Pain* 2005, **115**:273-283.

20. Ostelo RW, van Tulder MW, Vlaeyen JW, Linton SJ, Morley SJ, Assendelft WJ: **Behavioural treatment for chronic low-back pain.** *Cochrane Database Syst Rev* 2005.
21. Turk DC, Rudy TE: **Neglected topics in the treatment of chronic pain patients: Relapse, noncompliance, and adherence enhancement.** *Pain* 1991, 44:5-28.
22. Marklund S: **Risk- och friskfaktorer. Sjukskrivning och rehabilitering i Sverige.** In [In Swedish; English summary: *Sickness Absentism and Vocational Rehabilitation in Sweden*] RFV redovisar; 1997:6.
23. Jensen IB, Bodin L, Ljungquist T, Bergström G, Nygren Å: **Assessing the needs of patients in pain: a matter of opinion?** *Spine* 2000, 25:2816-2823.
24. **General practitioners** *Health and Ageing Factbook 2006. Chapter 7 – International comparisons* [http://www.health.gov.au/internet/main/publishing.nsf/Content/056CE2932CA4D491CA257183008169E2/\$File/2006%20Health%20and%20Ageing%20Factbook.pdf]. Australian Government. Department of health and ageing
25. Grunnesjö M, Borgefeldt JP, Svärdsudd KF, Blomberg SI: **A randomized controlled clinical trial of stay-active care versus manual therapy in addition to stay-active care: functional variables and pain.** *J Manipulative Physiol Ther* 2004, 27:431-441.
26. **The Health and Medical Service Act (1982:763)** 2003 [http://sweden.gov.se/sb/d/574/a/23125jsessionid=afUj3zB3N6v-]. Government Offices of Sweden: Ministry of Health and Social Affairs
27. Rae D: **Getting better value for money from Sweden's health-care system.** *Economics department working paper no. 443.* OECD 2005 [http://www.oilis.oecd.org/olis/2005doc.nsf/c707a7b4806fa95dc125685d005300b6/f501081ec882a6b8c1257088002cddb0/\$FILE/JT00189812.PDF].
28. Arrelöv B, Borgquist L, Ljungberg D, Svärdsudd K: **The influence of change of legislation concerning sickness absence on physicians' performance as certifiers. A population-based study.** *Health Policy* 2003, 63:259-268.
29. Altman DG: **Comparing groups – continuous data.** In *Practical statistics for medical research* 1st edition. London: Chapman & Hall; 1991:179-228.
30. Altman DG: **Comparing groups – categorical data.** In *Practical statistics for medical research* 1st edition. London: Chapman & Hall; 1991:229-276.
31. Kelly PJ, Lim LL: **Survival analysis for recurrent event data: an application to childhood infectious diseases.** *Statist Med* 2000, 19:13-33.
32. Brown H, Prescott R: **Applied mixed models in medicine.** Edinburgh: John Wiley & Sons LTD; 2001:199-259.
33. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF: **Improving the quality of reporting of randomized controlled trials. The CONSORT statement.** *JAMA* 1996, 276:637-639.
34. van Tulder MW, Assendelft WJ, Willem JJ, Koes BW, Bouter LM: **Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders.** *Spine* 1997, 22:2323-2330.
35. Waddell G: **Low back pain: a twentieth century health care enigma.** *Spine* 1996, 21:2820-2825.
36. Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ: **Progressive isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population.** *Spine* 1988, 13:998-1002.
37. Lindell O, Eriksson L, Strender LE: **The reliability of a 10-test package for patients with prolonged back and neck pain: Could an examiner without formal medical education be used without loss of quality? A methodological study.** *BMC Musculoskel Disord* 2007, 8:31.
38. Englund L, Svärdsudd K: **Sick-listing habits among general practitioners in a Swedish county.** *Scand J Prim Health Care* 2000, 18:81-86.
39. **Goals of public health** Swedish; *Mål för folkhälsan Governmental bill 2002/03:35* [http://fhi.se/upload/PDF/2004/omfhi/p200203_35.pdf].
40. Lidwall U: **The Swedish Social Insurance Agency and vocational rehabilitation – active vocational rehabilitation measures and return to work.** [In Swedish; summary in English] 2006 [http://forsakringskassan.se/filer/publikationer/pdf/ana0610.pdf]. Stockholm: Social Insurance Agency
41. Bergendorff S: **Rehabilitation – the history of a long period of suffering.** [In Swedish; *Rehabilitering – ett långt lidandes historia*]. Commission of Inquiry on Social Insurance [In Swedish; *Socialförsäkringsutredningen*]. *Samtal om socialförsäkring* 2006:10 [http://www.kfs.net/projekt/dokument/SOU_lidande_rehab.pdf].
42. Linton SJ, Boersma K, Jansson M, Svard L, Botvalde M: **The effects of cognitive-behavioral and physical therapy preventive interventions on pain-related sick leave: a randomized controlled trial.** *Clin J Pain* 2005, 21:109-119.
43. Schonstein E, Kenny DT, Keating J, Koes BW: **Work conditioning, work hardening and functional restoration for workers with back and neck pain.** *Cochrane Database Syst Rev* 2003.
44. Heymans MW, de Vet HC, Bongers PM, Knol DL, Koes BW, van Mechelen W: **The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial.** *Spine* 2006, 31:1075-1082.
45. Staal JB, Hlobil H, Köke AJ, Twisk JW, Smid T, van Mechelen W: **Graded activity for workers with low back pain: who benefits most and how does it work?** *Arthritis Rheum* 2008, 59:642-649.
46. Linton SJ: **Utility of cognitive-behavioral psychological treatments.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment* Edited by: Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:361-381.
47. Goossens M, Evers S: **Cost-effectiveness of treatment for neck and low back pain.** In *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment* Edited by: Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:399-419.
48. Wewers ME, Lowe NK: **A critical review of visual analogue scales in the measurement of clinical phenomena.** *Res Nurs Health* 1990, 13:227-236.
49. Hansson E, Hansson T, Jonsson R: **Predictors for work ability and disability in men and women with low-back or neck problems.** *Eur Spine J* 2006, 15:780-793.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2474/9/172/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



IV

RESEARCH ARTICLE

Open Access

Predictors of stable return-to-work in non-acute, non-specific spinal pain: low total prior sick-listing, high self prediction and young age. A two-year prospective cohort study

Odd Lindell*, Sven-Erik Johansson, Lars-Erik Strender

Abstract

Background: Non-specific spinal pain (NSP), comprising back and/or neck pain, is one of the leading disorders in long-term sick-listing. During 2000-2004, 125 Swedish primary-care patients with non-acute NSP, full-time sick-listed 6 weeks-2 years, were included in a randomized controlled trial to compare a cognitive-behavioural programme with traditional primary care. This prospective cohort study is a re-assessment of the data from the randomized trial with the 2 treatment groups considered as a single cohort. The aim was to investigate which baseline variables predict a stable return-to-work during a 2-year period after baseline: objective variables from function tests, socioeconomic, subjective and/or treatment variables. Stable return-to-work was a return-to-work lasting for at least 1 month from the start of follow-up.

Methods: *Stable return-to-work* was the outcome variable, the above-mentioned factors were the predictive variables in multiple-logistic regression models, one per follow-up at 6, 12, 18 and 24 months after baseline. The factors from univariate analyses with a *p*-value of at most .10 were included. The non-significant variables were excluded stepwise to yield models comprising only significant factors (*p* < .05). As the comparatively few cases made it risky to associate certain predictors with certain time-points, we finally considered the predictors which were represented in at least 3 follow-ups. They are presented with odds ratios (OR) and 95% confidence intervals.

Results: Three variables qualified, all of them represented in 3 follow-ups: *Low total prior sick-listing* (including all diagnoses) was the strongest predictor in 2 follow-ups, 18 and 24 months, OR 4.8 [1.9-12.3] and 3.8 [1.6-8.7] respectively, *High self prediction* (the patients' own belief in return-to-work) was the strongest at 12 months, OR 5.2 [1.5-17.5] and *Young age* (max 44 years) the second strongest at 18 months, OR 3.5 [1.3-9.1].

Conclusions: In primary-care patients with non-acute NSP, the strong predictors of stable return-to-work were 2 socioeconomic variables, *Low total prior sick-listing* and *Young age*, and 1 subjective variable, *High self-prediction*. Objective variables from function tests and treatment variables were non-predictors. Except for *Young age*, the predictors have previously been insufficiently studied, and so our study should widen knowledge within clinical practice.

Trial registration: Trial registration number for the original trial NCT00488735.

Background

For many years, spinal pain, comprising back and/or neck pain, was the leading disorder in long-term sick-listing, including disability pensions, in Sweden as all over the industrial world. In 2002, Sweden was the leading country

within the European Union in sick-listing for spinal pain [1], which in 2007 resulted in 11.9% of new disability pensions [2]. Following an international trend [3], the leading position of spinal pain in Sweden since 2005 has been overtaken by depression (in 2007 13.1% of new disability pensions) [2]. Most cases of spinal pain concern non-specific spinal pain (NSP) and are a matter for primary care [4]. In the management of disabling spinal pain, stable

* Correspondence: odd.lindell@ki.se

Center for Family and Community Medicine, Karolinska Institutet, Alfred Nobels allé 12, SE-141 83 Huddinge, Sweden

return-to-work is the ultimate objective [4]. As return-to-work is often followed by recurrences of work absence, longitudinal data are required to denote a stable return-to-work [5].

Cost-effectiveness in allocating treatment resources requires predictors of return-to-work to be collected by means of both questionnaires and function tests, i.e. tests in which the patient performs some kind of physical activity [6]. While the former are cheap, the latter require substantial personnel resources. Despite an immense amount of research, no gold standard for questionnaires and/or tests has been established for this purpose [6,7]. In the treatment of non-acute NSP, i.e. pain leading to full-time sick-listing for more than 3 weeks [8], evidence-based guidelines advocate a cognitive-behavioural therapeutic approach [4].

During 2000-2004, 125 patients with non-acute NSP were included in a randomized, controlled trial to compare a cognitive-behavioural programme with traditional primary care [9]. A package of function tests and a questionnaire were completed at baseline. The aim of this study was to answer the question "which are the predictors at baseline in non-acute NSP for stable return-to-work during a 2-year period after baseline: objective variables from function tests, socio-economic, subjective and/or treatment variables?"

Methods

On sick-listing and return-to-work in Sweden

As the employer has the financial responsibility for the 2 initial weeks of sick-listing in Sweden, the available data include only the sick-listing periods exceeding 2 weeks. For the unemployed subjects, however, those data include all periods. Sick-listing, as described in detail in a prior study [9], might have the degrees .25, .50, .75 or 1.00 (= full-time). The degree of return-to-work = 1.00 minus the degree of sick-listing, as defined by the Social Insurance Agency. For example, sick-listing = .75 equals return-to-work = .25 and full-time sick-listing equals non-return-to-work. In response to prolonged sick-listing, the Agency might consider a temporary or permanent disability pension (the temporary form being abolished in 2008), which might have the same degrees as the other forms of sick-listing.

Setting and source population

The setting was a suburban area in the Southern part of Stockholm County, including 9 municipalities with a population of 466,000, of whom 288,000 of working age (18-64 years) constituted the source population.

Patients

One hundred and twenty-five primary-care patients with non-acute NSP were recruited to a randomized controlled

trial, which in detail was described in a previous study [9], by 41 family doctors at 13 health centres between August 2000 and January 2004. Recruitment was non-systematic, i.e. it was up to the family doctor on the basis of her or his current motivation and available time to invite a potentially eligible patient. In summary:

The patients were allocated either to a cognitive-behavioural programme at a rehabilitation centre or continued traditional primary care. *The criteria for inclusion:* 1. Vocationally active, up to and including 59 years of age. 2. Sick-listed full-time for spinal pain at least six weeks (42 days) and at most two years (730 days). 3. Able to fill in forms. *The criteria for exclusion:* 1. Temporary disability pension or disability pension being paid or in preparation. 2. Primary need for action by a hospital specialist (for example, operation for intervertebral herniation (slipped disc)). 3. Pregnancy or diseases (other than spinal pain) that might make the rehabilitation programme impracticable (for example, advanced pulmonary disease). 4. Whiplash-associated disorders as a primary obstacle to working. 5. Previous rehabilitation at the rehabilitation centre. 6. Other multidisciplinary rehabilitation measures ongoing or planned.

The recruited patients were interviewed by telephone by a research assistant within 2 days. The patients who remained qualified saw the assistant at the health centre within 5 days. Before the assistant carried out the randomization, certain procedures were completed: the patient finished a questionnaire, including a pain drawing; the assistant categorized the pain as being back and/or neck pain, basing the decision on how the patient completed the pain drawing and by a short interview. The back was taken as the area below an imaginary line connecting the lower tips of the shoulder blades, including the lower half of the thoracic spine and the lumbar spine; and the neck was the area on and above this line, including the upper half of the thoracic spine and the cervical spine [10]; the patient also performed a package of 10 function tests as described in detail in a previous study [11].

Design

This prospective cohort study is a re-assessment of the data from the randomized controlled trial with the 2 treatment groups considered as a single cohort.

Outcome variable

Stable return-to-work The outcome variable was *Stable return-to-work*, which required that a return-to-work on a specific day lasted for at least 1 month. For example, a *Stable return-to-work* on 6 June required that the return-to-work continued at least up to and including 5 July. The reference to *Stable return-to-work* was *Non-return-to-work*, including non-return-to-work a specific day and return-to-work that day but with recurrence of

work absence the following month. Due to the responsibility of the employer, *Stable return-to-work* possibly contained a period of work absence of a maximum of 14 days during the follow-up month including the specific day. *Stable return-to-work* was analyzed in 4 specific days during a 2-year period, selected as 6, 12, 18 and 24 months after baseline.

Predictive variables

Objective variables Six reliable function tests from the 10-test package were used as objective variables. In a previous study, we had examined the reliability, including inter- and intra-rater reliability, of the package [11]. In summary, 2 examiners participated, an experienced physiotherapist and a research assistant. All the 5 tests that did not require manual fixation of the patient by the examiner were reliable. Only 1 of the 5 tests which required fixation was reliable. In conclusion, 6 of the

10 tests were reliable and could be used by an examiner lacking formal medical education (the research assistant) without loss of quality. Two of those tests included flexion to the right and to the left and rotation to the right and to the left, and a lift test comprised a lumbar and a cervical subtest. Nine subtests in total are given in Table 1.

Socioeconomic variables

These were collected from the questionnaire except data for the 2 sick-listing variables, which were collected from The Social Insurance Office. The sick-listing variables were: *Subacute NSP* = current, full-time sick-listing at baseline for NSP of 6-12 weeks (42-84 days) with the reference *Chronic NSP* = current, full-time sick-listing of more than 12 weeks up to and including 2 years (85-730 days) [8], and *Low total prior sick-listing* = at most 183 net days during the 2 years prior to baseline, including all diagnoses, with the reference *High total*

Table 1 Objective variables. Results of univariate-logistic regression, adjusted for gender and age, with p-values of at most .10

Subtests	Class limits	n	Prediction for <i>Stable return-to-work</i>											
			6 months			12 months			18 months			24 months		
			OR	p	95%CI	OR	p	95%CI	OR	p	95%CI	OR	p	95%CI
<i>Forward flexion</i> (centimeters (cm))	25-64	41							Ref.			Ref.		
	8-24	42	-	-	-	-	-	-	3.4	.01	1.3-8.8	2.6	.05	1.0-6.5
	0-7	41	-	-	-	-	-	-	2.1	NS	.8-5.6	1.3	NS	.5-3.2
<i>Modified Schober</i> (cm)	1-3	18												
	4-6	83	-	-	-	-	-	-	-	-	-	-	-	-
	7-19	23	-	-	-	-	-	-	-	-	-	-	-	-
<i>Lateral flexion right</i> (cm)	3-10	41							Ref.					
	11-15	39	-	-	-	-	-	-	2.3	.09	.9-6.2	-	-	-
	16-28	44	-	-	-	-	-	-	1.9	NS	.8-4.9	-	-	-
<i>Lateral flexion left</i> (cm)	2-11	41							Ref.					
	12-15	38	-	-	-	-	-	-	2.9	.03	1.1-7.6	-	-	-
	16-27	45	-	-	-	-	-	-	1.8	NS	.7-4.7	-	-	-
<i>Cervical rotation right</i> (degrees)	0-50	44										Ref.		
	51-60	43	-	-	-	-	-	-	-	-	-	2.6	.04	1.0-6.6
	61-80	37	-	-	-	-	-	-	-	-	-	2.7	.05	1.0-7.1
<i>Cervical rotation left</i> (degrees)	0-50	47												
	51-60	39	-	-	-	-	-	-	-	-	-	-	-	-
	61-80	38	-	-	-	-	-	-	-	-	-	-	-	-
<i>Abdominal endurance</i> (seconds)	0	46												
	1-14	40	-	-	-	-	-	-	-	-	-	-	-	-
	15-75	38	-	-	-	-	-	-	-	-	-	-	-	-
<i>PILE lumbar</i> (kilogram)	0-6	33												
	8-12	45	-	-	-	-	-	-	-	-	-	-	-	-
	14-44	46	-	-	-	-	-	-	-	-	-	-	-	-
<i>PILE cervical</i> (kilogram)	0-6	37	Ref.						Ref.					
	8-12	47	1.4	NS	.5-4.4	-	-	-	1.1	NS	.4-2.9	-	-	-
	14-44	40	2.9	.09	.9-9.5	-	-	-	2.8	.06	1.0-8.4	-	-	-

OR = Odds ratio. 95%CI = 95% confidence interval. Ref. = Reference, which always has OR = 1.0. NS = Non-significant ($p > .10$).
The rationales for the choice of the function tests were established in a previous study [11].

prior sick-listing ≥ 184 net days [12]. 'Net days' is the sick-listing expressed in whole days = crude days \times degree [13]. In total, 23 socioeconomic variables are presented in Table 2.

Subjective variables

Except for the division into back and/or neck pain, the subjective variables were collected exclusively from the

questionnaire. They included different aspects of pain, mental mood, comorbidity, loss of function due to NSP, health-related-quality of life, coping with pain, and a question about the probability of return-to-work: "What do you believe, honestly, is the probability that you will become so much better that you will be able to work at some time in the future?" [14]. *High self prediction*

Table 2 Socioeconomic variables

	n	Prediction for <i>Stable return-to-work</i>											
		6 months			12 months			18 months			24 months		
		OR	p	95%CI	OR	p	95%CI	OR	p	95%CI	OR	p	95%CI
<i>Man</i> [24,27,29,34]	56	-	-	-	2.1	.06	1.0-4.5	-	-	-	-	-	-
<i>Young age</i> (≤ 44 years)[26,30]	67	-	-	-	3.0	.005	1.4-6.7	2.9	.006	1.3-6.1	2.6	.01	1.2-5.4
<i>Non-immigrant</i> [12] ¹	90	-	-	-	-	-	-	-	-	-	-	-	-
<i>Co-habiting</i> [53] ²	85	-	-	-	-	-	-	-	-	-	-	-	-
<i>Living without children</i> [54]	55	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-bad economy</i> [17] ³	68	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-low education</i> [55] ⁴	80	2.2	.07	.9- 5.6	2.9	.02	1.2-7.1	3.0	.01	1.3-6.9	3.5	.004	1.5-8.0
<i>White-collar job</i> [56] ⁵	12 ⁵	-	-	-	-	-	-	-	-	-	-	-	-
Physical work conditions ⁶ :													
<i>No vibrations</i> [58]	84	3.3	.03	1.2-9.4	2.9	.04	1.0-7.0	-	-	-	-	-	-
<i>Light physical workload</i> [34]	21	-	-	-	-	-	-	-	-	-	-	-	-
<i>Varied work moments</i> [34]	46	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-sedentary work</i> [57]	88	-	-	-	-	-	-	-	-	-	-	-	-
<i>Comfortable w. postures</i> [34]	27	-	-	-	-	-	-	-	-	-	-	-	-
Psychosocial work conditions ⁷ :													
<i>No job strain</i> [60]	90	-	-	-	-	-	-	-	-	-	-	-	-
<i>Good social support</i> [61]	94	4.5	.02	1.2-16.2	-	-	-	-	-	-	2.7	.04	1.1-6.8
<i>Non-unemployed</i> [62]	95	.5	.10	.2-1.2	-	-	-	-	-	-	-	-	-
<i>No work trauma litigation</i> ⁸	97 ⁸	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-smoking</i> [26]	75	-	-	-	-	-	-	-	-	-	-	-	-
<i>No indication of alcohol overconsumption</i> [42] ⁹	107	-	-	-	-	-	-	-	-	-	-	-	-
<i>High physical activity</i> [65] ¹⁰	86	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-obese (BMI < 30[66])[37]</i>	94	.4	.05	.2-1.0	-	-	-	-	-	-	-	-	-
<i>Subacute NSP</i> ¹¹	38	3.4	.006	1.4-8.0	2.8	.02	1.2-6.3	5.4	< .001	2.2-13.0	3.1	.008	1.4-7.2
<i>Low total prior sick-listing</i> ¹²	57	3.1	.005	1.4-6.9	3.1	.005	1.4-6.9	7.7	< .001	3.3-18.1	4.9	< .001	2.2-11.0

¹ = Born in Sweden. Reference: *Immigrant* (n = 34).

² Includes living single with children. Reference: *Single* = living alone, without children (n = 39).

³ = "Neither bad nor good", "Good" or "Very good". Reference: *Bad economy* = "Very bad" or "Bad" (n = 56).

⁴ Reference: *Low education* = at most junior high school (n = 44).

⁵ Out of the 94 non-unemployed patients. Reference: *Blue-collar job* (n = 82).

⁶ "State the conditions that you regularly (not occasionally) are exposed to: ...Vibrations (from tools, vehicles etc.) ...Heavy lifting or greater muscle efforts ... Monotonous work moments ...Sedentary work ...Difficult work postures (bent, twisted, locked etc.): answer "No". References: "Yes" [57].

⁷ Psychological demands (5 items), decision latitude (6 items) and social support (6 items), total scores 5-20, 6-24 and 6-24 respectively; *No job strain* = non-scoring demands above the midpoint (> 13) and decision latitude below the midpoint (> 15); reference: *Job strain* = demands above + decision latitude below the midpoint (n = 34). *Good social support* = above the midpoint (> 15); reference: *Bad social support* = below the midpoint (n = 30) [59].

⁸ Out of 115 patients (9 patients scored "I don't know"). "Have you reported your pain as a work trauma?"; answer "No". Reference: "Yes" (n = 18) [63].

⁹ To drink alcohol corresponding to at least 1/2 bottle (= 37.5 centilitres) of strong spirits on one and the same occasion, less than 2-3 times monthly. Reference: *Indication of alcohol overconsumption* = at least 2-3 times monthly (n = 17) [64](+ personal communication Anders Romelsjö 27 Aug. 2007).

¹⁰ Physical activity, including walking > 30 minutes, twice/week or more. Reference: *Low physical activity*: once/week or less (n = 38).

¹¹ = a current, full-time sick-listing at baseline for NSP 42-84 days. Reference: *Chronic NSP* = a corresponding sick-listing of 85-730 days (n = 84) [31].

¹² = a prior 2-year sick-listing for all diagnoses of at most 183 net days. Reference: *High total prior sick-listing* ≥ 184 net days [12].

Further explanations in Table 1.

included the answering alternatives 'rather probable', 'probable' and 'very probable', and *Low self prediction* the alternatives 'rather improbable', 'improbable' and 'very improbable'. A similar type of question was used by Linton et al. [15], but included a future time-limit of 6 months, i.e. a much shorter period than our 2-year follow-up. We therefore chose the open-ended question from Eklund et al. [14]. A total of 16 subjective variables are shown in Table 3.

Treatment variables

Sixty-three of the 125 patients received *Cognitive-behavioural rehabilitation* and 62 patients received the reference treatment of *Traditional primary care*. The treatment options were described in detail in a previous study [9].

Statistics

STATA10.1 was used for the calculations [16].

Table 3 Subjective variables

	Class limits	n	Prediction for <i>Stable return-to-work</i>											
			6 months			12 months			18 months			24 months		
			OR	p	95%CI	OR	p	95%CI	OR	p	95%CI	OR	p	95%CI
<i>Pain just now</i> (VAS 1-100)[25]	70-100	41	Ref.											
	48-69	43	2.4	.09	9-6.9	-	-	-	-	-	-	-	-	-
	0-47	40	1.5	NS	.5-4.3	-	-	-	-	-	-	-	-	-
<i>Pain at worst last week</i> [25]	81-100	42	Ref.											
	68-80	43	2.5	.09	9-6.8	-	-	-	-	-	-	-	-	-
	0-67	39	1.4	NS	.5-4.2	-	-	-	-	-	-	-	-	-
<i>Intermittent pain</i> [15] ¹	-	39	-	-	-	-	-	-	2.3	.04	1.0-5.4	-	-	-
<i>Non-radiating pain</i> [17] ²	-	32	-	-	-	-	-	-	-	-	-	-	-	-
<i>Local pain</i> [25] ³	-	24	-	-	-	-	-	-	-	-	-	-	-	-
<i>Back-pain domination</i> [32] ⁴	-	86	9.0	.004	2.0-40.2	2.5	.05	1.0-6.4	-	-	-	-	-	-
<i>Time since start of NSP</i> (years)[27]	> 5	53							Ref.			Ref.		
	1.5-5	34	-	-	-	-	-	-	2.9	.03	1.1-7.4	2.2	.09	9-5.5
	< 1.5	37	-	-	-	-	-	-	1.5	NS	.6-3.6	1.1	NS	.5-2.7
<i>No surgery f. b/n pain</i> [50] ⁵	-	116	-	-	-	-	-	-	-	-	-	-	-	-
<i>No anxiety/depression</i> [15] ⁶	-	26	-	-	-	-	-	-	-	-	-	-	-	-
<i>Tired seldom</i> [67] ⁷	-	59	3.1	.01	1.3-7.6	-	-	-	1.9	.09	.9-4.2	-	-	-
<i>No comorbidity</i> [68] ⁸	-	79	-	-	-	-	-	-	-	-	-	-	-	-
<i>Non-severe functional impairment</i> (ODI) ⁹	-	78	2.1	.09	9-4.9	2.9	.01	1.3-6.8	2.5	.02	1.2-5.4	-	-	-
<i>Health-related quality of life</i>	0-359	42	Ref.			Ref.			Ref.					
	.360-629	46	2.8	.06	1.0-8.3	2.1	NS	.8-5.4	2.1	NS	.8-5.1	-	-	-
	.630-1.0	36	2.9	.06	.9-8.9	2.6	.06	1.0-7.1	3.0	.03	1.1-7.9	-	-	-
<i>State of health</i> (EQ-VAS)[21]	0-35	44	Ref.						Ref.					
	36-49	33	2.2	NS	.7-7.0	-	-	-	2.0	NS	.7-5.4	-	-	-
	50-100	47	3.6	.02	1.3-10.3	-	-	-	3.1	.01	1.3-7.7	-	-	-
<i>Non-catastrophizing</i> [70] ¹⁰	-	67	2.2	.08	.9-5.1	-	-	-	3.6	.002	1.6-8.0	2.3	.04	1.1-4.9
<i>High self prediction</i> [14]	-	95	4.2	.03	1.2-15.2	6.4	.002	1.9-21.0	4.4	.005	1.5-12.4	3.8	.008	1.4-10.2

¹ Reference: *Continual pain* = pain whenever awake (n = 95).

² Reference: *Radiating pain* = radiation of pain/numbness to the leg beneath the knee and/or the arm beneath the elbow (n = 92).

³ Pain in the back or the neck. Reference: *Widespread pain* = pain in both the back and the neck (n = 100).

⁴ Reference: *Neck-pain domination* (n = 38).

⁵ Reference: Surgery for back and/or neck pain at least once (for example, for a slipped disc) (n = 8).

⁶ Item 5 in EQ-5 D[20], alternative 1 = "I am not anxious or depressed". Reference: alternative 2: "... moderately..." or 3: "... extremely...".

⁷ One item from SF 36[67]: "Tired during the last four weeks: 'some of the time', 'a little bit of the time' or 'none of the time'". Reference: *Tired often* = 'all of the time', 'most of the time' or 'a good bit of the time'.

⁸ Reference: *Comorbidity* = any other, chronic disease except NSP or obesity (n = 45).

⁹ ODI (Oswestry Disability Index) scores general functional disability associated with back pain, 0-100%: 0-20% = minimal, 21-40% = moderate, 41-60% = severe, 61-100% = extremely severe to crippling disability[38]. Reference: *Non-severe functional impairment* = ODI < 41%[69]

¹⁰ Six catastrophizing thoughts, never-always, 0-6, are summarized, 0-36. *Non-catastrophizing* ≥ 15. Reference: *Catastrophizing* > 15 (n = 39).

Further explanations in Table 1.

Power calculation

The power calculation of the randomized controlled trial has been described in a previous study [9]. In this prospective cohort study we were reduced to analyze the number of patients who were already included in the randomized controlled trial. However, several prior prediction studies included a comparable number of patients, e.g. Eklund et al. [14] 149 patients, Lancourt et al. [17] 134 patients, and Linton et al. [15] 142 patients.

Stable return-to-work

Stable return-to-work for 6, 12, 18 and 24 months, and of disability pension in 24 months were calculated. The proportions were compared between the genders by univariate-logistic regression, adjusted for age (*Young age* = 18-44 years and *Older age* = 45-59 years) and are given with *p*-values [18]. In the logistic regression *Stable return-to-work* might have the values 1, including the degrees .25, .50, .75 and 1.00, or 0 = *Non-return-to-work*.

Multiple-logistic regression

We built 4 multiple-logistic regression models for each of the follow-ups at 6, 12, 18 and 24 months. The outcome variable was *Stable return-to-work*. The predictive variables were the above described objective, socioeconomic, subjective and treatment variables. Ordinal and continuous variables were divided into classes. The models were adjusted for gender and age. We first explored univariate analyses. The variables with a *p*-value of at most .10 are presented with odds ratios (OR), *p*-values and 95% confidence intervals (CI). They were included in a multiple model, from which the variables with *p*-values of .05 or higher were excluded stepwise to yield a model comprising only variables with *p*-values < .05. However, in the choice between a model with a larger number of variables including those with *p*-values of .05 or slightly above and a smaller model with *p*-values exclusively smaller than .05, the larger model was tested against the smaller model (STATA commands "estimates store full" and "lrtest full"). If that test produced a *p*-value smaller than .05, the larger model was chosen as the ultimate one, otherwise the smaller model [18]. All possible first-order interaction terms were tested in each model.

Although it is important that a multiple-logistic regression model includes all relevant predictor variables, it is also important that the model does not include more predictors than the given number of observations justify. The existence of sufficient events per variable was emphasized by Bagley et al. [19] in a large overview of logistic regression. The number of the less common of 2 possible outcomes (in our study *Stable return-to-work* or *Non-return-to-work*) divided by the number of predictor variables was recommended to be at least 10 and preferably more [20]. On the basis of the number of patients with *Stable return-to-work* (Table 4), the maximal possible numbers of predictors were calculated as 3, 5, 5-6 and 6 at 6, 12, 18 and 24 months, respectively. While the models of 18 and 24 months lived up to that with 5 and 4 variables each, the models of 6 and 12 months included 5 and 6 variables, which necessitated the exclusion of 2 and 1 predictors respectively. We excluded from the 6-month model *No vibrations*, OR 5.9 [1.7-20.8](95% CI), *p* = .006, since this variable was represented in only one of the other follow-ups; and *Tired seldom*, OR 3.3 [1.2-9.4], *p* = .02, since it was not found in other follow-ups. From the 12-month model *No vibrations*, OR 3.2 [1.1-9.3], *p* = .03, was excluded, since it was found in only one of the other follow-ups. By the exclusion of *No vibrations* one of the remaining variables, *Man*, became non-significant (*p* changed from .02 to .08), and was left outside the final presentation of the 12-month model. Like *No vibrations*, *Back-pain domination* was represented in only the 6- and 12-month follow-ups, but was retained in the models due to its outstanding OR in 6 months. The results of the final models are shown as OR with *p*-values and 95% CI with goodness-of-fit tests by Hosmer-Lemeshow, the percentages of correctly predicted patients and the areas under ROC-curves [18].

We found no comparable studies of return-to-work at several time-points; for example, Hansson et al. [21] analyzed return-to-work at 90 days, 12 and 24 months, but their study included ~1500 subjects and no objective variables. However, we appraised that the comparatively small number of cases in our study made it risky to associate certain predictors with certain time-points. We chose to

Table 4 *Stable return-to-work*

	6 months			12 months			18 months			24 months		
	n	<i>Stable return-to-work</i> n (%)	<i>p</i>	n	<i>Stable return-to-work</i> n (%)	<i>p</i>	n	<i>Stable return-to-work</i> n (%)	<i>p</i>	n	<i>Stable return-to-work</i> n (%)	<i>p</i>
All	124	33 (26.6)	-	123	48 (39.0)	-	122	55 (45.1)	-	122	58 (47.5)	-
Men	56	19 (33.9)	NS	55	27 (49.1)	NS	54	29 (53.7)	NS	54	30 (55.6)	NS
Women	68	14 (20.6)	NS	68	21 (30.9)	NS	68	26 (38.2)	NS	68	28 (41.2)	NS

NS = Non-significant (*p* ≥ .05).

All patients and gender. The proportions are compared by univariate logistic regression, adjusted for age.

take into final consideration only the variables that were represented in at least 3 of the 4 follow-ups.

Ethical approval

Approval for the study was given by The Research Ethics Committee, Karolinska University Hospital, Huddinge.

Results

A flow-chart of the study is shown in Figure 1.

Source population

From data in a cross-sectional study under preparation, the point prevalence of severe spinal pain in the source

population was estimated at 15.6% or ~45,000 subjects, and of full-time sick-listing due to spinal pain to .8% or ~2,300 subjects, including short- and long-term sick-listing. The data were collected from Statistics Sweden, a governmental authority [22]. The great majority of patients with disabling NSP recovers quickly. Roughly, after full-time sick-listing 1 week around 50% and after 12 weeks 90% of the patients have returned to work. Thereafter the recovering speed evidently levels off [8]. We estimated the point prevalence of non-acute NSP in the source population to around .2% or 500 subjects. We have no data of the prevalence over time.

Loss to follow-up

Three of the 125 patients, all males, deceased during the follow-up, 11, 12 and 22 months after baseline. The last deceased patient was excluded from the study because of an incomplete questionnaire. The other 2 subjects were analyzed up to their possible follow-ups. The questionnaires of 124 patients were analyzed and sick-listing data were collected at 6, 12, 18 and 24 months for 124, 123, 122 and 122 patients, respectively.

Study population

Of the 124 patients, *Subacute NSP* and *Chronic NSP* occurred in 38 (30.6%) and 86 (69.4%) patients, respectively. The current sick-listing period at baseline was m 172 [149-194], days. *Back-pain domination* and *Neck-pain domination* was seen in 86 (69.4%) and 38 (30.6%) patients, respectively. Twenty-four patients (19.4%) had *Local pain*, i.e. back or neck pain, and 100 patients (80.6%) had *Widespread pain*, i.e. both back and neck pain.

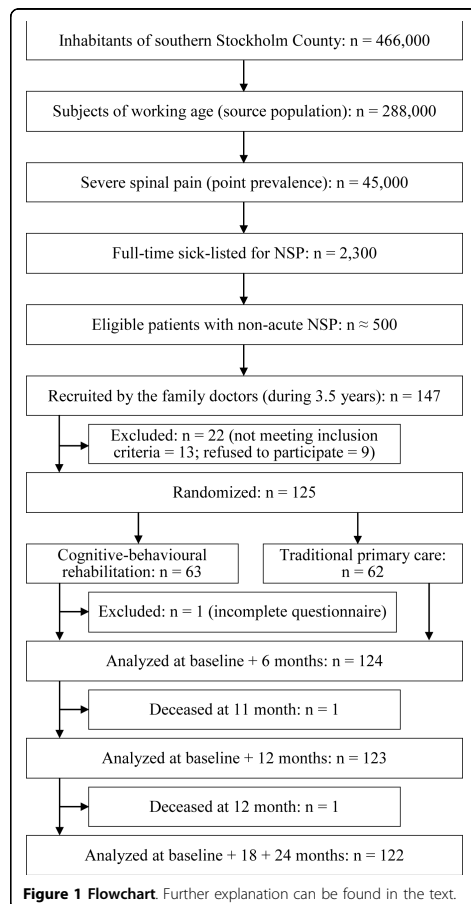
Stable return-to-work

Stable return-to-work gradually increased and was 58/122 (47.5%) at 24 months, a majority at full-time (43/58 = 74.1%). The proportions were generally higher for men, but the gender differences were non-significant (Table 4). At 24 months, disability pension (temporary or permanent) was received by 30/122 (22 full- and 8 half-time pensions), with a significantly higher proportion of women, 22/68 (32.4%), than men, 8/54 (14.8%) ($p = .04$).

Predictors of Stable return-to-work

In the univariate analyses, several objective, socioeconomic and subjective variables were associated with *Stable return-to-work* (Tables 1, 2, 3, while the treatment variables, *Cognitive-behavioural rehabilitation* or *Traditional primary care*, were not predictive in any of the follow-ups.

In the multiple-logistic models only socioeconomic and subjective variables remained, of which 3 variables were finally considered, all of them represented in



3 follow-ups (Table 5): *Low total prior sick-listing*, including all diagnoses, was the strongest predictor in 2 follow-ups, and *High self prediction* and *Young age* were the strongest and second strongest, respectively, in one of the follow-ups. In the models there were 3, 6, 10 and 10 first-order interaction terms, respectively, but none was predictive. The model fit was generally good and the proportions of correctly classified patients were satisfactory (on average 74.1%).

Discussion

The predictors of stable return-to-work were analyzed in 124 patients with non-acute NSP. Of the total of 50 variables, 2 socioeconomic variables, *Low total prior sick-listing* and *Young age*, and 1 subjective variable, *High self prediction*, were finally considered. None of the objective variables from function tests and of the treatment variables were predictive.

Predictors in the study compared with prior research

Young age is in line with several previous studies and reviews [1,23-27]. Also *High self prediction* is a well known predictor [14,15,28,29]. For example, the basic question in the clinical algorithm for return-to-work prediction by Dionne et al. [28] concerned the patient's own recovery expectations.

One of the most consistent predictors in previous research was low prior sick-listing for spinal pain [26,30,31]. While *Subacute NSP* was one of the strongest predictors in the univariate analyses, it was outflanked in the multiple context by *Low total prior sick-listing*, except at 6 months. According to one of the hitherto most extensive reviews of predictors of long-term sick-listing for spinal pain, prior sick-listing for all diagnoses has been insufficiently studied [1]. Our results indicate that it is very important to map prior sick-listing for all

diagnoses, not only for spinal pain. This is also in line with the Örebro Musculoskeletal Pain Questionnaire, a widely-used screening instrument [15], and with fairly recent studies with prolonged follow-ups [12,32].

Non-predictors in the study compared with prior research

Two non-predictors were in line with previous studies, *Comfortable work postures* [26,30] and *Good social support* [25,26]. *Non-Smoking* as a non-predictor is supported by some studies [33,34], but is contradicted by others: a large review of mostly cross-sectional studies indicated a possible association between NSP and cigarette smoking, but emphasized the lack of prospective studies [35]. Recent, prospective studies pinpointed cigarette smoking as a strong predictor of non-return-to-work in men, Dionne et al. [5], and as a moderate predictor of non-return in both sexes, Stillgate et al. [36].

Six of our non-predictors contradicted prior research:

Man and *Non-low education* were non-predictors, while prior research indicated them as predictors, at least for disability pension [1,12]. However, the proportion of disability pension at 24 months was significantly lower for men and *Non-low education* was close to qualify with a representation at both 18 and 24 months. It is logical that a disability pension will be granted only after prolonged sick-listing and that education might influence return-to-work comparatively late in a rehabilitation process, when the medical efforts have been replaced by vocational measures. Consequently, our findings might be in line with prior research, although a longer follow-up than 2 years is required to confirm this.

High physical workload, the reference to *Light physical workload*, is a well-established predictor of low return-to-work [26,30,34,37], but was non-predictive in our

Table 5 Predictors of Stable return-to-work

	Prediction for <i>Stable return-to-work</i>											
	6 months			12 months			18 months			24 months		
	OR	p	95%CI	OR	p	95%CI	OR	p	95%CI	OR	p	95%CI
<i>Young age</i>	-	-	-	2.8	.02	1.2-6.5	3.5	.001	1.3-9.1	2.7	.02	1.2-6.2
<i>Non-low education</i>	-	-	-	-	-	-	3.0	.04	1.1-8.2	2.9	.02	1.2-6.9
<i>Subacute NSP</i>	3.2	.02	1.3-8.2	-	-	-	3.0	.04	1.1-8.4	-	-	-
<i>Low total prior sick-listing</i>	-	-	-	2.7	.02	1.2-6.4	4.8	.001	1.9-12.3	3.8	.002	1.6-8.7
<i>Back-pain domination</i>	9.5	.004	2.0-44.4	2.9	.04	1.1-7.7	-	-	-	-	-	-
<i>Non-catastrophizing</i>	-	-	-	-	-	-	3.4	.01	1.3-9.1	-	-	-
<i>High self prediction</i>	4.1	.02	1.1-15.7	5.2	.009	1.5-17.5	-	-	-	2.7	.06	.9-7.8
Goodness-of-fit:												
Hosmer-Lemeshow		.70			.38			.29			.67	
Correctly classified (%)		78.2			71.5			73.0			73.8	
Area under ROC		.79			.79			.85			.79	

Multiple-logistic regression. The variables found in at least three follow-ups are in bold text.

study. The large majority of our patients (83.2%) had a *High physical workload* (compared to 15.4% of a population-based local sample in a cross-sectional study in preparation [22]). Thus, a variable of such overwhelming frequency might be non-discriminative, although it has a powerful effect on sick-listing.

Non-severe functional impairment, as measured by the Oswestry Disability Index [38-40], *Health-related quality of life*, according to EQ-5 D [21,41], and *State of health*, as expressed by EQ VAS [21], were comparatively strong predictors in the univariate analyses, but non-predictors in the final multiple-logistic models. This is contrary to previous studies [5,21,25,38-40], for which we can offer no explanation.

Non-predictors in the study that have previously been insufficiently studied

Many of our non-predictors that have been insufficiently studied in previous research might contribute to a widening of knowledge: *Non-immigrant*, *Co-habiting*, *Living without children*, *Non-unemployment*, *No work trauma litigation*, *Non-bad economy*, *Non-obese*, *No comorbidity*, *No surgery for spinal pain*, *Pain duration*, *Pain intensity*, *Local pain*, *Back-pain domination*, *High physical activity*, *Varied work moments*, *No job strain*, *No depression/anxiety* and *No indications of alcohol over-consumption* [1].

Concerning pain localisation and alcohol, prior studies are conflicting: While the predictive value of spinal pain localization has been questioned [1,15], recent research, including very large samples, supports the positive effect on return-to-work of *Back-* versus *Neck-pain domination* [21,32]. *Back-pain domination* in our study was near to qualify with a strong representation in 2 follow-ups. So, the non-prediction might be due to the comparatively low number of patients. While one study showed no association between alcohol over-consumption and sick-listing for spinal pain [41], another study found that alcohol abuse was higher among persons with chronic spinal pain [42]. A recent large study indicated that moderate alcohol consumption tended to decrease sick-listing for NSP, at least among women in the public sector [36].

Objective versus subjective variables

As few of the function tests commonly used in previous research were validated, it is difficult to judge from prior studies if objective variables are predictive [6]. For example, in a Cochrane review of specific spinal pain, subjective variables such as the state of health predicted return-to-work, but there was insufficient scientific support concerning objective variables, such as strength or motion range [7]. Our study strongly supports the

predictive value of subjective predictors and might widen the knowledge of objective variables as non-predictors.

Treatment as a predictor of return-to-work

For the entire group of patients, treatment was non-predictive. In a previous study [9], there were indications that patients with *Subacute NSP* had a greater return-to-work chance when they received the cognitive-behavioural programme. However, a more detailed evaluation of the possible positive effect on return-to-work of our programme requires other analyses than in the present study - for example, survival analysis as in the previous study [9] - and is a matter for future work.

Strengths of the study

The prospective design, with a comparatively long follow-up period, is a major strength of our study.

The generalisation of the results of previous research on the prediction of return-to-work in spinal pain is seriously limited by the under-representation of women [1]. Thus one strength of our study has been the good representation of women.

We have no data of the proportion of work obstacles due to back pain compared with neck pain in the source population. In previous research, the annual prevalence in industrial countries of work obstacles due to back pain and neck pain has been estimated to 8% and 2%, respectively [43]. We obtained a similar ratio, which might indicate that our study sample is representative of subjects with non-acute NSP.

Because we used data from the Social Insurance Office, no sick-listing data was missing, except the possible short-term relapses of non-return-to-work during the follow-up months. With the exclusion of one patient, the questionnaire data were complete.

The use of reliable function tests is a major strength. One of the examiners in our reliability study [11], the research assistant, also carried out the function tests in this study.

Limitations of the study

Some circumstances might have decreased the representativeness of the study sample, and increased the risk of bias. The above-mentioned annual prevalence of work limiting back pain and neck pain corresponds to ~23,000 and ~5,000, respectively, in the source population. Though these data include short-term sick-listing also, it is obvious that the study population of 125 patients recruited over a period of 3.5 years constituted a very low percentage of the eligible subjects. As a comparison, Dionne et al. [28] achieved a participation rate of 68.4% of eligible subjects. The inclusion was

non-systematic: a family doctor with a local reputation of great skills in spinal pain might attract more complex cases, and have a higher motivation for research and the recruitment of study patients. This might lead to spectrum bias, i.e. the effect the patient mix may have on the performance of tests, e.g., a package of predictors [44]. We were overoptimistic concerning the recruiting propensity of the family doctors and lacked resources to make them more compliant. This contributed to a prolonged inclusion period (3.5 years) that increased the probability of societal changes in rules and attitudes concerning sick-listing and might result in different return-to-work predictors in identical spinal pain due to inclusion either early or late in the recruitment period. The problem with protracted inclusion periods is shared with several other studies. For example, Lindström et al. [45] and Loisel et al. [46] used 2.5 years for the inclusion of 103 and 130 patients, respectively, and Jensen et al. [47] 3.5 years for 214 patients. As a comparison, Dionne et al. [28] used a systematic approach and recruited 1007 patients in about 1.5 years.

While it is advocated that predictive conclusions might be drawn exclusively from studies with a sick-listing-baseline on day zero [48], our patients had been sick-listed for at least 6 weeks at baseline, which might be seen as a limitation. However, even in the above-mentioned large review [1], several of the studies had baselines similar to ours [17,26,34] and arguably it is also of great interest to predict return-to-work in non-acute NSP.

Work satisfaction as a separate variable was not included. Since work satisfaction was indicated as a return-to-work predictor in several previous studies [25,49,50], it is a limitation.

There is no gold standard enabling the analysis of the time-points of return-to-work [51], but logically different predictors have a different impact in different time-points. While education might have a stronger influence comparatively late, pain and other subjective variables might affect the outcome early. It is also of great interest to know what variables predict return-to-work and when. For example, prediction of return-to-work, but not until 24 months, might be of no use concerning a patient close to old-age pension. A limitation of our study is that the follow-ups are not mutually compared, which should require a larger number of cases.

As cognitive-behavioural therapy, among other items, addresses dysfunctional beliefs [52], *Cognitive-behavioural rehabilitation* given to half of the patients might have a greater impact on the self prediction and result in an underestimation in the association between *High self prediction* and return-to-work. This might be a limitation of the study. However, as none of the treatment variables predicted *Return-to-work*, we

consider the potential bias achieved by the treatment as negligible.

Conclusions

In primary-care patients with non-acute, non-specific spinal pain, including back and/or neck pain, the strong predictors of stable return-to-work were 2 socioeconomic variables, *Low total previous sick-listing* (including all diagnoses) and *Young age* (max 44 years), and 1 subjective variable, *High self prediction* (the patients' own belief in return-to-work). Objective variables from function tests and treatment variables (a programme of cognitive-behavioural rehabilitation or traditional primary care) were non-predictors. Except for *Young age*, the predictors had been insufficiently studied in previous research. Hence, our study might contribute to a widening of knowledge within clinical practice, including the allocation of treatment resources.

Acknowledgements

This study was supported by grants from the Stockholm County Social Insurance Agency, Stockholm County Council, Ministry of Health and Social Affairs, Vårdal Foundation, Cardionics, Pharmacia (now part of Pfizer) and Grunenthal Sweden AB.

Authors' contributions

OL was the main investigator, carried out the study, performed the analysis and drafted the manuscript. SEJ contributed to the statistical analysis. LES, as supervisor of OL, participated in all phases of the study. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 18 December 2009 Accepted: 20 July 2010
Published: 20 July 2010

References

- Hansson T, Jensen I: *Sickness absence due to back and neck disorders. Sickness absence - causes, consequences, and physicians' sickness-certification practice* Scand J Public Health/Alexanderson K, Norlund A 2004, **32**(Suppl 63):109-151.
- Newly awarded disability pensions. Distribution between counties and diagnoses, 2006. [Swedish; Nybeviljade sjukersättningar/aktivitetsersättningar. Fördelning på län och diagnos, 2006] Social Insurance in Sweden. Statistics 2007, **7** [http://statistik.forsakringskassan.se/rfvexcel/statpub/Statistik_2007-02.pdf]. [In Swedish; Försäkringskassan. Statistik 2007:7].
- Key trends and outcomes. *Sickness, Disability and Work. Keeping on track in the economic downturn* OECD, Directorate for employment, labour and social affairs, Background paper: High-Level Forum, Stockholm 2009 [http://www.oecd.org/dataoecd/42/15/42699911.pdf].
- Waddell G, van Tulder M: *Clinical guidelines. Back Pain Revolution* Edinburgh: Churchill Livingstone/Waddell G, 2004, 283-322.
- Dionne CE, Bourbonnais R, Frémont P, Rossignol M, Stock SR, Nouwen A, Larocque I, Demers E: *Determinants of "return to work in good health" among workers with back pain who consult in primary care settings: a 2-year prospective study.* *Eur Spine J* 2007, **16**:641-655.
- Ljungquist T: *Physical performance tests and spinal pain. Assessing impairments and activity limitations.* PhD thesis Division of Physiotherapy, Neurotec Department, and Section for Personal Injury Prevention, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden 2002.
- Ostelo RW, de Ver HC, Waddell G, Kerckhoffs MR, Leffers P, van Tulder M: *Rehabilitation following first-time lumbar disc surgery: a systematic*

- review within the framework of the Cochrane collaboration. *Spine* 2003, **28**:209-218.
8. Waddell G: **The biopsychosocial model.** *Back Pain Revolution* Edinburgh: Churchill Livingstone/Waddell G, 2004, 265-282.
 9. Lindell O, Johansson SE, Strender LE: **Subacute and chronic, non-specific back and neck pain: cognitive-behavioural rehabilitation versus primary care. A randomized controlled trial.** *BMC Musculoskelet Disord* 2008, **9**:172.
 10. Platzer W: **Color Atlas and Textbook of Human Anatomy. Locomotor system** Georg Thieme Verlag Stuttgart 1978, 1:141.
 11. Lindell O, Eriksson L, Strender LE: **The reliability of a 10-test package for patients with prolonged back and neck pain: Could an examiner without formal medical education be used without loss of quality? A methodological study.** *BMC Musculoskelet Disord* 2007, **8**:31.
 12. Borg K, Hensing G, Alexandersson K: **Predictive factors for disability pension - an 11-year follow up of young persons on sick leave due to neck, shoulder or back diagnoses.** *Scand J Public Health* 2001, **29**:104-112.
 13. Arrelöv B, Borgquist L, Ljungberg D, Svärdsudd K: **Do GPs sick-list patients to a lesser extent than other physician categories? A population-based study.** *Fam Pract* 2001, **18**:393-398.
 14. Eklund M, Fugl-Meyer AR: **Vocational rehabilitation in northern Sweden. II. Some psycho-socio-demographic predictors.** *Scand J Rehab Med* 1991, **23**:73-82.
 15. Linton SJ, Halldén K: **Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain.** *Clin J Pain* 1998, **14**:209-215.
 16. StataCorp: *Stata Statistical Software: Release 9* College Station, TX: StataCorp LP 2005.
 17. Lancourt J, Kettelhut M: **Predicting return to work for lower back pain patients receiving worker's compensation.** *Spine* 1992, **17**:629-640.
 18. Hosmer DW, Lemeshow S: *Applied logistic regression* New York: John Wiley & Sons, 2000.
 19. Bagley SC, White H, Golomb BA: **Logistic regression in the medical literature: Standard for use and reporting, with particular attention to one medical domain.** *J Clin Epidemiol* 2001, **54**:979-985.
 20. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR: **A simulation study of the number of events per variable in logistic regression analysis.** *J Clin Epidemiol* 1996, **49**:1373-1379.
 21. Hansson E, Hansson T, Jonsson R: **Predictors for work ability and disability in men and women with low-back or neck problems.** *Eur Spine J* 2006, **15**:780-793.
 22. **Living conditions. Appendix 16.** *The Swedish Survey of Living Conditions. Design and methods* Official Statistics of Sweden. Statistics Sweden [http://www.scb.se/statistik/LE/LE0101/_dokument/Appendix_16.PDF].
 23. Oleinick A, Gluck JV, Guire K: **Factors affecting first return to work following a compensable occupational injury.** *Am J Ind Med* 1996, **30**:540-555.
 24. Hagen KB, Holte HH, Tambs K, Bjerkedal T: **Socioeconomic factors and disability retirement from back pain: a 1983-1993 population-based prospective study in Norway.** *Spine* 2000, **25**:2480-2487.
 25. van der Giesen AM, Bouter LM, Nijhuis FJ: **Prediction of return-to-work of low back pain patients sicklisted for 3-4 months.** *Pain* 2000, **87**:285-294.
 26. Krause N, Dasinger LK, Deegan LJ, Rudolph L, Brand RJ: **Psychosocial job factors and return-to-work after compensated low back injury: a disability phase-specific analysis.** *Am J Ind Med* 2001, **40**:374-392.
 27. Karlsson N, Borg K, Carstensen J, Hensing G, Alexanderson K: **Risk of disability pension in relation to gender and age in a Swedish county; a 12-year population based, prospective cohort study.** *Work* 2006, **27**:173-179.
 28. Dionne CE, Bourbonnais R, Frémont P, Rossignol M, Stock SR, Larocque I: **A clinical return-to-work rule for patients with back pain.** *CMAJ* 2005, **172**:1559-1567.
 29. Du Bois M, Szpalski M, Donceel P: **Patients at risk for long-term sick leave because of low back pain.** *Spine* 2009, **34**:350-359.
 30. Dasinger LK, Krause N, Deegan LJ, Brand RJ, Rudolph L: **Physical workplace factors and return to work after compensated low back injury: a disability phase-specific analysis.** *J Occup Environ Med* 2000, **42**:323-333.
 31. Waddell G, Burton K: **Occupational health guidelines for the management of low back pain at work: evidence review.** *Occup Med* 2001, **51**:124-135.
 32. Bergström G, Bodin L, Bertilsson H, Jensen IB: **Risk factors for new episodes of sick leave due to neck or back pain in a working population. A prospective study with an 18-month and a three-year follow-up.** *J Occup Environ Med* 2007, **49**:279-287.
 33. Boshuizen HC, Verbeek JH, Broersen JP, Weel AN: **Do smokers get more back pain?** *Spine* 1993, **18**:35-40.
 34. Hansson TH, Hansson EK: **The effects of common medical interventions on pain, back function, and work resumption in patients with chronic low back pain: A prospective 2-year cohort study in six countries.** *Spine* 2000, **25**:3055-3064.
 35. Goldberg MS, Scott SC, Mayo NE: **A review of the association between cigarette smoking and the development of nonspecific back pain and related outcomes.** *Spine* 2000, **25**:995-1014.
 36. Skillgate E, Vingård E, Josephson M, Holm LW, Alfredsson L: **Is smoking and alcohol consumption associated with long-term sick leave due to unspecific back or neck pain among employees in the public sector? Results of a three-year follow-up cohort study.** *J Rehabil Med* 2009, **41**:550-556.
 37. Fransén M, Woodward M, Norton R, Coggan C, Dawe M, Sheridan N: **Risk factors associated with the transition from acute to chronic occupational back pain.** *Spine* 2002, **27**:92-98.
 38. Fairbank JC, Couper J, Davies JB, O'Brien JP: **The Oswestry low back pain disability questionnaire.** *Physiotherapy* 1980, **66**:271-273.
 39. Fairbank JC, Pynsent PB: **The Oswestry disability index.** *Spine* 2000, **25**:2940-2953.
 40. Du Bois M, Szpalski M, Donceel P: **Patients at risk for long-term sick leave because of low back pain.** *Spine* 2009, **34**:350-359.
 41. Holmberg SA, Thelin AG: **Primary care consultation, hospital admission, sick leave and disability pension owing to neck and low back pain: a 12-year prospective cohort study in a rural population.** *BMC Musculoskelet Disord* 2006, **7**:66.
 42. Demyttenaere K, Bruffaerts R, Lee S, Posada-Villa J, Kovess V, Angermeyer MC, Levinson D, de Girolamo G, Nakane H, Mneimneh Z, Lara C, de Graaf R, Scott KM, Gureje O, Stein DJ, Haro JM, Bromet EJ, Kessler RC, Alonso J, Von Korff M: **Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys.** *Pain* 2007, **129**:332-342.
 43. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA: **Comprehensive review of epidemiology, scope, and impact of spinal pain.** *Pain Physician* 2009, **12**:E35-70.
 44. Willis BH: **Spectrum bias - why clinicians need to be cautious when applying diagnostic test studies.** *Family Practice* 2008, **25**:390-396.
 45. Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson A: **The effect of graded activity on patients with sub-acute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach.** *Phys Ther* 1992, **72**:279-293.
 46. Loisel P, Abenhaim L, Durand P, Esdaile JM, Suissa S, Gosselin L, Simard R, Turcotte J, Lemaire J: **A population-based, randomized clinical trial on back pain management.** *Spine* 1997, **22**:2911-2918.
 47. Jensen IB, Bergström G, Ljungquist T, Bodin L: **A 3-year follow-up of a multidisciplinary rehabilitation programme for back and neck pain.** *Pain* 2005, **115**:273-283.
 48. Steenstra JA, Verbeek JH, Heymans MW, Bongers PM: **Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature.** *Occup Environ Med* 2005, **62**:851-860.
 49. Staal J, Hlobil H, van Tulder MW, Köske AJ, Smid T, van Mechelen W: **Return-to-work interventions for low back pain: a descriptive review of contents and concepts of working mechanisms.** *Sports Med* 2002, **32**:251-267.
 50. Lakke SE, Soer R, Takken T, Reneman MF: **Risk and prognostic factors for non-specific musculoskeletal pain: a synthesis of evidence from systematic reviews classified into ICF dimensions.** *Pain* 2009, **147**:153-164.
 51. Wasiak R, Young AE, Roessler RT, McPherson KM, van Poppel MN, Anema JR: **Measuring return to work.** *J Occup Rehabil* 2007, **17**:766-781.
 52. Waddell G, Watson PJ: **Rehabilitation.** *Back Pain Revolution* Edinburgh: Churchill Livingstone/Waddell G, 2004, 371-399.
 53. Biering-Sørensen F, Thomsen C: **Medical, social and occupational history as risk indicators for low-back trouble in a general population.** *Spine* 1986, **11**:720-725.
 54. Mastekaasa A: **Parenthood, gender and sickness absence.** *Soc Sci Med* 2000, **50**:1827-1842.
 55. Åstrand NE, Isacson SO: **Back pain, back abnormalities, and competing medical, psychological, and social factors as predictors of sick leave,**

- early retirement, unemployment, labour turnover and mortality: a 22 year follow up of male employees in a Swedish pulp and paper company. *Br J Ind Med* 1988, **45**:387-395.
56. Waddell G, Waddell H: **A review of social influences on neck and back pain and disability.** *Neck and back pain. The scientific evidence of causes, diagnosis, and treatment* Philadelphia: Lippincott, Williams & Williams/Nachemson A, Jonson E 2000, 13-55.
 57. Linton SJ: **Risk factors for neck and back pain in a working population in Sweden.** *Work & Stress* 1990, **4**:41-49.
 58. Tomqvist EW, Kilbom A, Vingård E, Alfredsson L, Hagberg M, Theorell T, Waldenström M, Wiktorin C, Hogstedt C, MUSIC-Norrälja Study Group: **The influence on seeking care because of neck and shoulder disorders from work-related exposures.** *Epidemiology* 2001, **12**:537-545.
 59. Ghaffari M, Alipour A, Farshad AA, Jensen I, Josephson M, Vingård E: **Effect of psychosocial factors on low back pain in industrial workers.** *Occup Med* 2008, **58**:341-347.
 60. Theorell T, Perski A, Åkerstedt T: **Changes in job strain in relation to changes in psychological state: a longitudinal study.** *Scand J Work Environ Health* 1988, **14**:189-196.
 61. Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM: **Systematic review of psychosocial factors at work and private life as risk factors for back pain.** *Spine* 2000, **25**:2114-2125.
 62. Selander J, Marnett SU, Bergroth A, Ekholm J: **Return to work following vocational rehabilitation for neck, back and shoulder problems: risk factors reviewed.** *Disabil Rehabil* 2002, **24**:704-712.
 63. Trief P, Stein M: **Pending litigation and rehabilitation outcome of chronic low back pain.** *Arch Phys Med Rehabil* 1985, **66**:95-99.
 64. Romelsjö A, Leifman H, Nyström S: **A comparative study of two methods for the measurement of alcohol consumption in the general population.** *Int J Epidemiol* 1995, **24**:929-936.
 65. Oleske DM, Lavender SA, Andersson GB, Morrissey MJ, Zold-Kilbourn P, Allen C, Taylor E: **Risk factors for recurrent episodes of work-related low back disorders in an industrial population.** *Spine* 2006, **31**:789-798.
 66. Neovius M, Janson A, Rössner S: **Prevalence of obesity in Sweden.** *Obes Rev* 2006, **7**:1-3.
 67. Sullivan M, Karlsson J, Ware JE Jr: **The Swedish SF-36 Health Survey. I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden.** *Soc Sci Med* 1995, **41**:1349-1358.
 68. Nordin M, Hiebert R, Pietrek M, Alexander M, Crane M, Lewis S: **Association of comorbidity and outcome in episodes of nonspecific low back pain in occupational populations.** *J Occup Environ Med* 2002, **44**:677-684.
 69. Sanderson PL, Todd BD, Holt GR, Getty CJ: **Compensation, work status, and disability in low back pain patients.** *Spine* 1995, **20**:554-556.
 70. Jensen IB, Linton SJ: **Coping strategies questionnaire (CSQ): reliability of the Swedish Version of the CSQ.** *Scand J Behaviour Therapy* 1993, **22**:139-145.

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2296/11/53/prepub

doi:10.1186/1471-2296-11-53

Cite this article as: Lindell et al.: Predictors of stable return-to-work in non-acute, non-specific spinal pain: low total prior sick-listing, high self prediction and young age. A two-year prospective cohort study. *BMC Family Practice* 2010 **11**:53.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



