Back and neck pain. Factors of importance for the prognosis

Tony Bohman

Stockholm 2013
Till mamma och pappa

“The mind is like a parachute.
It doesn’t work unless it’s open”
Frank Zappa
ABSTRACT

Back pain and neck pain are very common and among the most frequent causes of sick-leave and disability pension, thereby greatly affecting the individual and the community. This stresses the need for prognostic research regarding these conditions. **Aim:** The main aim of the present work was to study prognostic factors, including manual therapy, for back and neck pain. The specific aims were, to investigate the influence of regular leisure physical activity and the body mass index (BMI) on the recovery from persistent low back pain (*Study I*), to study the influence of healthy lifestyle behaviour on the prognosis of occasional low back pain (*Study II*), to explore the long-term effects (up to one year) of napraphatic manual therapy for patients with non-specific back and/or neck pain (*Study III*), and to develop a prediction model for the recovery from whiplash-associated disorders (WAD) in patients who consulted physiotherapy (*Study IV*). **Methods:** *Studies I* and *II* were based on the Stockholm Public Health Cohort, and comprised data from four questionnaire-based public-health surveys conducted between 2002 and 2010. *Study I* included 1836 individuals reporting persistent low back pain at baseline in 2002 and answering the follow-up in 2007, while *Study II* involved 8994 individuals with occasional low back pain at baseline in 2006 responding to the 2010 survey. Information on exposures and potential confounders was collected at baseline. The exposures were, regular leisure physical activity and BMI (*Study I*), and “healthy lifestyle behaviour”, a combination of four lifestyle factors (smoking habits, alcohol consumption, leisure physical activity and consumption of fruit and vegetables) (*Study II*). Both *Studies I* and *II* assessed men and women separately. *Study III* was based on a Swedish randomized controlled trial of 409 patients with non-specific back and/or neck pain. It compared napraphatic manual therapy with evidence based support on staying active and on pain coping strategies. Questionnaires at 26 and 52 weeks provided the follow-up data. *Study IV* included 680 patients with WAD consulting physiotherapy, using data retrieved from the Saskatchewan Government Insurance study, Canada (1997-1999). A prediction model for recovery from WAD was developed and internally validated by assessing twenty-five possible prognostic factors, using survival analyses. **Results and Conclusions:** Regular leisure physical activity improved recovery from persistent low back pain among women. No such association was found among men, or between BMI and recovery regardless of sex (*Study I*). Healthy lifestyle behaviour decreased the risk of long duration troublesome low back pain among women with occasional low back pain. No clear association was found among men (*Study II*). Compared to evidence-based care, napraphatic manual therapy implied greater long-term improvement in pain and disability for patients with non-specific back and/or neck pain (*Study III*). The prediction model developed includes seven clinically important prognostic factors, and has acceptable predictive ability (*Study IV*). The conclusions in this thesis are, that lifestyle factors are of importance for the prognosis of low back pain among women, that combined manual therapy, such as napraphathy, has a long-term effect on non-specific back and/or neck pain, and that the present prediction model for recovery from WAD has acceptable predictive ability but has to be further validated to be used in clinical practice.
SVENSK SAMMANFATTNING

Smärta i ländrygg och nacke är mycket vanligt och bland de mest frekventa orsakerna till sjukskrivning och förtidspension i Sverige och internationellt. Konsekvenserna av dessa tillstånd påverkar såväl den enskilda individen som samhället i stort. Därför är forskning om vad som har betydelse för prognosen av ländryggs- och nacksmärta viktig.

Det övergripande syftet med detta arbete var att studera faktorer, inklusive manuell terapi, som kan påverka prognosen för smärtor i ländrygg och nacke. De specifika målen med de i arbetet ingående studierna var, att undersöka påverkan av regelbunden fysisk aktivitet på fritiden respektive BMI (the body mass index) på tillfrisknande från ihållande ländryggssmärta (Studie I), att studera påverkan av hälsosam livsstil på tillfrisknande från nacksmärta efter en whiplash skada, s.k. whiplash-associated disorders (WAD), för patienter som konsulterat en sjukgymnast (Studie IV).


Resultaten i Studie I visade att regelbunden fysisk aktivitet förbättrar tillfrisknande från ihållande ländryggssmärta för kvinnor. Inget sådant samband upptäcktes för män. Vidare, upptäcktes inget samband mellan BMI och tillfrisknande varken hos kvinnor eller hos män. Studie II visade att en hälsosam livsstil minskar risken för långvarig

Konklusionerna i denna avhandling är, att livsstilsfaktorer är av betydelse för prognosen av ländryggssmärta hos kvinnor, att kombinerad manuell terapi, som exempelvis naprapati, har en långvarig effekt på icke-specifika rygg- och/eller nacksmärta och att modellen som utvecklades har godtagbar säkerhet att förutsäga tillfrisknande från WAD hos patienter som konsulterar en sjukgymnast men att den bör testas ytterligare för att kunna användas i kliniska situationer.
LIST OF PUBLICATIONS

This thesis is based on the following papers, referred to in the text by their Roman numerals (I - IV).


Additional analyses, not previously published, are added in the thesis.

*Study II* may not be the final version for publication.
CONTENTS

1 Introduction ........................................................................................................... 1

2 Background ........................................................................................................... 2
  2.1 Back and Neck Pain ......................................................................................... 2
    2.1.1 The biopsychosocial model ..................................................................... 2
    2.1.2 Definitions of back and neck pain .......................................................... 2
    2.1.3 Operational definitions in this thesis ....................................................... 5
  2.2 Back and neck pain in the population .............................................................. 6
    2.2.1 The course of back and neck pain ......................................................... 7
  2.3 Prognostic research ......................................................................................... 7
    2.3.1 Prognostic course studies ....................................................................... 9
    2.3.2 Prognostic factor studies ....................................................................... 9
    2.3.3 Outcome prediction studies (prediction models) ..................................... 9
  2.4 Prognostic research related to this thesis ....................................................... 9
    2.4.1 Treatment (prognostic course studies) ................................................... 9
    2.4.2 Prognostic factor studies ....................................................................... 11
    2.4.3 Prediction models (outcome prediction studies) ................................... 12
  2.5 Association and causality .............................................................................. 12
  2.6 Rationale ........................................................................................................ 13

3 Aims ...................................................................................................................... 14
  3.1 Main aim ......................................................................................................... 14
  3.2 Specific aims .................................................................................................. 14

4 Material and methods .......................................................................................... 15
  4.1 Study design and source population .............................................................. 15
    4.1.1 The Stockholm Public Health Cohort (Studies I and II) ....................... 15
    4.1.2 The Swedish BJÖRN trial (Study III) .................................................. 15
    4.1.3 The Saskatchewan Government Insurance study (Study IV) ............ 16
  4.2 Data collection and Participants .................................................................... 16
    4.2.1 Studies I and II ...................................................................................... 16
    4.2.2 Study III ............................................................................................... 17
    4.2.3 Study IV ............................................................................................... 19
  4.3 Measurements and interventions ................................................................... 20
    4.3.1 Studies I and II ...................................................................................... 20
    4.3.2 Study III ............................................................................................... 23
    4.3.3 Study IV ............................................................................................... 24
  4.4 Statistics and comparisons ............................................................................ 25
    4.4.1 Studies I and II ...................................................................................... 27
    4.4.2 Study III ............................................................................................... 28
    4.4.3 Study IV ............................................................................................... 28
  4.5 Ethics ............................................................................................................... 29

5 Results ................................................................................................................... 30
  5.1 Study I .......................................................................................................... 30
  5.2 Study II ......................................................................................................... 32
  5.3 Study III ....................................................................................................... 35
  5.4 Study IV ....................................................................................................... 38

6 Discussion ............................................................................................................. 40
6.1 Study I .................................................................................................................................40
  6.1.1 Findings and implications .................................................................................................40
  6.1.2 Methodological considerations .........................................................................................42
6.2 Study II ..................................................................................................................................44
  6.2.1 Findings and implications .................................................................................................44
  6.2.2 Methodological considerations .........................................................................................45
6.3 Study III ................................................................................................................................46
  6.3.1 Findings and implications .................................................................................................46
  6.3.2 Methodological considerations .........................................................................................47
6.4 Study IV ................................................................................................................................48
  6.4.1 Findings and implications .................................................................................................48
  6.4.2 Methodological considerations .........................................................................................49
6.5 External validity .....................................................................................................................50
6.6 Future perspectives .................................................................................................................51

7 Conclusions .............................................................................................................................53
8 Acknowledgements ....................................................................................................................54
9 References ...............................................................................................................................57
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>The body mass index</td>
</tr>
<tr>
<td>BP</td>
<td>(Low) back pain</td>
</tr>
<tr>
<td>CE</td>
<td>Change in estimate method</td>
</tr>
<tr>
<td>CES-D</td>
<td>Centre for Epidemiological Studies Depression Scale</td>
</tr>
<tr>
<td>CPQ</td>
<td>Chronic Pain Questionnaire</td>
</tr>
<tr>
<td>GEE</td>
<td>General Estimating Equation</td>
</tr>
<tr>
<td>HLB</td>
<td>Healthy lifestyle behaviour</td>
</tr>
<tr>
<td>HRR</td>
<td>Hazard rate ratio</td>
</tr>
<tr>
<td>LBP</td>
<td>Low back pain</td>
</tr>
<tr>
<td>LTLBP</td>
<td>Long duration troublesome low back pain</td>
</tr>
<tr>
<td>MD</td>
<td>Medical doctor</td>
</tr>
<tr>
<td>NRS</td>
<td>Numeric rating scale</td>
</tr>
<tr>
<td>OLBP</td>
<td>Occasional low back pain</td>
</tr>
<tr>
<td>PA</td>
<td>Regular leisure (time) physical activity</td>
</tr>
<tr>
<td>PBP</td>
<td>Persistent (low) back pain</td>
</tr>
<tr>
<td>PROGRESS</td>
<td>The Prognosis research strategy group</td>
</tr>
<tr>
<td>PT</td>
<td>Physiotherapist</td>
</tr>
<tr>
<td>RD</td>
<td>Risk difference</td>
</tr>
<tr>
<td>RR</td>
<td>Risk ratio</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>SGI</td>
<td>The Saskatchewan Government Insurance study (cohort)</td>
</tr>
<tr>
<td>SPHC</td>
<td>The Stockholm Public Health Cohort</td>
</tr>
<tr>
<td>SRH</td>
<td>Self-rated health</td>
</tr>
<tr>
<td>WAD</td>
<td>Whiplash-associated disorders</td>
</tr>
<tr>
<td>WDQ</td>
<td>Whiplash Disability Questionnaire</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

During my almost twenty years as a clinically practising physiotherapist, specialising in manual therapy, I have met patients with all kinds of back and neck pain. With increasing experience I noticed that when seeing two patients with similar descriptions of their problems and similar clinical findings I often have a strong sense of which one will have a good prognosis and which one will not. I find it unsatisfactory just to rely on a subjective feeling and this experience has aroused my interest for more knowledge about what factors actually influence the prognosis for back and neck pain. And how could such knowledge be used to enhance the management of these patients?

Starting to search for evidence regarding prognostic factors I, surprisingly, found that the knowledge was sparse. Since spinal pain is very common globally and may be an increasing health problem as the population ages, increasing our knowledge of prognostic factors is of great importance in secondary prevention\(^1\)-\(^3\). For example, better understanding of modifiable prognostic factors related to lifestyle may help affected individuals and health-care providers in managing back and neck pain. Further, increasing our knowledge of treatment and other prognostic factors would be valuable for health-care providers trying to predict the outcome for the individual patient, and for policy-makers in their decision regarding the handling of these problems.

Epidemiology has been defined in many ways as for example: “the study of the distribution and determinants of disease frequency in man” or simply “the study of the occurrence of illness”\(^4\)-\(^5\). Epidemiology is a scientific discipline ideal for studying the prognosis of disease and health problems.

In 2008, I got the opportunity to satisfy my curiosity about prognostic factors for back and neck pain by applying for this doctoral project in epidemiology, designed by my current supervisors. The project was financed by the Health Care Sciences Postgraduate School at Karolinska Institutet, Stockholm, Sweden. Luckily I was accepted, and hopefully this has resulted in a thesis that may be a contribution to the knowledge of prognostic factors of importance for back and neck pain.
2 BACKGROUND

2.1 BACK AND NECK PAIN

Pain is a basic mechanism for protecting us against possible body damage. It is an individual and subjective experience and as such a multidimensional phenomenon. Pain is an important survival mechanism but problematic and dysfunctional when it becomes long-lasting. The International Association for the Study of Pain (IASP) has defined pain as:

“An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”.

The definition incorporates pain distributed through nerve fibres from peripheral body tissues to the brain cortex (nociceptive pain). Further, as it defines pain as an experience, an individual may have pain even in the absence of detectable nociception. The experience of pain can be described by three components; the sensory, the affective and the cognitive. These components accentuate the complexity of pain in a biopsychosocial context. Pain can also be classified according to its aetiology: nociceptive, neuropathic (due to damage of the nerve fibre), idiopathic (unknown origin) and psychological.

Nociceptive back and neck pain are mainly due to stimuli from the musculoskeletal system (muscles, tendons, ligaments, joints, spinal bone and cartilage tissue). It can also be caused by nociceptive stimuli from internal organ, for example, kidneys, urogenital organs, the heart and vertebral vessels. It is then usually termed “referred pain”. Like pain in general, back and neck pain can likewise be caused or influenced by psychological mechanisms.

2.1.1 The biopsychosocial model

In the present work we studied prognostic factors for back and neck pain in the context of the biopsychosocial model proposed by Mosey in 1974 and Engel in 1977. The biopsychosocial approach recognises the role of biological, psychological and social factors, and their interaction in the context of back and neck pain and the related disability that can arise. The model is widely accepted and integrated in many recommendations and clinical guidelines as well as in the International Classification of Functioning, Disability and Health (ICF, WHO 2001). It has been emphasized that pain and disability are mutually differing subjective issues and that both entities are affected by physiological as well as mental factors. Even though both are related to the underlying physical disorder, some individuals have pain but little disability while others have disability out of proportion to their pain. This stresses the importance of assessing both pain and disability as well as physical, psychological and social factors in prognostic back and neck pain research.

2.1.2 Definitions of back and neck pain

The definitions and classification of back and neck pain are complex and somewhat confusing and there is still a lack of consensus, despite promising attempts.
The term ‘back pain’ can be confusing as it most commonly means low back pain but sometimes incorporates other parts of the back of the body such as the thorax and neck region. In a Delphi study from 2008, 82% of the participating experts agreed on back pain as only meaning pain in the lower back (low back pain)²⁶.

Low back pain is commonly defined as pain and discomfort located below the costal margin and above the inferior gluteal folds, with or without leg pain¹⁹,²⁰. A definition for neck pain is harder to find but the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and its Associated Disorders suggests; pain located in the anatomic region of the neck (between the superior nuchal line and the spine of the scapula from behind and covering the throat from the front), with or without radiation to the head, trunk and upper limbs¹⁷.

Further, spinal pain is most often classified into specific and non-specific pain using the “diagnostic triage”. Specific pain will then be equivalent to pain attributed to recognisable or known specific pathology, and non-specific equivalent to all other spinal pain. Guidelines and reviews recommend using the diagnostic triage and the concept of “flags” in the clinical assessment and management of specific and non-specific back and neck pain¹¹-¹⁴. Then the presence of red flags may indicate specific pathology and yellow, blue and black flags may be used for further consideration of possible factors that may influence non-specific pain (Table 1).

<table>
<thead>
<tr>
<th>The diagnostic triage</th>
<th>The concept of “flags”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific pain</strong></td>
<td></td>
</tr>
<tr>
<td>Specific spinal pathology</td>
<td>“Red flags”:</td>
</tr>
<tr>
<td></td>
<td>e.g. onset age &lt; 20 or &gt; 55 years, non-mechanical pain,</td>
</tr>
<tr>
<td></td>
<td>thoracic pain, previous history of carcinoma, steroids or HIV,</td>
</tr>
<tr>
<td></td>
<td>feeling unwell, weight loss, widespread neurological symptoms,</td>
</tr>
<tr>
<td></td>
<td>structural spinal deformity</td>
</tr>
<tr>
<td>Nerve root pain/radicular pain</td>
<td>“Yellow flags”:</td>
</tr>
<tr>
<td></td>
<td>psychological, social and behavioural factors</td>
</tr>
<tr>
<td></td>
<td>“Blue flags”:</td>
</tr>
<tr>
<td></td>
<td>individual perceptions and attitudes about the workplace</td>
</tr>
<tr>
<td></td>
<td>“Black flags”:</td>
</tr>
<tr>
<td></td>
<td>workplace organisational or environmental factors</td>
</tr>
</tbody>
</table>

Moreover, back and neck pain is commonly classified by duration. Mostly used are the terms, acute, sub-acute and chronic. Nachemsson and Jonsson proposed duration of 0-3 weeks for acute, 4-12 weeks for sub-acute and more than 12 weeks for chronic pain, respectively²¹. More recently Guzman and colleagues suggested the following categorization of neck pain: transitory pain lasting less than 7 days; short-duration pain
lasting 7 days or more, but less than 3 months; long-duration pain lasting 3 months or more\textsuperscript{17}. A disadvantage of the former classification is that “chronic” has a tendency to raise ideas of un-changeability.

Other duration-based definitions often used in back and neck pain literature include persistent pain (pain without periods of complete recovery or pain present on at least half the days in a 6-month period)\textsuperscript{15,17,22}, recurrent pain (two or more episodes of pain with a full recovery in between)\textsuperscript{17,18} and occasional pain (pain present on less than 30 days during the past 6 months)\textsuperscript{15}.

Many definitions are used for recovery. Kamper and colleagues found that almost every study assessing recovery from low back pain during the past 10 years did so differently\textsuperscript{23}. The duration-based definition, “At least 30 days free of pain”, has been proposed by some studies but many other definitions take disability, return to work or other dimensions into consideration\textsuperscript{18,24}.

Neck pain as a result of a motor vehicle collision or other similar mishaps is usually termed whiplash-associated disorders (WAD). The term was adapted by the Quebec Task Force on Whiplash-Associated Disorders in 1995\textsuperscript{25}. WAD includes the symptomatology related to a whiplash injury. To distinguish the injury mechanism from the symptomatology the Quebec Task Force suggested the following definition;

\begin{center}
Whiplash is an acceleration-deceleration mechanism of energy transferred to the neck. It may result from a rearend or side-impact motor vehicle collision, but can also occur during diving or other mishaps. The impact may result in bony or soft-tissue injuries (whiplash injury), which in turn may lead to a variety of clinical manifestations (Whiplash-Associated Disorders).
\end{center}

The pathophysiology of WAD is not well understood, but its aetiology likely combines physical and psychological causes. The proportion of patients with WAD in whom the symptoms are determined by a persistent lesion is still unknown when many patients are without detectable anatomic pathology. Nevertheless, Curatolo and colleagues found support for a lesion based model of WAD in their non-systematic review of basic and clinical science related to whiplash injury\textsuperscript{26}.

The Quebec Task Force also suggested a five category classification of WAD, to be used mostly in the acute phase, based on clinical signs and symptoms\textsuperscript{25}. This classification has been revised by the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and its Associated Disorders to include only four categories\textsuperscript{17};

Grade I: Neck pain and associated disorders with no signs or symptoms suggestive of major structural pathology and no or minor interference with activities of daily living.

Grade II: Neck pain and associated disorders with no signs or symptoms of major structural pathology, but major interference with activities of daily living.

Grade III: Neck pain and associated disorders with no signs or symptoms of major structural pathology, but presence of neurologic signs such as decreased deep tendon reflexes, weakness, or sensory deficits.
Grade IV: Neck pain and associated disorders with signs or symptoms of major structural pathology.

The Neck Pain Task Force suggests that this revised classification should apply to all neck pain and its associated disorders, not only WAD. Comparing the definitions above, neck pain classified as Grade I and II is equivalent to non-specific neck pain while Grade II and IV would be equivalent to specific neck pain.

2.1.3 Operational definitions in this thesis

This thesis seeks to follow the definitions and suggestions above as far as possible. The operational definitions in all four studies are based on self-reported information and, in Study III, an additional clinical examination. Throughout, we have used back pain and low back pain as synonyms indicating pain in the lower back.

2.1.3.1 Studies I and II
Persistent low back pain at baseline in Study I was defined as low back pain every day during the past six months, and the outcome, recovery, was defined as no periods of considerably disturbing back pain lasting for 7 days or more, during the latest 5-year period. In Study II, occasional low back pain at baseline was defined as low back pain, on average, up to a couple of days per months during the past 6 months, while the outcome, long duration troublesome low back pain, was defined as low back pain that decreased workability or interfered with other daily activities to some or to a high degree, on average a couple of days per week or more often during the past 6 months. According to these definitions, low back pain at baseline could, in both studies, be either specific or non-specific with or without disability while the outcomes in addition incorporate a dimension of disability, physical as well as psychological.

2.1.3.2 Study III
Participants included in Study III were considered to have non-specific back and/or neck pain lasting for two weeks or longer of the kind that brought about marked dysfunction at work and/or in leisure time. The evaluation followed the criteria for specific/non-specific pain. The primary outcomes in Study III corresponded to a clinically significant improvement in pain or disability and “totally recovered”, measured with the Chronic Pain Questionnaire (CPQ) and a modified version of the Whiplash Disability Questionnaire (WDQ).

2.1.3.3 Study IV
Included in Study IV were patients with WAD, where WAD was defined as an affirmative answer to the question “Did the accident cause neck or shoulder pain?” Patients hospitalized for more than 2 days after the injury were excluded since this indicated a more severe trauma. Therefore, according to the definitions above, our inclusion/exclusion criteria mostly, but not exclusively, incorporated participants with non-specific WAD (Grade I and II) into the study sample. The outcome, recovery from WAD, was measured with the question, “How well do you feel that you are recovering
from your injuries”? Participants were defined as recovered if they answered, “All better (cured)” or “There has been quite a bit of improvement”.

### 2.2 BACK AND NECK PAIN IN THE POPULATION

The majority of back and neck pain is non-specific, with specific pain estimated at about 10 percent but not exceeding 15 percent\(^3,14\).

Estimating prevalence and incidence in the adult population is problematic as back and neck pain has an episodic course and the problems are common even in children and adolescents\(^3\). Moreover, due to substantial heterogeneity concerning definitions and classification of back and neck pain the estimated prevalence and incidence reported in epidemiological studies vary considerably\(^3,16\).

In 2010 Hoy and colleagues published two reviews regarding the epidemiology of low back pain and neck pain\(^27,28\). They found the one-year incidence to be between 1.5 percent and 36 percent in low back pain, and between 10.4 percent and 21.3 percent in neck pain. The one year prevalence was estimated to range from 0.8 percent to 82.5 percent in low back pain, and from 4.8 percent to 79.5 percent in neck pain. In a 2003 World Health Organization (WHO) report the one year prevalence for low back pain was estimated to be between 50 to 60 percent with the highest prevalence between 18 to 34 years and the lowest from 65 years and above\(^3\). The Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders summarized the one year prevalence of neck pain to range from 30 percent to 50 percent\(^2\). It seems to be agreed that the incidence and prevalence of back pain, as well as neck pain, is higher among women than among men\(^27-30\).

Activity limiting back and neck pain are reported to be less common in the population, with a prevalence up to about 10 percent\(^2,14\). Notably though is that among people with neck pain as many as 70 percent reported the pain to be moderate to severely activity-limiting\(^28\). In a 2006 survey of chronic pain in 16 European countries 19 percent of the participants had moderate to severe pain that affected their quality of life, and almost half of those suffered from pain from lower and/or upper back\(^31\).

According to the Task Force on Neck pain the incidence of WAD in the western world has increased during the past 30 years with an annual cumulative incidence of at least 0.003 percent in North America and Western Europe\(^32\). In a Swedish back and neck pain compilation of evidence, three studies found WAD to be reported in 47, 58 and 62 percent of all traffic-related injuries in Sweden, Norway and Great Britain, respectively\(^33\). In the early 2000s the annual costs of WAD has been estimated to be somewhere between $3-14 billion in the United States and Europe\(^34,35\).

A WHO report from 2003 stated that no increase in the prevalence of spinal pain has occurred during the past 30 years but they, and other, believe that with an aging population the global prevalence may increase in the future\(^1,3\). Even though back and neck pain may be viewed as benign and self-limiting health problems the impact on the individual and the society are substantial due to for example quality of life and economy. Worldwide, The Global Burden of Disease Study 2010 ranks low back pain
as first and neck pain as fourth in “years lived with disability” where disability refers to any sort of health loss. Furthermore, from 1990 to 2010, “years lived with disability” has increased with 42 percent for back and neck pain taken together. The somewhat stable prevalence in contrast to the increase in estimated health loss may reflect a change in social security and health benefits throughout the world. Combined, back and neck pain are probably the health problem causing most sick-leave and disability pension globally. In a Swedish report from 2000 the sum of direct and indirect costs for healthcare, sick leave and disability pension due to back pain was more than three times higher than for cancer diseases.

2.2.1 The course of back and neck pain

In 1987 Waddell concluded that “low back pain by its natural history appears to be a universal, benign, self-limiting condition. Indeed, low back pain may from one perspective be regarded as normal.” Many studies reveal that the natural and clinical course of back and neck pain is favourable, and that most pain will resolve within a few weeks. For example, studies suggest that only about 2-7 percent of people will develop chronic pain. In contrast, research show that may individuals experience pain a long time after an initial pain episode and that recovery is prolonged, as for example in WAD. Carroll and colleagues found 50-75 percent of people with an initial episode of neck pain reported pain 1-5 years later, and Itz and colleagues found that 65 percent reported back pain one year after onset. One reason for this could be the typical recurrent course of back and neck pain long known. Consequently, the course of low back pain is more often viewed as a long term, recurrent condition that may have one of several trajectories. Studying the clinical course of low back pain, Dunn and colleagues classified four trajectories (clusters) named; severe-chronic, fluctuating, persistent-mild and recovering. Analogous, Tamcan and colleagues described four trajectories for the “natural” course of low back pain; severe-persistent, moderate-persistent, fluctuating and mild-persistent. There is reason to believe that the course of neck pain could be described similarly.

2.3 PROGNOSTIC RESEARCH

The term prognosis refers to the future consequence (outcome) of an already established health problem or disease, unlike the term risk, which refers to the development of a health problem or disease in healthy individuals (Figure 1).

![Figure 1. Risk and prognosis.](image)

Prognostic research is important in many aspects of secondary prevention. It could be used to improve information to affected individuals and help health-care providers in their management of patients with, for example; information, recommendations and treatment decisions. Further, public health and health-care policy-makers need
knowledge of the average prognosis, important prognostic factors and treatment effects in order to make decisions. Knowledge from prognostic research is also beneficial for the design and interpretation of trials.

At the beginning of 2013 the Prognosis research strategy group (PROGRESS) published a series of articles thoroughly summarizing the field of prognostic research \(^{49-52}\). The PROGRESS group define prognostic research as; “investigation of the relations between future outcomes (endpoints) among people with a given baseline health state (start point) in order to improve health”. The definition implies that prognostic research has to use longitudinal study design. The authors emphasize the importance of prognostic research. They also state that prognostic research needs to improve, and propose a framework of four themes: 1) fundamental prognostic research, 2) prognostic factor research, 3) prognostic model research and 4) stratified medicine research. Previously in 2010, Hayden and colleagues, using somewhat different nomenclature, described the first three themes in relation to low back pain (Figure 2) \(^{53}\). The present work uses the wording suggested by Hayden and colleagues.

**Figure 2. Three basic types of low back pain prognostic study** \(^{53}\). With permission.
2.3.1 Prognostic course studies

Prognostic course studies aim to establish the average prognosis in a population with certain characteristics. They are important as knowledge bases for the other types of prognostic study. Results from prognostic course studies are often referred to as the “course” of a health condition in different populations and could, for example, be used as reference data in a clinical trial\(^{42}\). One can also perform prognostic course studies with treatment as one of the characteristic in the population.

2.3.2 Prognostic factor studies

The PROGRESS group define a prognostic factor as; “any measure that, among people with a given health condition, is associated with a subsequent clinical outcome”\(^{50}\). Prognostic factor studies are most often observational and can be divided into two design phases (Figure 2): 1) exploration, designed to identify prognostic factors associated to the outcome, 2) confirmation, designed to test the influence (sometimes referred to as the independent effect) of a single prognostic factor by controlling for confounding. A third design phase has been proposed where, in addition to confounders, mediating and effect-modifying factors are considered to afford better “understanding of prognostic pathways”\(^{54}\). Studies I and II are based on the phase 2 design (confirmation phase).

2.3.3 Outcome prediction studies (prediction models)

In *Study IV* we developed a prediction model (referred to as “prognostic model” by the PROGRESS group), which is the first phase of the third theme (Figure 2). A prediction model is a combination of prognostic factors that can be used to predict the outcome for risk groups or individual patients. The second phase of the third theme is aimed to externally validate the developed prediction model in other sample similar to the sample used for the development\(^{49,55}\). Further, a fully validate prediction model (also referred to as a prediction tool or rule) should have been tested in clinical settings by, for example, comparing the predictive ability with the prediction made by the clinicians or by the predictive ability of earlier “golden standards”\(^{55}\). A prediction tool/rule aims to assist in identifying risk groups of patients or to help health-care providers with their prediction of an individual patient’s outcome. Prediction tools/rules can also be used as base for stratified medicine research, the fourth theme proposed by the PROGRESS group\(^{52}\).

2.4 PROGNOSTIC RESEARCH RELATED TO THIS THESIS

2.4.1 Treatment (prognostic course studies)

There are many different treatment alternatives for back and neck pain ranging from “advice to stay active” to surgery. Treatments aim to target the different dimensions of back and neck pain. For example, medication to reduce the symptom, behavioural therapy to enhance the ability to cope with pain and manual therapy, used to decrease pain by restoring “normal” function. The evidence base for treatment consists of a growing number of clinical trials and systematic reviews from which results are transformed into recommendations in clinical guidelines\(^{11,12,56}\).
Advice to stay active and support on how to cope with pain was the treatment compared to naprapathic manual therapy in Study III. At the time of the data collection for Study III, there was evidence for advice and support from health-care providers aiming to help the patient to understand that recovery is enhanced by staying active and living as normally as possible\cite{21}. These recommendations still hold as important and essential and are often included in clinical guidelines, especially for acute and sub-acute pain\cite{11,12}.

Manual therapy is a common treatment alternative for patients with non-specific back or neck pain. The aim of manual therapy is to decrease pain and disability by targeting the neuromusculoskeletal system to restore normal function. Manual therapy for back and neck pain commonly includes manipulation and/or mobilisation of spinal joints and massage, and sometimes with additional muscle-stretching. Several health-care providers use manual therapy. The most common in Sweden are physiotherapists, naprapaths, chiropractors and osteopaths. Swedish physiotherapists using manual therapy are often specifically trained in orthopaedic manual (manipulative) therapy (OMT). The manual therapy treatment in this work was performed by naprapaths.

Naprapathic manual therapy has its origin in the USA and was introduced in Sweden in 1970. Since 1994 naprapaths have been a part of the Swedish health-care system and licenced by the National Board of Health and Welfare for treating patients with musculoskeletal pain and pain-related disability. Naprapaths in Sweden have a four-year full-time education with an additional year within the licenced health care system. So far, in addition to Study III, there are only two published studies on naprapathic manual therapy. Using the same trial as in Study III, Skillgate and colleagues found naprapathic manual therapy to reduce pain and disability in the short-term (7 and 12 weeks) compared to evidence based care by physicians\cite{57}. Lilje and colleagues found naprapathic manual therapy to give significantly decreased pain and improved function compared to conventional orthopaedic care for outpatients with musculoskeletal problems\cite{58}.

The most well-studied techniques used in manual therapy are spinal manipulation, spinal mobilisation and massage. Results from recent systematic reviews are somewhat inconclusive but as a summary, massage and manual therapy seem to have beneficial short-term effects on sub-acute and chronic low back pain and disability while the effect on neck pain is less clear. The long-term effects are unclear and need to be studied more\cite{56,59-65}. Bronfort and colleagues summarised the scientific evidence for the effectiveness of manual treatment in a variety of health conditions\cite{66}. Their findings regarding conditions studied in the present work are presented below (Table 2).

<table>
<thead>
<tr>
<th>Spinal conditions</th>
<th>Intervention</th>
<th>Evidence a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute LBP b</td>
<td>Spinal manipulation/mobilisation</td>
<td>Positive</td>
</tr>
<tr>
<td>Chronic LBP</td>
<td>Spinal manipulation/mobilisation</td>
<td>Positive</td>
</tr>
<tr>
<td>Chronic LBP</td>
<td>Massage</td>
<td>Positive</td>
</tr>
<tr>
<td>Acute/sub-acute NP c</td>
<td>Thoracic spinal manipulation/ mobilisation</td>
<td>Positive</td>
</tr>
<tr>
<td>Acute WAD</td>
<td>Mobilisation with exercise</td>
<td>Positive</td>
</tr>
<tr>
<td>Chronic NP</td>
<td>Spinal manipulation/mobilisation with exercise</td>
<td>Positive</td>
</tr>
<tr>
<td>NP of any duration</td>
<td>Cervical spinal manipulation/mobilisation</td>
<td>Favourable</td>
</tr>
<tr>
<td>Chronic NP</td>
<td>Massage</td>
<td>Positive</td>
</tr>
</tbody>
</table>

a Moderate and high quality positive evidence; regarded as effective and could be recommended to patients as an effective treatment. Inconclusive, but favourable evidence; not supported by evidence to be effective, but may be recommended to patients as an option in absence of an effective treatment. b Low back pain. c Neck pain.

2.4.2 Prognostic factor studies

Evidence shows that psychosocial factors are associated to poor outcome of low back pain (LBP)27,67. In a comprehensive “review of reviews” from 2009, Hayden and colleagues found nine factors consistently reported in the reviews to be associated to a poor outcome in acute and sub-acute low back pain: older age, poor general health, increased psychological or psychosocial stress, poor relations with colleagues, physically heavy work, worse functional disability, sciatica, and the presence of worker’s compensation68. The authors observed differences in the systematic review methods, and in the prognostic studies included in the reviews which may explain the few factors consistently reported. In a systematic review of prospective cohort studies Kent and Keating likewise concluded that differences in methodology may be the reason why they could not find evidence that any factor was of prognostic value for poor recovery in recent-onset LBP69. Iles and colleagues reviewed prognostic studies including a measure of recovery expectation and activity limitation outcomes (e.g. return to work) for non-chronic LBP70. They found a strong association between recovery expectation and the outcomes. A 2012 review of prognostic factors for recovery in chronic non-specific low back pain either found no association to recovery or conflicting results for all factors studied71. The authors considered the included studies to be limited by methodological weakness.

Hayden and colleagues found conflicting evidence for BMI to influence the prognosis of LBP68. They did not include leisure physical activity in their report. However, one of the reviews included in the report, concluded that being active in physical fitness or sports does not seem to be a prognostic factor for the duration of sick leave due to LBP in patients sick-listed with acute LBP72. Based on seven prospective studies on patients with LBP, Hendrick and colleagues found moderate evidence for no association between day-to-day physical activity (occupational, sports and leisure activity) and LBP outcomes73.
Knowledge concerning influence of combinations of lifestyle factors, as a proxy for healthy lifestyle behaviour, on the prognosis of low back pain is lacking. Nevertheless, a combination of healthy lifestyle factors such as non-smoking, physical activity, healthy diet and moderate or no alcohol use seems to influence the risk and the prognosis in several diseases (e.g. cancer, type 2 diabetes mellitus and cardiovascular disease) as well as mortality\textsuperscript{74-77}.

2.4.3 Prediction models (outcome prediction studies)

When searching the literature we found only three reports developing a prediction model for WAD, and only one presenting an external validation of such a prediction model\textsuperscript{78-81}. Due to different settings of the studies, and differences in definitions of WAD, potential prognostic factors assessed, statistical methods used and length of follow-up, there is no point in presenting the results of these studies in this work. Notable though, is that they all included older age and factors related to the initial neck pain in their models.

2.5 ASSOCIATION AND CAUSALITY

When studying prognostic factors (exposures) we consider whether the outcome of a health problem, e.g. recovery from low back pain, occurs more often or less often, alternatively sooner or later, in the exposed group than in the unexposed group. If either is true there is an association between the prognostic factor and the outcome. To infer that the prognostic factor (X) may cause the outcome (Y), certain criteria must be fulfilled; the prognostic factor must be present before the outcome, the association must not depend on another factor that associates both to the exposure and to the outcome (confounding) and the association should not be affected by other systematic errors such as misclassification and selection bias\textsuperscript{82}.

![Figure 3. Methodological considerations when studying association and causality between a prognostic factor (X), and the outcome of the health problem (Y). 3a: confounding. 3b: mediation. 3c: effect measure modification.](image)

A confounding factor (Figure 3a) is a factor that affects the outcome but does not take part in the same casual pathway as the prognostic factor in question. Thus a factor that is an intermediate step in the pathway between the prognostic factor and the outcome, referred to as a mediator (Figure 3b), is not to be considered as a confounder\textsuperscript{82}. A related concept is that of a collider which is a factor influenced by the prognostic factor as well as the outcome, and therefore partly a consequence of, rather than the cause of, the outcome (not shown in figure)\textsuperscript{83}.

There are many ways to deal with confounding\textsuperscript{82}. In Studies I and II we have adjusted for confounding in the multivariable analyses. The randomisation in Study III is a well-
known principle to avoid the risk of confounding, at least if the sample size is large\textsuperscript{82}. In smaller samples and/or if randomization is not completely successful, factors associated to exposure and outcome may be unevenly distributed among exposed and unexposed groups. In such cases adjusting for confounding may be necessary even in a randomized controlled trial. In multivariable analysis it is important to decide which factors are confounders, mediators or colliders. This is because adjusting for confounding is essential to minimize bias. However, adjusting for mediators or colliders may instead introduce bias\textsuperscript{82,83}.

Another issue when studying association and causality is “interaction”, a term used for many concepts. In epidemiology two concepts for “interaction” are commonly used; effect measure modification and sufficient-cause interaction\textsuperscript{82}. The former (also known as statistical interaction) refers to the situation where a measure of effect from one factor is changed by values of another factor (Figure 3c). It has been described as “departure from additivity of effects on the chosen outcome scale”. This implies that effect measure modification is scale-dependent, i.e. two factors can interact on the relative scale but not on the absolute scale and vice versa. In Studies II and IV effect measure modification on the relative scale was considered in the analyses.

Sufficient-cause interaction (sometimes referred to as biological interaction) between two contributory causes occurs when the effect of one factor depends on the presence of another factor on the absolute scale\textsuperscript{82}. Positive sufficient-cause interaction (also known as synergism) arises when the joint effect of the two factors is greater than the sum of their independent absolute effect. Negative sufficient-cause interaction (antagonism) arises when the joint effect is smaller than the added independent effects.

2.6 RATIONALE

Back and neck pain is very common and greatly affects the individual as well as the community. The condition is strongly linked to quality of life and is among the most common causes of sick leave and disability pension. This makes research regarding its prognosis important.

The studies reported in this thesis are intended to deepen the knowledge of factors that may influence the prognosis of back and neck pain and how these factors can be used. This may guide health-care providers working with secondary prevention of back and neck pain, and contribute to health-policy recommendations and guidelines. Increased knowledge may also facilitate the identification of individuals that have a better chance of recovery as well as those at higher risk of long-term pain. Further, it may help to find and apply interventions to enhance the possibility of recovery from back and neck pain.
3 AIMS

3.1 MAIN AIM

The main aim of the studies presented in this thesis was to study prognostic factors, including manual therapy, for back and neck pain.

3.2 SPECIFIC AIMS

*Study I:* to investigate the influence of regular leisure time physical activity and BMI on recovery from persistent back pain among men and women in a general population

*Study II:* to explore the influence of healthy lifestyle behaviour on the prognosis of occasional low back pain among men and women in a general population

*Study III:* to study the long-term effects (26 and 52 weeks respectively) of naprapathic manual therapy for patients with non-specific back and/or neck pain

*Study IV:* to develop a predictive model for the recovery from WAD in a sample of patients who consulted physiotherapy within six weeks of their injury
4 MATERIAL AND METHODS

4.1 STUDY DESIGN AND SOURCE POPULATION

4.1.1 The Stockholm Public Health Cohort (Studies I and II)

Studies I and II were based on data from the Stockholm Public Health Cohort (SPHC)\(^84\). This population-based cohort was set up within the framework of Stockholm County Council public-health surveys. It was managed by Statistics Sweden together with researchers from the Department of Public Health Sciences at Karolinska Institutet, Stockholm. So far, baseline and follow-up surveys from 2002 to 2010 are included in the cohort.

The SPHC includes three sub-samples, of which the sub-sample with baseline in 2002 was used in Study I and the one with baseline in 2006 was used in Study II (Table 3).

Table 3. Baseline sample size and number of responders at recruitment in the Stockholm Public Health Cohort.

<table>
<thead>
<tr>
<th>Questionnaire-based waves of data collection</th>
<th>Sub-cohort by year of recruitment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2002</td>
</tr>
<tr>
<td>Recruitment in 2002</td>
<td>Baseline sample</td>
</tr>
<tr>
<td></td>
<td>Responders</td>
</tr>
<tr>
<td>Recruitment in 2006</td>
<td>Baseline sample</td>
</tr>
<tr>
<td></td>
<td>Responders</td>
</tr>
<tr>
<td>Follow-up in 2007</td>
<td>Responders</td>
</tr>
<tr>
<td>Recruitment and</td>
<td>Baseline sample</td>
</tr>
<tr>
<td>follow-up in 2010</td>
<td>Responders</td>
</tr>
</tbody>
</table>

\(^1\) Baseline sample and responders in 2002 and at follow-up in 2007 included in Study I.

\(^2\) Baseline sample and responders in 2006 and at follow-up in 2010 included in Study II.

At baselines (2002 and 2006) the source populations were residents, 18 to 84 years old, of Stockholm County, Sweden, an urban region with approximately 1.4 million adults at the time of sampling. Approximately 50 000 individuals were randomly selected, stratified by region (municipality and urban districts), to be included in each sub-cohort. The sample size was estimated to enable enough power for studying the prevalence of important health factors. Compared to consensus data from Stockholm County, the SPHC participants were more likely to be female, be born in Sweden, have higher education and income, and be more than 45 years old\(^84\).

4.1.2 The Swedish BJÖRN trial (Study III)

Study III was based on a pragmatic randomized controlled trial, “the Swedish BJÖRN trial”, which was performed in Stockholm, Sweden between March 2005 and October 2006. The main source population was approximately 40,000 employees at two public companies in Stockholm. Participants were recruited by advertising and potential participants were asked to contact the study administration if they fulfilled the inclusion criteria, which 522 persons did.
4.1.3 The Saskatchewan Government Insurance study (Study IV)

This closed cohort study includes information from the Saskatchewan Government Insurance study (SGI). The SGI is a population-based inception cohort study of 8634 individuals injured in traffic collisions, between 1997 and 1999, in the province of Saskatchewan, Canada. Members of the SGI cohort were residents of Saskatchewan, 18 years of age and older, who reported their collision to the SGI. In addition, individuals who consulted a health-care provider for their injury were also included in the cohort, because providers were mandated to inform SGI of all traffic injuries. Excluded were individuals with serious unassociated illness, individuals who did not understand English and those with Workers’ Compensation claims since those were covered under a different system.

4.2 DATA COLLECTION AND PARTICIPANTS

4.2.1 Studies I and II

The baseline and the follow-up questionnaires in the SPHC comprised self-reported information on lifestyle, demographic and socioeconomic characteristics, physical and psychological health and work-related factors. The self-reported data were supplemented with information from Swedish national and regional registers.

The study sample in Study I consisted of the 1836 participants who reported persistent low back pain (PBP) at baseline in 2002 and answered the follow-up questionnaire in 2007. In Study II the study sample comprised the 8994 participants with occasional low back pain at baseline in 2006, answering the follow-up questionnaire in 2010. Participants without information on the exposure and/or the outcome were excluded from both studies (Figure 4).
Figure 4. Flowchart describing the inclusion process for the study sample in Study I and Study II.


PBP (Study I) and occasional low back pain (Study II) at baseline were defined using a modified question from the Standardized Nordic Questionnaire:

“During the previous six months, have you experienced pain in your lower back? If you have experienced pain on several occasions, try to estimate an average.”

Response alternatives:

“No”

“Yes, a few days in the past six months”

“Yes, a few days per month”

“Yes, a few days per week”

“Yes, every day”

Participants in Study I answering “Yes, every day” were considered to be suffering from PBP and participants in Study II answering “Yes, a few days in the past six months” or “Yes, a few days per month” were considered to have occasional low back pain.

4.2.2 Study III

The inclusion criteria for participants in Study III were “having back and /or neck pain, now and the previous two weeks or longer, of the kind that brought about marked
dysfunction at work and/or in leisure time”. Participants who met the inclusion criteria were asked to contact the study administration where the first-step exclusions were made (symptoms too mild, pregnancy, specific diagnoses such as acute slipped disc or spinal stenosis, inability to understand Swedish, and recent visits to a manual therapist with the exception for massage). Further, they were scheduled for an appointment at the study centre where they gave their informed consent and answered an extensive self-administered questionnaire. Next, one of four experienced physicians performed a medical examination (about 20 minutes), made a diagnosis, and prescribed medication if necessary. Further exclusions were made based on the following exclusion criteria: too mild symptoms (the physician’s subjective opinion based on the estimated pain and disability in the questionnaires filled in before the examination, and the results of the medical history and physical examination), evidence-based advice during the past month, surgery in the painful area, acute disc herniation, spondylolisthesis, stenosis or “red flags”\textsuperscript{21}. Participants finally included were considered to have non-specific back and/or neck pain.

The 409 participants included (patients) were assigned to one of two groups by randomization. An assistant not involved in the project prepared 500 opaque, sequentially-numbered sealed envelopes with cards numbered 1 or 2 (randomized by a computer), indicating the two interventions. The patients were sequentially numbered in the order they came to the study centre and received the assigned envelope with the corresponding number. The unmasking was performed by the physician after the medical examination, so that the assistant, the physician and the patient were all blind to the group assignment until after all patient baseline data had been collected.

Self-administered questionnaires, web-based or postal, were collected three, seven, twelve, twenty-six and 52 weeks after the inclusion. Figure 5 describes the flow of participants and attrition through each stage of the trial.

![Figure 5. Flowchart showing the progress of participants through the trial.](image-url)
4.2.3 Study IV

Baseline data were available from self-administered SGI forms filled in by the participants, supplemented with questions constructed for research purposes. The SGI form included questions on socio-demographics, injury-related pain intensity and location, activity limitations, comorbid health conditions, pre- and post-injury general health, health care provision, depressive symptomatology and work status. Follow-up data were collected in computer-assisted telephone interviews at six weeks and three, six, nine and 12 months. The data used in Study IV were from the six-weeks and the three- and 6-month follow-ups.

The baseline study sample of 680 subjects included participants with WAD who consulted a physiotherapist (PT) between the date of the collision and the date for reporting to SGI. Participants with WAD (patients) were defined as those who answered “Yes” to the question “Did the accident cause neck or shoulder pain?” Excluded patients were those not in a motor vehicle when injured and those reporting their injury to SGI more than 42 days after the collision. Further, patients hospitalized for more than 2 days after the injury were excluded since this indicated a more severe trauma (Figure 6).

**Figure 6. Inclusion process and progress of patients in the study sample.**
4.3 MEASUREMENTS AND INTERVENTIONS

4.3.1 Studies I and II

4.3.1.1 Prognostic factors

Potential prognostic factors (exposures) were measured at baseline in 2002 for Study I and in 2006 for Study II. The prognostic factors considered were regular leisure time physical activity and the body mass index in Study I and “healthy lifestyle behaviour” in Study II.

Regular leisure time physical activity (PA) in Study I
PA was categorized into four levels using the answers based on the question; “During the previous 12 months, how physically active have you been during leisure time? If your activity differs between e.g. summer and winter, please estimate the average activity”. The PA levels were:

Sedentary; “You mostly devote yourself to reading, TV, movies or other sedentary activity during leisure time. You walk, cycle or are active in other ways less than 2 hours a week”.

Low; “You walk, cycle or are active in other ways at least 2 hours a week, mostly without sweating. Also include walking or cycling to and from work, Sunday walks, ordinary gardening, fishing, table tennis and bowling”.

Moderate; “You are physically active regularly, 1-2 times a week at least 30 minutes each session with running, swimming, tennis, badminton or other activity that makes you sweat”.

High; “You devote yourself to e.g. running, swimming, tennis, badminton, aerobic exercise or similar on average at least 3 times a week, each session lasting at least 30 minutes”.

The question assessing PA is useful for categorizing adults into different levels of PA, based on the physical activity in the different groups as measured by accelerometer88. Moreover, Leijon and colleagues found the PA measure to have moderate criterion validity compared to a structured personal interview in a working population with musculoskeletal complaints. They found no sex differences regarding the validity of the question89. Orsini and colleagues found the question to have acceptable validity among older women90.

The body mass index (BMI) in Study I
BMI (kg/m2) was calculated using self-reported weight and height and categorized into normal weight (BMI < 25) and overweight (BMI ≥ 25). The normal-weight category included 31 underweight participants (BMI < 18.5) and the overweight category included 270 obese participants (BMI ≥ 30)91. BMI is widely used and recommended as a baseline exposure measurement in cohort studies with BP as outcome92. Still there is a potential risk of misclassification when using self-reported weight and height to calculate BMI as weight tends to be underestimated and height overestimated93. Further, BMI cut-offs have low sensitivity to identifying adiposity (high body fat percentage) in general and overestimating adiposity in trained subjects94,95.
Healthy lifestyle behaviour (HLB) in Study II

Four binary healthy lifestyle factors were constructed where cut-offs (healthy/not healthy) were set in accordance with recommendations for a health-enhancing lifestyle made by Swedish authorities and the WHO\textsuperscript{96-99}. HLB was a combination of these binary factors and was categorised into five levels according to the number of healthy lifestyle factors included, i.e. from none to four (HLB0 to HLB4). A healthy lifestyle behaviour with regard to each of the healthy lifestyle factors considered was defined by: non-smoking, no risk consumption of alcohol (\(\leq 168\) g 100\% alcohol/week for men and \(\leq 108\) g 100\% alcohol/week for women, and consuming alcohol corresponding to \(\approx\) half a bottle of spirits on the same occasion less than once a month), recommended level of leisure physical activity (at least 150 minutes at moderate intensity or 75 minutes at high intensity per week or a combination of these activities), and recommended consumption of fruit and vegetables (\(\geq\) a total of 4 servings of fruit and vegetables/day, equal to \(\approx 400\) g/day). Leisure physical activity was measured with the same question as in Study I in combination with a question about the number of days per average week with at least 30 minutes of physical activity that made the responder grow warm. The questions regarding consumption (number of servings per day) of fruit and vegetables had acceptable reliability and validity and the “period-specific normal week” method used to measure weekly alcohol consumption is recommended by Rommelsjö and colleagues\textsuperscript{100,101}. For a detailed description of the construction of HLB, see the Appendix in the Study.

4.3.1.2 Potential confounders

Information on potential confounders was gathered at baseline in each study. Potential confounders were chosen based on prior research regarding the prognosis of spinal pain, empirical relevance and the availability in the questionnaires\textsuperscript{44,68,102}. Seventeen potential confounders were assessed in Study I as well as in Study II. Study I: civil status, country of birth, socioeconomic status, current occupation, smoking habits, alcohol consumption, neck pain, chronic illness, psychological wellbeing, emotional and instrumental social support, time spent doing housework, physical workload, sick leave and three psychosocial work-related factors. Study II; civil status, country of birth, socioeconomic status, disposable income, financial stress, education, neck pain, chronic illness, pain from hip, thigh or knee, headache or migraine, rheumatoid arthritis, living alone, living with children, hours of sleep a typical night during the working week, frequency of stress, psychological wellbeing and emotional social support.

The questions used for measuring potential confounders have, since 1975, been used in several Swedish national and local public health surveys. They have on several occasions been tested and improved by Statistics Sweden’s test centre and several questions have proved to have acceptable psychometric properties. In addition, information on civil status, country of birth, education, socioeconomic status and disposable income were collected from Swedish national registers known to have high quality. Psychological wellbeing was measured with the twelve-item general health questionnaire (GHQ-12) regarded as valid to detect cases with reduced psychological wellbeing\textsuperscript{103,104}. The social- support questions (emotional and instrumental) originate from an instrument (ISSI) described by McDowell and tested,
as a short form (SS-13), by Undén and Orth-Gomer in a population of Swedish industrial workers.\textsuperscript{104,105}

4.3.1.3 Collider and effect measure modification

Self-rated health (SRH) was not treated as a potential confounder. The reason was that poor SRH is shown to be a consequence of, rather than the cause of, pain in long-duration pain conditions such as the PBP in Study \textsuperscript{106}. If this is true, SRH could be regarded as a collider in Study I and should not be adjusted for in the analyses.\textsuperscript{82,83}

In Study II relevant effect measure modification between possible confounding variables and the exposure (HLB) were considered. The variables considered were: age, socioeconomic status, disposable income, financial stress, education, neck pain, chronic illness, pain from hip/thigh or knee, rheumatoid arthritis, living with children, hours of sleep a typical night during the working week, frequency of stress and psychological wellbeing.

4.3.1.4 Outcomes

Study I

The follow-up questionnaire in 2007 supplied information about the outcome “recovery from PBP” defined as reporting: “No periods of considerably disturbing back pain lasting for 7 days or more, during the latest 5-year period.”

The outcome was constructed using a combination of two questions: “During the last 5-year period have you had back pain for at least 3 consecutive months that has disturbed you considerably?”, and “During the last 5-year period have you had back pain, in at least 7 consecutive days but less than 3 consecutive months, that has disturbed you considerably?”

We defined participants as recovered if they answered “No” to both questions.

Study II

Data on the outcome “long duration troublesome low back pain” (LTLBP) was collected from the follow-up questionnaire in 2010 and was assessed with a two-part question:

a) “Have you had any pain in your lower back in the past 6 months? If you have experienced pain on several occasions, try to estimate an average.”

Response alternatives:
“No.”
“Yes, a few days per month or more seldom.”
“Yes, a few days per week or more often.”

b) “If Yes: Has these troubles resulted in a decreased workability or interference with other daily activities?”

Response alternatives:
“Yes, to a high degree.”
“Yes, to some degree.”
“Not at all.”

Participants answering “Yes, a few days per week or more often” and “Yes, to a high degree” or “Yes, to some degree” were defined as having LTLBP. The question is
modified from the Standardized Nordic Questionnaire and the second part incorporates a dimension of disability suggested to be of importance when defining LBP.

### 4.3.2 Study III

#### 4.3.2.1 Interventions

**Naprapathic Manual Therapy (Index Group)**

For patients in the Index Group, one of eight experienced, licensed naprapaths was contacted for a first appointment within a week. A maximum of six treatments was given within six weeks at the naprapaths’ private clinics, and a combination of napraphathic manual techniques (such as spinal manipulation/mobilization, massage and stretching) was given adapted to the patient’s condition. Preventive and rehabilitating advice on physical activity and ergonomics appropriate for the patient were often given. Each appointment lasted for about 45 minutes.

**Evidence-Based Care Provided by a Physician (Control Group)**

For patients in the Control Group, evidence-based care was given by one of four experienced physicians. Evidence-based care in this study is defined as support and advice on staying active and on pain coping strategies, according to guidelines and evidence-based reviews. This was given in conjunction with the medical examination (an additional 15 minutes) at the study centre. The aim was to empower the patient with understanding of the importance of staying active and living as normal a life as possible, including work and physical activities. The care also aimed to improve the patient’s pain-coping strategies. Advice on exercise was general and adapted to the patient’s condition. A booklet with examples of exercises and general information on back and neck pain was provided and additional consultation was offered if necessary.

The treatments in both groups were conformed to the patients’ condition, but standardized as far as possible. The naprapaths were told only to use techniques they had learned at the education centre. The content in the evidence-based advice and support was carefully discussed in groups with the physicians in order to make the care reliable.

#### 4.3.2.2 Outcomes

The primary outcomes pain and disability were measured with the Chronic Pain Questionnaire (CPQ) with three items on pain and three on disability, with a numerical 11-point scale. The wording of the questions was changed to concern the previous four weeks instead of the previous six months. A pain score was constructed from the mean of the three pain items and a disability score from the mean of the three disability items. Disability was also measured in a more detailed way with a modified version of the Whiplash Disability Questionnaire (WDQ), with 13 items, each with a numerical 11-point scale. In the context of Study IV the items were modified by replacing the word “whiplash” with “back pain” or “neck pain.” This disability score was the mean of the 13 items.

Four dichotomized outcomes were defined in advance based on what is believed to correspond to a clinically significant improvement:
1) improvement in pain: at least a two-step decrease in pain score (CPQ), compared to baseline.
2) improvement in disability I: at least a one-step decrease in disability score (CPQ), compared to baseline.
3) improvement in disability II: at least a one-step decrease in disability score (WDQ), compared to baseline.
4) totally recovered: a pain score of less than or equal to 1 and a disability score equal to 0 (CPQ).

The secondary outcome was health status measured with the Medical Outcomes Study Short Form-36 Health Survey (SF-36) after 26 and 52 weeks\textsuperscript{120,121}.

4.3.3 Study IV

4.3.3.1 Prognostic factors for the prediction model

The baseline questionnaire comprised information on potential prognostic factors to be included in the prediction model. The factors were selected based on literature, empirical relevance and availability in the questionnaire\textsuperscript{45,122-124}. The selected factors were grouped into three domains that represented the sequential approach of the medical history obtained by a physiotherapist during an initial consultation. The rationale for using this approach was to determine whether expanding the breadth of information collected during the medical history improved the physiotherapists’ ability to predict recovery.

The three domains were:

- **Sociodemographics**: Age, sex, civil status, education and work status.
- **Collision, symptoms, comorbidity and health care**: Number of days to reporting the collision, average pain intensity in neck/shoulder and low back and/or of headache, pain other than neck and back pain, feeling of numbness/tingling or pain in arms or hands, pain when moving the neck, reduced neck movement, sleeping problems, musculoskeletal problems and headache before the collision, restrictions in daily home or leisure-time activities and number of visits to physiotherapists and medical doctors since collision.
- **General health and psychology**: Recovery expectations, current and pre-collision general health, anxiety or worry caused by the trauma and depressive mood.

Pain intensity was measured on an 11-point numeric rating scale (NRS: 0-10) where “0” means “no” pain and “10” means “pain as bad as it could be”. The NRS is a reliable and valid method for assessing pain in various patient categories\textsuperscript{22,125}. Neck/shoulder pain intensity (from now referred to as neck pain intensity) was categorised into no (0), mild (1-4), moderate (5-7) and severe (8-10) pain according to suggestions from Fejer and colleagues\textsuperscript{117}. The same categories were used for headache and back pain intensity. Current and pre-collision general health was measured using an item from the Short Form 36; “In general, how would you say your health is now?” and “How was your health the month before the collision?”\textsuperscript{126}. Depressive mood was assessed with the 20-item Centre for Epidemiological Studies Depression Scale (CES-D). The scale is considered reliable and valid for measuring depressive symptomatology in both healthy and ill populations. The cut-off point of 16 with scores of 16 or above indicating depressed mood was used\textsuperscript{127}.
4.3.3.2 Effect measure modification

Effect measure modification between neck pain intensity and recovery expectations as well as depressive mood was considered clinically relevant to test.

4.3.3.3 Outcome

The self-reported outcome “recovery from WAD” was collected from the follow-up interviews six weeks, three months and six months after the collision. Recovery was measured with the global recovery question: “How well do you feel that you are recovering from your injuries”? Patients answering “All better (cured)” or “There has been quite a bit of improvement” were defined as recovered.

The question has adequate reliability and validity for use in epidemiological studies of WAD128,129. Time to recovery, used in the “Time to event analyses”, was days between the collision and the first follow-up where the patients were defined as recovered.

4.4 STATISTICS AND COMPARISONS

I performed all the statistical analyses in Studies I, II and IV, supported by my supervisors, co-authors and experienced biostatisticians at the Karolinska Institutet. The analyses in Study III were performed by a statistician blinded to the allocation of patients into the treatment groups.

All p-values were two-sided and the analyses were completed using SAS® version 9.1 to 9.3 (Cary, NC: SAS Institute) and STATA/IC® version 12.1 (Stata Corp LP, USA).

Table 4. Overview of the main statistical methods used in the present work.

<table>
<thead>
<tr>
<th>Statistical methods</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Log-binomial regression</td>
<td>●</td>
</tr>
<tr>
<td>General estimating equation (GEE)</td>
<td></td>
</tr>
<tr>
<td>Time to event analyses</td>
<td></td>
</tr>
<tr>
<td>Kaplan-Meier</td>
<td></td>
</tr>
<tr>
<td>Cox proportional regression</td>
<td></td>
</tr>
<tr>
<td>Concordant statistics (c-index)</td>
<td></td>
</tr>
<tr>
<td>Bootstrap method</td>
<td></td>
</tr>
</tbody>
</table>

The effect measures used in Studies I, II and III were risk ratios (RR) and risk differences (RD) with corresponding 95 percent confidence intervals (95% CI). The RR represents the relative effect between exposed and unexposed groups and is calculated by dividing the incidence proportion (or cumulative incidence) of the exposed with the incidence proportion of the unexposed;

\[
RR = \frac{Ae / Ne}{Aue / Nue}
\]
Where \(Ae\) and \(Aue\) are the number of subjects experiencing the outcome among exposed and unexposed, and \(Ne\) and \(Nue\) corresponds to the number of subjects among the exposed and unexposed at baseline, respectively. Risk ratios of 1.0 mean that the exposure has no effect on the outcome. An RR above 1.0 indicates that the exposed group has an increased risk for the outcome and an RR below 1.0 means that the exposed group has a decreased risk.

The RD reflects the absolute effect of the exposure on the outcome and is measured by the difference between the incidence proportion in the exposed and the unexposed group:

\[
RD = \frac{Ae}{Ne} - \frac{Aue}{Nue}
\]

In Studies I and II we have used the log-binomial regression to estimate crude and adjusted RRs and RDs with corresponding 95% CI. The log-binomial regression is a generalized linear model with a binomial random component and with the log as link function\(^{82}\). The model can be used for RR and RD estimations of a dichotomous outcome, and models the probability of the outcome given the exposure and other covariates (confounders) included in the model. It has been criticized for producing adjusted RR with CI that may be narrower than what is true. This bias McNutt and colleagues found to be minor\(^{130}\). Furthermore, sometimes the log-binomial regression does not converge. Despite these drawbacks, several authors recommend the model for use in cohort studies with a common outcome\(^{130,131}\).

The General Estimating Equation (GEE) method, also referred to as the marginal logistic model, provides a natural way of extending standard regression analysis to longitudinal data\(^{132}\). The GEE needs no distributional assumptions and takes the correlation between the repeated measurements into account. The method was used to analyse the effect on dichotomized pain and disability over the total follow-up time in Study III.

In time to event analyses we need to know the event (outcome), the time to event and the “censoring time”. In Study IV the event was recovery from WAD and the time to event was the time from inclusion to the recovery. The censoring time for a subject in Study IV was the time from inclusion to the end of follow-up if no event was observed, or the mid-point between the last completed follow-up and the next follow-up if the subject was lost to follow-up. The Kaplan-Meier estimator was used to estimate the proportion of patients recovering from WAD at a certain times after inclusion into the study (see graph in Study IV). It also provided an estimate of the median time to recovery. Further, the Cox proportional hazard regression was used to estimate the associations between prognostic factors included in the prediction models and recovery from WAD\(^{133}\). The relative effect was reported as beta coefficients (\(\beta\)) with standard error (SE) and hazard rate ratios (HRR) with 95% CI. The Cox regression is a semi-parametric model where the baseline hazard function does not have to be specified, but where parametric assumptions about the effects of the prognostic factor of the hazard function have to be made. This assumes that the hazard rates predicted by the model are proportional over time (i.e the ratio is constant over the follow-up time)\(^{133}\).
In Study IV, the predictive ability of the prediction models was measured with the Harrell’s concordant statistics (c-index) with 95% CI. A model with a c-index of 0.5 has no predictive ability while a c-index of 1.0 indicates perfect predictive ability\textsuperscript{134}. Another common concordant statistics is the receiver operating characteristic curve (ROC curve) where the area under the curve, possible to estimate alongside a logistic regression, is equivalent to the c-index estimated in conjunction with a Cox model\textsuperscript{135}. Concordant statistics have been criticised to be insensitive and other predictive measures have been introduced recently, for example the net reclassification improvement (NRI) and integrated discrimination improvement (IDI)\textsuperscript{136,137}. Nevertheless, Kamper and colleagues suggests the c-index to be used in prediction studies of WAD\textsuperscript{138}. The authors, in agreement with Harrell, also recommend the use of the bootstrap method to internally validate a prediction model\textsuperscript{134,138}.

### 4.4.1 Studies I and II

#### 4.4.1.1 Main analyses

Log-binomial regression models were used to estimate the association between the exposures and the outcome, analysing men and women separately. First crude regression models, only including the exposures and the outcome, were built. To determine the role of a potential confounder factor we included them, one at a time, in the crude model. If the inclusion changed the crude estimates by 10 percent or more, the factor was considered to be a confounder and was included in the final adjusted model; the change in estimate (CE) method\textsuperscript{82,139-142}. All final models were adjusted for age (continuous age in Study I and age categorized into 10-year intervals in Study II).

#### 4.4.1.2 Additional analyses

**Study I:**
The same adjusted models as in the main analyses were used stratified by poor and good self-rated health (SRH) using the strata as an approximation for more or less severe PBP at baseline. The rational for these analyses was to tell whether the effect of PA and/or BMI was different in participants with more or less severe PBP at baseline, as data on pain intensity was missing.

**Study II:**
A likelihood ratio test was used to assess effect measure modification between the exposure and possible confounders as well as confounders included in the final models\textsuperscript{135}. If including a product term representing the effect measure modification significantly improved the model it should be used in the final analysis. Further, a test for trend in the associations between the exposure and the outcome was performed, and a Chi-square test to assess whether the overall adjusted risk differed between men and women\textsuperscript{135}.

*Additional analyses not reported in the manuscript*

Using the same crude and adjusted regressions as in the main analyses, additional analyses with a different reference group were made for women. Here the reference
group was women with no healthy lifestyle factor and women with one healthy lifestyle factor (HLB0/HLB1) instead of only HLB0 as in the main analyses. The effect of attrition was assessed, using Chi-square tests, by comparing the distribution of the four healthy lifestyle factors included in the exposure, HLB, in participants who were lost to follow-up to the distribution in the study sample.

4.4.2 Study III

4.4.2.1 Main analyses

All the analyses were performed using an “intention to treat” principle\(^\text{143}\). Changes in mean scores at follow-up compared to baseline and differences in changes between groups were calculated with a t-test. To compare the groups regarding the dichotomized outcomes, RRs and RDs with corresponding 95% CI were calculated. A GEE was performed to analyse the effect on dichotomized pain and disability over the total follow-up time. The final model for improvement in pain and disability, respectively, included the following terms in addition to the treatment variable: location of pain (back or neck), follow-up occasion and an interaction term between location of pain and follow-up occasions.

In the analyses of the outcomes “improvement in pain” and “improvement in disability,” patients with scores at baseline less than required to attain these improvements were excluded. Crude SF-36 data were transformed and standardized using recommended procedures and to obtain the dichotomization regarding good health, values from a Swedish population were used\(^\text{120,121}\).

4.4.2.2 Additional analyses

Baseline factors that differed between the treatment groups were considered with regard to their potential confounding effect by means of Mantel Haenszel’s method\(^\text{82}\). If a factor changed the result by 10 percent or more it was considered to be a confounder and should be adjusted for\(^\text{82}\). Using the same logic, the factors were also included one by one in the GEE model.

To estimate the impact of missing responses, additional sensitivity analyses were performed, using multiple imputations with the “predictive mean matching method”\(^\text{144}\).

4.4.3 Study IV

4.4.3.1 Main analyses

The median time to recovery of our sample was estimated with the Kaplan-Meier method. Patients who were lost to follow-up were censored at the mid-point between the last completed follow-up and the next follow-up time.

The model building strategy for time to event data used in Study IV is recommended by Hosmer and colleagues\(^\text{133}\). First, each potential confounder was tested for equality of survivor functions using a log-rank test. Factors with p-value \(\leq 0.2\) were considered for the following multivariable analyses\(^\text{133,145}\). Secondly, collinearity between potential confounders was tested with the Spearman pairwise correlation using a correlation coefficient greater than 0.5 as criterion for collinearity\(^\text{146}\). If factors correlated, the factor that was judged to be the least important from a clinical perspective was excluded from further analyses. Thirdly, the proportionality assumption for each
prognostic factor was verified using Schoenfeld residuals against time\textsuperscript{133,134}. Fourthly, the presence of clinically relevant statistical interactions between neck pain intensity and recovery expectations or depression were tested, and an interaction significant at $p \leq 0.05$ was included in further analyses\textsuperscript{133}. When developing the predictive models (multivariable phase) a manual backward selection procedure was used, based on Cox proportional hazard regression\textsuperscript{133}. The procedure included the development of three models with possible prognostic factors from the domains representing the sequential gathering of medical history done by a physiotherapist. In model 1, sociodemographic factors were considered. In the selection procedure the factors with the highest $p$-values was excluded one by one until all the prognostic factors themselves (or at least one category) had a $p$-value of $< 0.1$\textsuperscript{145,147}. Likelihood ratio test statistics were used to compare the model before and after exclusion of a factor with a $p$-value greater than 0.05 indicated that excluding the variable did not significantly change the fit of the model\textsuperscript{134,148}. In model 2, significant factors from model 1 and factors related to collision, symptoms, comorbidity and health care were considered using the same method as for building model 1. The final prediction model (model 3) was developed with the same methodology, using the remaining factors from model 2 and factors concerning general health and psychology. Signs of collinearity during the multivariable phase were assessed using the variance inflation factor of more than 10 as the criterion for collinearity, and the presence of collinearity was handled as in the univariate phase described above. The predictive ability of the models was measured with the c-index and further internal validated by using 500 bootstrap replicates to get a bias-corrected c-index. This c-index indicates the predictive ability of the model in similar WAD populations as in this study. Furthermore, overfitting was assessed by computing the shrinkage factor\textsuperscript{134,149}. A shrinkage factor of 1.0 indicates perfect fit of the model while a factor of for example 0.8 indicates that 20\% of the inference is due to overfitting. Finally, the overall goodness of fit of the prediction model was evaluated by plotting Cox-Snell residuals and by computing a score test based on Martingale residuals\textsuperscript{133,134}.

4.4.3.2 Additional analyses

To evaluate the impact of missing data we repeated the univariate and multivariable analyses with a “complete study sample”. In this sample we excluded patients who were lost to follow-up and/or with missing data on potential prognostic factors.

4.5 ETHICS

*Study I and Study II* were approved by the Regional Ethical Review Board in Stockholm, Diary No. 2009/457-31 and Diary No. 2013/497-32. *Study III* was approved by the Ethics Committee of the Karolinska Institutet, Diary No. 03-657. *Study IV* was approved by the University Health Network Research Ethics Board, Toronto, Ontario, Canada (REB 10-0216-AE). The original inception cohort study (SGI) was approved by the Research ethics Boards of the University of Saskatchewan and the University of Alberta.
5 RESULTS

This section presents the main results of the studies included in the thesis. Detailed results are reported in the manuscripts at the end of the thesis.

Table 5. Overview of the study samples used in this thesis

<table>
<thead>
<tr>
<th>Study</th>
<th>Region of source population</th>
<th>Study sample size (n)</th>
<th>Proportion of women (%)</th>
<th>Mean age, years (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Stockholm, Sweden</td>
<td>1836</td>
<td>66</td>
<td>55 (15)</td>
</tr>
<tr>
<td>II</td>
<td>Stockholm, Sweden</td>
<td>8994</td>
<td>56</td>
<td>48 (16)</td>
</tr>
<tr>
<td>III</td>
<td>Stockholm, Sweden</td>
<td>409</td>
<td>71</td>
<td>47 (11)</td>
</tr>
<tr>
<td>IV</td>
<td>Saskatchewan, Canada</td>
<td>680</td>
<td>69</td>
<td>39 (15)</td>
</tr>
</tbody>
</table>

5.1 STUDY I

At baseline in 2002, 22 percent men and 25 percent women reported sedentary physical activity. Low, moderate and high PA was reported by 51, 16 and 11 percent men and 50, 16 and 9 percent women respectively. The mean BMI was 26 (SD: 4). Sixty-six percent of the men and 50 percent of the women were classified as being overweight. At the follow-up in 2007, twenty-one percent of the participants had recovered from PBP. Table 6 presents the crude and adjusted results from analyses for the associations between exposures (PA and BMI) and outcome (recovery from PBP). Neck pain was the only variable found to be a confounder and only among men. In addition, age was included in the adjusted analyses.
Table 6. Crude and adjusted RR of recovery from PBP associated with levels of PA and BMI.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=632)</td>
<td>(n=1204)</td>
</tr>
<tr>
<td></td>
<td>Recovered (n=142)</td>
<td>Recovered (n=245)</td>
</tr>
<tr>
<td></td>
<td>Not recovered (n=490)</td>
<td>Not recovered (n=959)</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td></td>
<td>Crude(^a) RR</td>
<td>Adjusted(^b) RR</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td></td>
<td>Crude(^a) RR</td>
<td>Adjusted(^b) RR</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>PA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedentary(^c)</td>
<td>30 (22)</td>
<td>106 (78)</td>
</tr>
<tr>
<td>Low</td>
<td>72 (22)</td>
<td>253 (78)</td>
</tr>
<tr>
<td>Moderate</td>
<td>24 (24)</td>
<td>77 (76)</td>
</tr>
<tr>
<td>High</td>
<td>16 (23)</td>
<td>54 (77)</td>
</tr>
<tr>
<td>BMI(^i)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight(^c)</td>
<td>91 (22)</td>
<td>325 (78)</td>
</tr>
<tr>
<td>Normal</td>
<td>51 (24)</td>
<td>165 (76)</td>
</tr>
</tbody>
</table>

Note: Risk ratio (RR) together with corresponding 95% confidence interval (95% CI).

\(^a\) Crude log-binomial regression model including PA and BMI.

\(^b\) Adjusted log-binomial regression model including PA, BMI, age and neck pain for men and PA, BMI and age for women.

\(^c\) Reference category (RR = 1.0).

Compared to sedentary leisure time, the chance of recovery from PBP was greater for women that were physically active during leisure time. No analyses indicated that PA was associated with recovery from PBP among men or that BMI was associated with recovery from PBP, among men or among women.

In addition we made analyses stratified by poor and good self-rated health (SRH) used as a substitute for more or less severe back pain at baseline. In these analyses the RRs for the association between levels of PA and recovery from PBP were well above 1.0, though not statistically significant, both for women with good SRH (RRs from 1.35 to 1.66) and women with poor SRH (RRs from 1.34 to 1.77).
5.2 STUDY II

At baseline in 2006, about 15 percent of the participants were 65 years or older. Figure 7 shows the distribution of the four healthy lifestyle factors combined to form the exposure healthy lifestyle behaviour, HLB, and Figure 8 the distribution of HLB.

Figure 7. Distribution of healthy lifestyle factors. PA: Leisure physical activity. F/V: Fruit and vegetables.

Figure 8. Distribution of the exposure healthy lifestyle behaviour categories (HLB0 – HLB4).
At follow-up in 2010, 9 percent of men and 11 percent of women reported LTLBP. The final log-binomial regressions were adjusted by socioeconomic status and age in 10-year categories (Table 7). Women with a healthy lifestyle behaviour had a decreased risk of LTLBP compared to women with an unhealthy lifestyle (test for trend: p=0.006). No clear associations between healthy lifestyle behaviour and LTLBP were found among men. There was no of statistical significant effect measure modification found.

Table 7. Association between healthy lifestyle behaviour (HLB) and long duration troublesome back pain (LTLBP) at follow-up 2010 in men and women with occasional low back pain at baseline 2006.

<table>
<thead>
<tr>
<th>Healthy lifestyle behaviour</th>
<th>MEN (n=3646)</th>
<th>WOMEN (n=4658)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude (n/n)</td>
<td>Adjusted (Age, SES)</td>
</tr>
<tr>
<td>LTLBP/no LTLBP</td>
<td>RR (95% CI)</td>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>HLB0</td>
<td>14/155</td>
<td>1.0</td>
</tr>
<tr>
<td>HLB1</td>
<td>71/812</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>(0.56, 1.68)</td>
<td>(0.59, 1.76)</td>
</tr>
<tr>
<td>HLB2</td>
<td>133/1476</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.59, 1.69)</td>
<td>(0.62, 1.78)</td>
</tr>
<tr>
<td>HLB3</td>
<td>60/818</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>(0.47, 1.44)</td>
<td>(0.48, 1.48)</td>
</tr>
<tr>
<td>HLB4</td>
<td>6/101</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.27, 1.71)</td>
<td>(0.30, 1.89)</td>
</tr>
</tbody>
</table>

Note: Log-binomial regression estimating the risk ratio (RR) with 95% confidence interval (95% CI).

a Reduced number of observations due to missing information about socioeconomic status (SES) (men n=292 and women n=398).

b HLB0 = no healthy lifestyle factor, HLB1 = 1 healthy lifestyle factors, HLB2 = 2 healthy lifestyle factors, HLB3 = 3 healthy lifestyle factors, HLB4 = 4 healthy lifestyle factors.

c Numbers of participants with and without long duration troublesome low back pain (LTLBP) at follow-up in 2010.

Compared to the proportion of LTLBP among women with unhealthy lifestyle behaviour, the adjusted proportion of LTLBP was 5 percent lower (RD: -0.05, 95% CI: -0.12, 0.01) for women in the category HLB1, 7 percent lower (RD: -0.07, 95% CI: -0.13, -0.01) for women in HLB2 and in HLB3 and 8 percent lower (RD: -0.08, 95% CI: -0.15, -0.02) for women in category HLB4.

Women had an overall higher adjusted risk of LTLBP than men (p=0.001). Also, the risk of LTLBP was 8 percent for men in HLB0 and 6 percent for men in HLB4 while women had a risk of LTLBP in HLB0 and HLB4 of 17 percent and 8 percent, respectively.
Additional result, not reported in the manuscript
There was also a decreased risk of LTLBP comparing women with more than one healthy lifestyle factor to women with no or one healthy lifestyle factor (HLB0/HLB1) (test for trend: \(p=0.024\)). Compared to women with no or one of the healthy lifestyle factors (HLB0/HLB1), women with all four factors (HLB4) had a 32 percent lower risk of LTLBP at follow-up in 2010 (Figure 9).

![Figure 9. Adjusted RR with 95% CI for the association between healthy lifestyle behaviour and long duration troublesome low back pain among women. Women with no or one healthy lifestyle factor (HLB0/HLB1) constitute the reference group (RR=1.0).](image)

Non-participants had significantly lower proportions of healthy lifestyle factors than the study sample (\(p < 0.01\) for all four factors). The differences in proportions were 8% for non-smoking, 16% for no risk consumption of alcohol, 6% for leisure physical activity and 5% for consumption of fruit and vegetables.
5.3 STUDY III

The patients in Study III were mainly suffering from neck pain (58%) and the majority (56%) had had pain for more than one year. Figure 10 shows the course of pain and disability (CPQ) for the intervention groups, from inclusion to the one-year follow-up.

Figure 10. The mean score of pain and disability during one year after inclusion.

There were statistically significant differences in changes in mean pain intensity and disability between the groups favouring the Index Group (naprapathic manual therapy) at 26 and 52 weeks (Table 8).
Table 8. Baseline values of pain and disability for the index and control groups, changes in the mean of the outcomes for patients taking part in the follow-up at 26 and 52 weeks, compared with baseline, and differences in mean change between groups.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>26 weeks</th>
<th>52 weeks</th>
<th>Diff. in change†</th>
<th>Diff. in change†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline value</td>
<td>Change*</td>
<td>Diff. in</td>
<td>Change*</td>
<td>Diff. in</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>Pain (CPQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index group</td>
<td>5.5</td>
<td>2.6</td>
<td>2.5</td>
<td>1.0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>n=204</td>
<td>n=188</td>
<td>n=182</td>
<td>n=182</td>
<td>n=158</td>
</tr>
<tr>
<td>Control group</td>
<td>5.4</td>
<td>1.6</td>
<td>2.0</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>n=203</td>
<td>n=176</td>
<td>n=158</td>
<td>n=158</td>
<td>n=157</td>
</tr>
<tr>
<td>Disability (CPQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index group</td>
<td>2.7</td>
<td>1.4</td>
<td>1.5</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>n=206</td>
<td>n=186</td>
<td>n=184</td>
<td>n=184</td>
<td>n=157</td>
</tr>
<tr>
<td>Control group</td>
<td>2.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>n=202</td>
<td>n=176</td>
<td>n=157</td>
<td>n=157</td>
<td>n=150</td>
</tr>
<tr>
<td>Disability (WDQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index group</td>
<td>3.0</td>
<td>1.3</td>
<td>1.4</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>n=206</td>
<td>n=189</td>
<td>n=186</td>
<td>n=186</td>
<td>n=150</td>
</tr>
<tr>
<td>Control group</td>
<td>3.0</td>
<td>0.8</td>
<td>0.8</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>n=203</td>
<td>n=178</td>
<td>n=150</td>
<td>n=150</td>
<td>n=140</td>
</tr>
</tbody>
</table>

* The difference in the group mean of the outcomes at follow-up compared to baseline.
† The difference between the Index Group and Control Group with regard to change of group mean at follow-up.
A higher proportion in the Index Group had a clinically important improvement in pain intensity, disability on CPQ and disability on WDQ at 26 weeks, as well as at 52 weeks (Table 9). Further, a higher proportion in the Index Group was totally recovered at 26 weeks (RD = 11%, 95% CI: 4-19) and at 52 weeks (RD = 7%, 95% CI: (-1)-15).

Table 9. Proportion of patients who reached clinically significant improvements in the intervention groups, and risk difference (RD), with corresponding 95% confidence intervals (95% CI) at 26 and 52 weeks follow-ups.

<table>
<thead>
<tr>
<th>Improvement</th>
<th>Index Group (imp/not imp)†</th>
<th>Control Group (imp/not imp)†</th>
<th>RD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>26 weeks</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain*</td>
<td>65% (120/65)</td>
<td>44% (78/99)</td>
<td>21% (10-30)</td>
</tr>
<tr>
<td>Disability* (CPQ)</td>
<td>74% (110/38)</td>
<td>63% (82/48)</td>
<td>11% (4-22)</td>
</tr>
<tr>
<td>Disability* (WDQ)</td>
<td>66% (114/59)</td>
<td>45% (75/90)</td>
<td>21% (10-31)</td>
</tr>
<tr>
<td><strong>52 weeks</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain*</td>
<td>67% (120/59)</td>
<td>50% (79/79)</td>
<td>17% (7-27)</td>
</tr>
<tr>
<td>Disability* (CPQ)</td>
<td>75% (109/37)</td>
<td>58% (68/49)</td>
<td>17% (5-28)</td>
</tr>
<tr>
<td>Disability* (WDQ)</td>
<td>68% (116/55)</td>
<td>49% (72/75)</td>
<td>19% (8-30)</td>
</tr>
</tbody>
</table>

* A clinically important decrease corresponding to at least a two-step decrease in pain score from baseline, or at least a one-step decrease in disability score from baseline, respectively.
† Numbers of patients very much improved/not very much improved in the intervention groups.

None of the baseline factors that differed between the intervention groups changed the estimate considerably and therefore there was no need to adjust for confounding from these factors.

The GEE analyses showed that differences between the groups considered over one year were statistically significant regarding improvement in pain (p=0.002), disability on CPQ (p=0.005) and on WDQ (p<0.001), favouring the Index Group.

Sensitivity analyses performed to evaluate potential bias from loss of follow-up showed no systematic differences in results between analyses with or without imputed primary outcome values.

Health related quality of life (SF-36) were better in the Index Group at 26 weeks and at 52 weeks follow-ups, but the differences were statistically significant only regarding the dimensions bodily pain and social function.
5.4 STUDY IV

Patients with WAD had a mean baseline neck pain intensity of 6.8/10 (SD: 2.0). Low back pain and headache as a result of the collision were reported by 55 percent and 83 percent of the patients respectively. About 67 percent believed that they would get better, either soon or slowly, while about 1 percent felt they would never get better and the rest did not know. Median time between collision and baseline was 14 days. All patients had also visited a medical doctor (MD) and the mean numbers of MD and physiotherapist (PT) visits during that period were 1.9 (SD 1.1) and 2.6 (SD 2.1) respectively. At the six month interview 71 percent had recovered, with a median time to recovery of 97 days.

The prediction model development resulted in a model with seven prognostic factors (Table 10).

Table 10. The final prediction model (model 3) as a result from the multivariable analyses of recovery from WAD. Models 1 and 2 in the published article are not presented.

<table>
<thead>
<tr>
<th>Prognostic factors</th>
<th>Final model (n=633)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta ) (SE)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.01 (0.00)</td>
</tr>
<tr>
<td>No. of days to reporting the collision</td>
<td>-0.02 (0.01)</td>
</tr>
<tr>
<td>Neck pain intensity</td>
<td></td>
</tr>
<tr>
<td>Mild*</td>
<td>0.0</td>
</tr>
<tr>
<td>Moderate</td>
<td>-0.43 (0.14)</td>
</tr>
<tr>
<td>Severe</td>
<td>-0.50 (0.15)</td>
</tr>
<tr>
<td>Low back pain intensity</td>
<td></td>
</tr>
<tr>
<td>No pain**</td>
<td>0.0</td>
</tr>
<tr>
<td>Mild</td>
<td>0.17 (0.15)</td>
</tr>
<tr>
<td>Moderate</td>
<td>-0.16 (0.12)</td>
</tr>
<tr>
<td>Severe</td>
<td>-0.41 (0.15)</td>
</tr>
<tr>
<td>Pain other than neck and back pain</td>
<td></td>
</tr>
<tr>
<td>No**</td>
<td>0.0</td>
</tr>
<tr>
<td>Yes</td>
<td>-0.35 (0.11)</td>
</tr>
<tr>
<td>Headache before collision</td>
<td></td>
</tr>
<tr>
<td>Absent**</td>
<td>0.0</td>
</tr>
<tr>
<td>Mild</td>
<td>0.28 (0.11)</td>
</tr>
<tr>
<td>Severe</td>
<td>-0.03 (0.15)</td>
</tr>
<tr>
<td>Recovery expectations</td>
<td></td>
</tr>
<tr>
<td>Better soon**</td>
<td>0.0</td>
</tr>
<tr>
<td>Better slowly/Never better</td>
<td>-0.66 (0.12)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>-1.09 (0.14)</td>
</tr>
</tbody>
</table>

Note: Overall Goodness of fit for the final model was adequate according to the Cox-Snell residual plot and the score test (p= 0.66). A shrinkage factor of 0.93 indicated that the final model was robust.

* Numbers of subjects are less than the study population (n=680) due to missing answers for prognostic factors in the backward selection procedures.

95% CI: 0.984, 0.996.

Reference category.
The predictive ability (c-index) of the final model showed an acceptable level of 0.68 (95% CI: 0.65, 0.71). Internal validation, using the bootstrap method, showed a robust final model with acceptable ability to predict self-reported recovery from WAD in similar populations of WAD patients (c-index: 0.67 (95% CI: 0.63, 0.70)). The internally validated c-index for model 1 was 0.55 (95% CI: 0.51, 0.58) and for model 2 0.63 (95% CI: 0.59, 0.66). During the multivariable phase of the model building, neck pain intensity and current general health showed signs of collinearity, and as neck pain intensity was considered more clinically relevant, current general health was excluded from further analyses. We found no statistical significant effect measure modification or violations of the proportionality assumption.

The sensitivity analyses using a complete study sample (patients lost to follow-up and/or with missing data on prognostic factors) resulted in a model with the same prognostic factors and a similar c-index as in the main analyses.
6 DISCUSSION

In a discussion of the studies included in this thesis there are many issues to address. I discuss here some of them that I find most relevant and interesting concerning the validity of our results.

In this work, using different study designs, we found factors and treatment that seem to be of importance for the prognosis of back and neck pain. In Study I leisure time physical activity was shown to improve recovery from persistent low back pain among women. Study II implied that for women with occasional low back pain, healthy lifestyle behaviour protects against long duration troublesome low back pain four years later. The results indicate the same association among men though not statistically significant. In Study III, naprapathic manual therapy resulted in greater long-term improvement for patients with non-specific back and/or neck pain compared to patients treated with evidence-based care provided by physicians. The prediction model in Study IV included seven prognostic factors clinically relevant to assess in a physiotherapy consultation for patients with WAD.

6.1 STUDY I

6.1.1 Findings and implications

Women with non-sedentary leisure time had a greater chance of recovery from PBP than sedentary women. There were no associations found among men. The chance of recovery was 46 percent higher for women with low PA (low intensive activity ≥ 2h/week) which may correspond to daily-life activities such as walking or biking to and from work, daily walks, gardening or similar. Moderate and high levels of PA improved recovery even more, 51 and 67 percent respectively. In addition to daily-life activities these levels may incorporate more vigorous activity such as fitness and strength training even though the specific form of activity was not stated in the question. This is in line with evidence saying that advice to stay active as well as exercise affect low back pain.

Interestingly, low PA, as defined in this study, affects low back pain even though it probably reflects physical activity beneath the recommendations for health, which correspond to 150 minutes of moderate physical activity per week. This implies that even small changes in PA can have a good effect on PBP for women, which is valuable in both a public-health and a clinical perspective. It has been suggested that specific back exercise is effective for reducing “chronic” back pain. Nevertheless, following non-specific low back pain patients exposed to PA or specific back exercise for 18 months, Hurwitz and colleagues found that PA reduces the likelihood of concurrent and subsequent low back pain, while back exercise increases the likelihood. Considering Hurwitz’s and our results maybe individuals with persistent back pain should be recommended general physical activity and exercise instead of back exercise?

Are there biological mechanisms that can explain the influence of PA on recovery from PBP found among women? We can only speculate, but physical activity may increase circulation and enhance the production of endorphins suggested to reduce
back pain and possibly reverse connective tissue fibrosis and neurally-mediated inflammation probably linked to back pain\textsuperscript{150,151}. Further, experimental studies indicate that physical activity reduces pain to a higher degree among women than among men\textsuperscript{152}. Could that contribute to the differences found between sexes?

BMI was not associated with recovery from PBP, either among men or among women. Even though, based on our results and previous studies, we cannot say that BMI is not a prognostic factor for back pain, the lack of association may challenge the common belief that it is.

We believe PBP indicates back pain severe enough to have negative consequences for the affected individual as well as for the community. It is therefore important to study. Moreover, as no association between PA and PBP was found among men, the result indicated neither benefit nor harm from PA for men. Therefore we feel confident, considering other benefits of physical activity, to recommend PA both to men and women with PBP.

Methodological differences such as study design, definition of back pain and outcome made it hard to find studies comparable to Study I. However, we found four similar studies with conflicting results. One of them concerned a general population while the others studied patients with back pain\textsuperscript{150,153-155}. In the general population study, with individuals reporting low back pain the previous month, levels of PA was not associated to low back pain one year later\textsuperscript{153}. One study with low back pain patients found that increased levels of leisure physical activity decreased the pain while one showed no association\textsuperscript{150,155}. In the fourth study sedentary lifestyle and low level PA resulted in a higher disability after one year compared to more vigorous PA for sick-listed low back pain patients\textsuperscript{154}. The authors found no relation between BMI and disability or pain at the one year follow-up.

Heneweer and colleagues found a U-shaped relation between physical activity and chronic low back pain in a population-based cross sectional study\textsuperscript{156}. Individuals with sedentary lifestyle and individuals with high levels of physical activity had more low back pain than those with intermediate levels, and the difference was more pronounced among women. In a clinical perspective this U-shaped relationship may be probable also for individuals with PBP, resulting in worse recovery associated with low physical activity as well as high activity. This relationship was not confirmed in Study I as it indicates a dose-response relationship, among women, with better chance of recovery by level of PA. The different results may be due to different intensity of physical activity in the highest level of activity in Study I compared to that in the Dutch study. Further, participants in the Dutch study were older (> 25 years of age) and work related physical activity was included.

The results in Study I could not confirm an association between BMI and recovery from PBP among men or among women. Research on BMI as a prognostic factor for back pain is sparse but among the few studies we found, two showed no association between BMI and low back pain outcomes\textsuperscript{157,158}. One showed obesity (BMI \textsuperscript{\geq} 30 kg/m\textsuperscript{2}) to be associated with persistency of LBP\textsuperscript{159}.  

41
Using the same cohort (SPHC 2002/2007) studying persistent neck pain, we made similar findings as in Study I. PA improved the recovery from persistent neck pain among women of working age. Likewise, the result could not confirm this association among men, or any association between BMI and recovery from persistent neck pain among either sex\textsuperscript{160}.

### 6.1.2 Methodological considerations

In Study I, the intensity of PBP at baseline may be related to the outcome, recovery from PBP, and also to the baseline level of the exposure, PA. In the published paper we have discussed this problem in the terms of reverse causation\textsuperscript{82}. Reverse causation has to be considered in a study such as ours. For example, Hurwitz and colleagues found indications of reverse causation when studying the effect of PA on low back pain in patients randomized to chiropractic or medical care\textsuperscript{150}. However, further reflecting about the situation in Study I we believe the problem to be more related to unmeasured confounding as we lacked baseline information on low back pain intensity or low back pain prior to 6 months, factors that could influence the level of PA. If women with severe pain at baseline had low levels of PA and/or women with mild pain had high levels of PA, this may consequently lead to overestimation of the result. Nevertheless, as discussed in the paper, this may be counterbalanced if some women with more severe PBP were more active than women with less severe PBP. Something that may, for example, be a result of advice to stay active given by health-care providers. Further, the additional analyses stratified on self-rated health (as a substitute for severity of PBP) indicated that being physically active is beneficial for women irrespectively of the severity of the PBP at baseline. Also, as studies show conflicting results, the presence of decreased levels of physical activity among individuals with PBP has not yet been confirmed\textsuperscript{161-164}. What is more, if unmeasured confounding of this type was the main reason for the result in women, would it not be reasonable to find the association among men as well?

The internal validity of a prognostic study depends, among other methodological issues, on the extent to which measures of exposure and outcome are valid and reliable, and whether the results may justify a link between exposure and outcome. Threats to internal validity are commonly divided into three major categories; information bias (e.g. misclassification of exposure and outcome), confounding and selection bias\textsuperscript{82}.

Misclassification can be either non-differential or differential. Differential misclassifications where the classification of exposure depends on the outcome or the classification of outcome depends on the exposure can either over- or underestimate a true effect. Due to the longitudinal design differential misclassification is most often unlikely in prognostic studies but should still be considered. Non-differential misclassification where the classification of exposure is unrelated to the outcome, and vice versa, tends to dilute a true effect, at least for extremes\textsuperscript{82}. The exposures, PA and BMI may have been affected by non-differential misclassification due to inaccuracies in self-reporting or imperfect recall. This could explain the lack of association between BMI and recovery from PBP in our study. Further, if men tend to misclassify their PA to a higher degree than women, non-differential misclassification may partly explain the different findings between the sexes.
The fact that we used a subsample from original cohort and further attrition in the study sample could have introduced some degree of bias due to selection, and if so affected our results. However, some facts contradict this. For selection bias to be present recovery from PBP should have affected attrition differently in exposed and unexposed groups, which we find unlikely. Given that the associations found are true, we believe that our results are valid also for subjects that dropped out or were excluded. In addition, Pizzi and colleagues concluded that bias due to subsampling in a cohort study is probably weak, especially if effect sizes are small.

The longitudinal design of the study supports a causal relationship between PA and recovery from PBP among women. The large number of potential confounders assessed strengthens the internal validity, although we cannot rule out residual or unmeasured confounding, for example, back pain episodes prior to 6 months before baseline or baseline back pain intensity as mentioned above. Further, we consider our overall sample size to be large for a study concerning PBP for at least 6 months. Still the analyses of some exposure categories were based on relatively small numbers (e.g. 27 recovered women with high PA), and should therefore be interpreted with caution. The questions used to assess exposure and potential confounders have, since 1975, been used in many Swedish national and local public-health surveys. They have been subjected to continuous tests (e.g. cognitive testing) and improvements by Statistics Sweden’s test centre, and several questions have acceptable psychometric properties. In addition to the self-reported information, data were collected from Swedish national registers of high quality. Despite this the measurements used may not have been optimal in terms of validity and reliability.

Recovery from PBP in this study was defined as not having experienced periods of considerably disturbing low back pain of seven days or more during the previous five years at follow-up in 2007. This is a very stringent definition as it incorporates the recurrent course of low back pain, why we believe it to support the association between PA and PBP found among women.

Finally, as the information about back pain concerns 6 months before baseline and the exposure PA included activity 12 months prior to inclusion, our results may be interpreted somewhat differently from how we did. Thus, one could imagine that the participants were free of pain before the “past 6 months”, so that an alternative interpretation of the results may be; Women with non-sedentary leisure time who developed PBP had a greater chance of recovery from PBP than sedentary women who developed PBP. The issue merits some reflection but we believe that findings from Tamcan and colleagues support our Study I interpretation of the results. When studying the one-year course of low back pain in the general population, they identified four different clusters of low back pain; fluctuating pain, severe, moderate and mild persistent pain. These clusters remained in the same trajectories during the follow-up time. Supported by their findings, we find it most probable that many of the individuals reporting PBP the latest 6 months in Study I had experienced the pain for longer.
6.2 STUDY II

6.2.1 Findings and implications

Among women with occasional low back pain healthy lifestyle behaviour decreased the risk of long duration troublesome low back pain (LTLBP) four years later. Healthy lifestyle behaviour comprised four healthy lifestyle factors: non-smoking, no risk consumption of alcohol, recommended level of leisure physical activity and recommended consumption of fruit and vegetables. The same tendency was indicated among men, but the results were not significant. Women with one healthy lifestyle factor (HLB1) had a 35 percent decreased risk of LTLBP, while women with all four healthy lifestyle factors (HLB4) had a 52 percent decreased risk of LTLB compared to women with no healthy lifestyle factor (HLB0). In absolute terms (i.e. RD) the result showed that of 100 women with occasional low back pain, five fewer will develop LTLBP if they have one healthy lifestyle factor instead of none and 8 fewer will have LTLBP if they have all four factors instead of none.

An interesting finding was that men with “unhealthy” lifestyle behaviour (HLB0) and women with “optimal” lifestyle behaviour (HLB4) had an equal risk, 8 percent, of LTLBP.

The proportions of men and women in the reference group with “unhealthy” lifestyle behaviour (HLB0) were only 5 percent and 3 percent, respectively (Figure 8). In a public-health perspective this is quite a small part of the population. Therefore we made additional analyses, using a larger group as reference: participants with no or one healthy lifestyle factor, HLB0/HLB1 (21% women and 29% men) (Figure 9). Compared to this reference group of women, adding one healthy lifestyle factor (HLB2) decreased the risk of LTLBP by 23 percent and having “optimal” lifestyle behaviour (HLB4) decreased the risk by 32 percent.

No matter which reference category we used in our analyses adding only one healthy lifestyle factor to the reference groups had a substantial protective effect something we believe important for the individual as well as in a public-health perspective.

When combining potential prognostic factors as in HLB, one must consider a possible synergistic effect that enhances the effect of the separate factors alone. Further, individuals with healthy lifestyle behaviour may have other characteristics that add to the observed inverse association found, for example, favourable coping strategies, positive pain behaviour or other factors that positively affect their health. Some of these factors were tested as potential confounders (e.g. psychological distress and stress) but not all could be so tested.

Only about 40 percent of the women and 30 percent of the men had 3 or 4 healthy lifestyle factors, which demonstrates the health-care and public-health challenge to enhance healthy lifestyle behaviour in the population.

Our study results showing that healthy lifestyle behaviour influence the prognosis of LBP is new and important knowledge with the potential to alleviate a very common public health problem, more prevalent among women. Our findings may have
implications in a public-health as well as in a clinical perspective. This is even more so if our results would apply to individuals without back pain.

The findings also indicate an association between healthy lifestyle behaviour and LTLBP among men. Considering this together with the obvious effect of healthy lifestyle on other health problems, the work to encourage both men and women to adopt to healthy lifestyles should certainly be continued. As mentioned in the present Introduction, knowledge concerning influence from combinations of lifestyle factors on the prognosis of low back pain is lacking. To our knowledge this is the first study concerning the influence of such a combination on the prognosis of LBP assessing men and women separately. We found no study with a similar definition of healthy lifestyle behaviour as in Study II regarding the prognosis of low back pain. However, in one study the authors investigated healthy lifestyle behaviour, similar to our exposure, and the risk of developing chronic low back pain among employees. In line with our results, they found that an “optimal” lifestyle (non-smoking, adequate physical activity, five servings of fruit and vegetables per day and limited or no alcohol consumption) decreased the risk of chronic low back pain by 66 percent compared to employees with an “unhealthy” lifestyle.

6.2.2 Methodological considerations

Exposure information may have been misclassified, as we used self-reported information. For example, some participants may wish to present themselves in a favourable light and overestimate their healthy lifestyle (social desirability) or some may have difficulties understanding the questions and therefore report less well. Moreover, we used two self-reported questions to construct each of the healthy lifestyle factors (alcohol consumption, physical activity and consumption of fruit and vegetables - for details, see the Appendix to Study II). This process may have resulted in suboptimal cut-offs for being healthy or not and further increased the risk of misclassification. However, it is unlikely that such misclassification would happen differentially according to the outcome of LTLBP for which reason it most likely would dilute a true association, at least when comparing extremes. If men tend to misclassify their healthy lifestyle factors to a greater extent than women this dilutive effect from misclassification may partly explain why we did not find any associations among men. Further, if men misclassify the outcome, LTLBP, more than women do, this may somewhat explain the low risk of developing LTLBP found among men. Considering the same discussion in Study I, can we trust men? Are they constantly misclassifying lifestyle factors and pain?

As the study population included individuals not working (65 years and older) we did not assess work-related factors as potential confounders. Socioeconomic status (SES), as was adjusted for in the analyses, is in part related to work and may somewhat compensate for other factors related to work.

About 34 percent of participants in the baseline survey were not part of the study sample due to attrition and exclusion (Figure 4). Non-participants had significantly lower proportions of healthy lifestyle factors than the study sample (p > 0.01). For the reasons discussed in Study I we believe that this selection bias affects mainly
descriptive results such as prevalence, rather than explaining the observed association between health lifestyle behaviour and LTLBP.

Several methodological issues (see Study I) support the validity of the association between healthy lifestyle behaviour and LTLBP found among women in Study II. These include the longitudinal study design, the large sample size, the large number of potential confounders assessed, the dose-response relationship and variables with acceptable quality.

6.3 STUDY III

6.3.1 Findings and implications

In this pragmatic randomized controlled trial we aimed to study the long-term effects (26 and 52 weeks) of naprapathic manual therapy (Index group) on non-specific back and/or neck pain. The control treatment consisted of support and advice to stay active and on pain coping strategies (Control group). Both groups showed improvements regarding pain intensity and disability. Differences in pain intensity and disability at 26 weeks and 52 weeks, compared to baseline, were clinically and statistically significantly larger in favour of naprapathic manual therapy. A greater proportion of the Index group was totally recovered at 26 weeks (11%) and at 52 weeks (7%). Further, the differences between the groups considered over one year were statistically significant also when consideration was taken of the correlation between the repeated measurements at 3, 7, 12, 26 and 52 weeks. In a previous report, using the same trial and studying short-term effects (up to 12 weeks), Skillgate and colleagues found naprapathic manual therapy to be, statistically and clinically, significantly more effective than the control treatment regarding pain intensity, disability and perceived recovery.

In both treatment groups the decrease in mean pain and disability score was largest during the first seven weeks and then leveled off. This is a course of pain well-known for back pain.

The results in Study III are unique since the long-term effects of naprapathic manual therapy have never been scientifically evaluated before. Together with the previously reported positive short-term effects, we believe the study supports recommending a combination of manual therapies as an alternative to consider in primary health-care for these patients. Since back and neck pain are among the most common reasons for seeking primary health-care, our results may be of broad interest and importance. Moreover, the long-term effects of treatments are of special interest from a public-health as well as a health-economic perspective. As an example, following neck pain patients, Korthals-de Bos and colleagues found manual therapy (mobilisation and stabilisation) to be more cost-effective than physiotherapy and standardised care provided by a general practitioner.

As this is the first published study to evaluate the long-term effect of naprapathic manual therapy we cannot compare the results to previous findings, even though some studies have reported on the long-term effects of manual therapy including combinations of manual techniques. In the study by Korthals-Bos, mentioned above, the effects of manual therapy on neck pain, was still significant at 26 weeks but
diminished at 52 weeks, compared to the other interventions\textsuperscript{169}. The UK BEAM pragmatic trial estimated the effect of adding exercise classes, a spinal manipulation package (a combination of several manual techniques), or spinal manipulation followed by exercise, to “best care” in general practice for patients presenting with back pain\textsuperscript{170}. Relative to “best care,” the spinal manipulation package with or without exercise showed only a small benefit at the 12-month follow-up. Studying neck pain, Dziedzic et al. found no additional effect after six months when adding a combination of manual techniques to advice and exercises\textsuperscript{171}. Hoving et al. compared a manual-therapy strategy (combination of manual techniques and coordination or stabilization techniques), physiotherapy strategy and continued care by the general practitioner for neck pain patients\textsuperscript{172}. Short-term recovery was faster with the manual-therapy strategy, but the differences between the treatments groups at the 12-months follow-up were small. In two relatively small trials, combined manual therapies were more effective for neck pain than a minimal intervention approach, but only marginally more effective for low back pain than advice to stay active\textsuperscript{173,174}. In summary, our results indicate a more obvious benefit from combined manual therapy in the long-term, than reported in previously published trials.

6.3.2 Methodological considerations

The pragmatic design in this trial was used to compare two treatment “concepts” as used in everyday clinical practice. It was not meant to evaluate the different components in the compared treatments, so we cannot tell what in the manual treatment that has the positive effects found. The pragmatic design allowed for different numbers of treatments, different length of treatment sessions and treatment strategies adapted to the individual patient’s condition as performed in clinical practice. The differences were permitted in order to evaluate the treatments as used in clinical practice. For this reason, possible placebo effects due to the hands-on methods and the intense patient-therapist interaction in the Index group may explain some of the differences in effect found between the groups.

A placebo effect may also arise due to the participants’ prior expectations regarding the effect of the specific treatments. These expectations were not measured prior to inclusion as a detailed explanation of the control treatment would have been equal to exposing it to all participants. However, before inclusion, participants were asked which intervention they preferred. Sixty percent wanted naprapathic manual therapy, which may indicate bias due to expectation.

How can manual therapy delivered within the initial six weeks still have a significantly better effect than the control treatment up to one year? We believe that the superior short-term effects of manual therapy compared to the control treatment may have enabled patients to return faster to a “normal” life and to be more physically active\textsuperscript{57}. Further, the extended patient-therapist contact may have helped patients to practise a more internal locus of control regarding the management of the back and neck pain problems, thus preventing recurrent episodes of pain. Another possible explanation is that participants received additional naprapathic manual therapy after the base line intervention. At the 26-weeks follow up 11 percent in the Index group and 6 percent in the Control group had taken additional naprapathic manual therapy the preceding six
months. The proportion at 52 weeks was 14 percent in the Index group and 4 percent in the Control group. Thus we do not believe that additional treatment explains all the difference in long-term effect found between the groups in the study.

At the time of the implementation of the trial, advice to stay active was the treatment alternative with the strongest scientific evidence of effect for acute and sub-acute non-specific back and neck pain. This was thus the natural choice as control group. When it turned out that more than 50 percent of the patients had had pain for more than a year one may argue that advice to stay active was not the best choice of control treatment and this choice may contribute to the results. Nevertheless, we are confident that staying active is probably a very important matter even in long-term pain conditions.

*Study III* has strength to mention. The internal validity of the study is supported by the large number of participants included, a low proportion of dropouts (21% in the Control group and 10% in the Index group), and additional confounder analyses made. Further, the results are endorsed by the use of several important outcomes such as improvement in pain and disability, and health related quality of life (SF-36). The results are presented with consideration taken not only of statistical significance, but also of clinically important changes in pain and disability. Through the use of GEE analyses we also considered the correlation between the repeated measures of the outcomes over the one year follow-up.

A recent Cochrane review of spinal manipulative therapy for chronic low back pain judged that the Skillgate and colleagues report on short-term effects, built on the same trial as *Study III*, had a low risk of bias.

### 6.4 STUDY IV

#### 6.4.1 Findings and implications

In *Study IV* we developed and internally validated a prediction model for recovery from WAD in patients consulting physiotherapy within six weeks of the injury. The prediction model had an acceptable ability to predict recovery from WAD (internal validated c-index: 0.67; 95% CI 0.63,0.70). Further, the model was robust and had a good fit. Seven prognostic factors were included: age, number of days to reporting the collision, initial neck and back pain, pain other than neck and back pain, headache before the collision and recovery expectations. The predictive ability, measured with the c-index, increased for each of the three models during the building of the final prediction model. This indicates that expanding the breadth of information in the medical history during a physiotherapy consultation improves the ability to predict recovery of these patients.

As we only completed the development phase of the prediction model, it should not yet be recommended for use in the management of patients with WAD. First, the model has to be validated in similar WAD populations (externally validated) and then it has to be tested in clinical settings to determine its impact on practice pattern, outcome and costs of care (impact analyses). Therefore or model should be treated with caution and could, at the most, give physiotherapists some guidance regarding factors important to assess in the initial medical history information when predicting recovery from WAD.
The model could likewise be of value for physicians as all patients in the study sample had visited a physician in addition to a physiotherapist. The results may also indicate prognostic factors to be considered in future research regarding the prognosis of WAD.

As patients with WAD frequently visit physiotherapists, a fully-validated prediction model (often referred to as a prediction tool/rule) for use in the medical history proceeding would be of great value. Such a tool could be used for predicting the outcome of individual patients or for identifying risk groups of patients with WAD consulting physiotherapy. Moreover, the developed prediction model may be beneficial in other clinical settings, for example in patients consulting chiropractic for other neck pain than WAD; but before such use it has to be fully validated in these specific settings.

We are unaware of any other study developing a prediction model for patients with WAD who consult physiotherapy. However, all prognostic factors in the prediction model developed have been reported, more or less consistently, to be associated with the outcome of WAD, for which reason we believe our model to represent a valuable step in the development of a prognostic tool for use with these patients.

6.4.2 Methodological considerations

Information on some prognostic factors possible for inclusion in the prediction model was missing, for example data regarding specific treatment prior to baseline, lifestyle factors and coping strategies. If they have the potential to change recovery from WAD including them may have changed the model and possibly improved the predictive ability. Further, use of the information from the baseline SGI form as a proxy for medical history collected by a physiotherapist could have influenced the model. This because patients with WAD may answer differently in a clinical situation, than when answering a self-reported questionnaire. For instance, Carragee found self-reported information on pain and other comorbidities prior to collision to be underreported in patients with acute back and neck pain after a motor vehicle collision.

Other studies propose that depressed mood may be associated to recovery from WAD. Our final model did not include depressed mood. It is possible that the physiotherapist’s active clinical management of these patients attenuates the effect of depressed mood on recovery.

In Study IV, 21 percent of the patients were not recovered at the six-month follow-up. This seems to be a low number as one systematic review reported that approximately 50 percent of subjects with WAD had symptoms one year after injury, and another reported about 40 percent not recovered at 6 months. One reason for this discrepancy could be that the recovery definition in Study IV is less stringent than in the reviews. Another possibility is that ‘recovery’ in Study IV was set to the first time during the follow-up when patients reported recovery. This may not reflect the recurrent course of neck pain, and may thus overestimate the recovery rate.
Developing a prediction model using multivariable regression will be a compromise between a model with good predictability and a parsimonious model (i.e. model fit and statistical significance). We followed the common recommendation to use the more conservative criterion of a p-value less than 0.1 for retention in the model, which results in a model with better predictability than when using a p-value less than 0.05\textsuperscript{133,145,147}. However, using this conservative criterion may result in a large model which would therefore be difficult to use in clinical settings. The STarT Back tool, a prediction tool recommended for categorizing low back pain patients in primary care, includes nine factors and we therefore believe that our model with seven factors, may be useful\textsuperscript{180}. Further, we refrained from using a stepwise automated backward process in favour of a manual process as the former produces a model based only on statistical significance without considering the clinical situation of a physiotherapy consultation.

Several things support internal validity in Study IV. When building the prediction model we used procedures and statistical methods recommended in the literature\textsuperscript{133,134,145-149}. The study also fulfilled most of the criteria suggested for an optimal design of prediction studies of WAD and of physiotherapy: the use of an inception cohort, a clearly defined population-based sample, sufficiently long follow-up period (\(\geq 6\) months) and well described measurements that are valid, reliable and clinically relevant. We had a sufficiently large sample size for the multivariable regression used (\(> 10\) events/\(\beta\)-coefficient in a survival analyses); we did not include factors too strongly related to the outcome or to another factor in the model (collinearity); we did internal validation by bootstrapping, and psychological and psychosocial assessments were included in the analyses\textsuperscript{138,181}.

The follow-up rate of 88 percent, and the sensitivity analyses using a complete study sample (without patients who were lost to follow-up and/or with missing data on the prognostic factors) resulting in a model corresponding to the main model, provide confidence that our outcome was not affected by selection bias.

6.5 EXTERNAL VALIDITY

External validity addresses the extent to which findings in a study may be generalized to other settings (populations and time), than the source population\textsuperscript{82}. In epidemiology there are two main issues concerning generalizability of results. When performing descriptive research, the source population has to be representative of the target population, as for example in prevalence and incidence studies. However, in aetiological research, as in Studies I, II and III, representativeness is less important. Instead, generalization of findings should be based more on scientific knowledge, understanding and even assumptions about nature\textsuperscript{82}. Important though, a prerequisite for the generalization of an aetiological study is that internal validity and the role of chance have not seriously affected the findings. A general rule is to be cautious in transferring the results from a single study to other populations.

If considering possible biological effects on low back pain from the exposures in Studies I and II (PA, BMI and healthy lifestyle behaviour) it is reasonable to believe that our findings are valid for a Swedish urban population with specific and non-
specific low back pain. Further, using this approach, there is no reason to believe that the rural community in Sweden or populations in other countries similar to Sweden, would be different. Nevertheless, even though we have considered other competing factors such as psychosocial influences in the present work they may work differently in other settings and additional studies may be needed to support the present findings.

Over 50 percent of the participants in Study III had had pain for more than one year and many have been on sick-leave during the preceding 6 months. Moreover, many had sought care due to their back and neck pain previously. Therefore, we believe the results may apply to patients in primary care, a belief further strengthens by the pragmatic design, imitating the care in an outpatient clinic. One may speculate that recruitment by advertising mainly attracts participants with mild pain, not in need of treatment, but we consider the above facts contradict this speculation.

In 1995 when the SGI cohort used in Study IV was formed, SGI was the only insurance company in Saskatchewan, Canada. Further, health-care providers were mandated to report whiplash injuries to SGI, thus the population of Saskatchewan was the source population of the cohort. Excluded from the SGI cohort were individuals with serious unassociated illness, those who did not understand English, and Workers Compensation claims. Included in Study IV were patients with WAD (Grades I and II) consulting physiotherapists within six weeks of their injury. In addition those patients had visited a physician. Strictly speaking, therefore, in this development phase the present prediction model is only valid in this context until externally validated in other settings, for example, other health-care providers and/or other geographical regions.

6.6 FUTURE PERSPECTIVES

Over the past decade prognostic research regarding back and neck pain has increased and resulted in a numerous prognostic course studies, randomized controlled trials, prognostic factor studies and prediction models. Our knowledge has increased but considering the volume of research the results are somewhat disappointing. The effects of interventions are often small and no treatment has been found to be superior to others either for back pain or for neck pain. Further, the evidence for specific prognostic factors of importance is limited and very few prediction models are used in clinical practice. Reviewers of back and neck pain studies often find it hard to draw conclusions about evidence due to methodological limitations. To improve our results and increase our knowledge we need to minimize such limitations with a consensus on what definitions, measurements and methods to use in the future. Meanwhile maybe we should emphasise the transition of the results so far into clinical practice?

During the work with this thesis we identified areas in which more research is warranted, and have some suggestion on how future research may be formed:

The influence of lifestyle-related factors on low back pain needs to be further studied especially as these are a modifiable and adaptable to use in self-care.
The fact that the results differ between men and women as found in Studies I and II should be considered in future prognostic research. Maybe there are biological mechanisms involved in back and neck pain that affect men and women differently. Therefore, combining epidemiology and biological research could be a future possibility to reach further understanding of the reason for the discrepancies found.

Instead of focusing research on specific prognostic factors maybe clustering factors related to the same entity, e.g. lifestyle or psychology, can help us find subgroups of patients suitable for interventions or recommendations for self-care.

There is a need for more studies regarding the long-term effects of manual therapy, and cost-effective analyses should be performed along with the trials.

I believe that future clinical trials should focus on “treatment packages” as is done in clinical practice. No matter whether it concerns manual therapy, cognitive therapy or “usual care” treatment consists of combinations of interventions, not a single one. Further, the non-specific effect of the patient-therapist interaction is always a part of the treatment and as such not distinguishable from the other parts.

Focus should be on external validation of existing and promising prediction models and getting them into practices, not on developing new ones. A good example is the StarTBack tool for low back pain

We need to reach consensus on what definition of back and neck pain to use, and on what available instruments to use when defining the pain and assessing its severity and its impact on disability; also decide on definitions for recurrent pain and recovery, and how to measure it. Promising attempts exists, but we need to use them unanimously. By improving definitions of back and neck pain, more specific sub-groups of patients can be identified and involved in research. In such sub-groups, prognostic factors and treatment effects could be studied in a more effective and valid way.

We need to reach consensus on what factors to use in the assessment of exposures in prognostic factor research and how to measure them. If such measures have poor validity they need to be improved or replaced. There are, for example, available questionnaires to measure physical activity; but the assessment may be improved using available technology such as GPS, heart rate monitors and “apps” accessible in cell phones.

Recognising back and neck pain as a recurrent condition, longitudinal studies with repeated measurements of the pain and disability may improve our knowledge. Moreover, measuring prognostic factors, e.g. physical activity, repeatedly over the follow-up time would help us understand their influence on back and neck pain. To do so, methods including longitudinal analyses such as general estimating equation (GEE) could be used.

Quantitative and qualitative methods could be combined (“mixed methods”) to better understand the patient’s perspective on pain, disability and recovery and how this may influence the prognosis of back and neck pain.
7 CONCLUSIONS

In conclusion, the present work shows that lifestyle factors are of importance for the prognosis of low back pain among women, that combined manual therapy, such as naprapathy, has a long-term effect on non-specific back and/or neck pain, and that the prediction model for recovery from WAD among patients consulting physiotherapy, has an acceptable predictive ability, but has to be further validated to be used in clinical practice. The differences in findings between men and women indicated in Studies I and II are interesting and may be important to consider in the design of future studies of back and neck pain prognosis.

Study I
Regular leisure time physical activity improves recovery from persistent low back pain among women. No such association was found among men. There seem to be no associations between BMI and recovery from persistent low back pain, among men or among women.

Study II
Healthy lifestyle behaviour decreased the risk of developing long duration troublesome low back pain among women with occasional low back pain. No clear associations were found among men.

Study III
Compared to evidence-based care provided by a physician, naprapathic manual therapy implies a greater long-term improvement (26 and 52 weeks) in pain and disability for patients with non-specific back and/or neck pain.

Study IV
The prediction model built in Study IV is the first step in the development of a fully-validated prediction tool/rule. The model includes seven clinically important prognostic factors for predicting recovery from WAD among patients consulting physiotherapy. It has an acceptable predictive ability, is robust and has a good fit. To be fully incorporated in clinical practice the model has to be validated in other populations and tested in clinical settings.
8 ACKNOWLEDGEMENTS

My time as a PhD student at the Institute of Environmental Medicine, KI, has been very stimulating and a tremendous challenge. I would not have been able to finish this work without the support I have had from all you skilful, helpful and friendly people. I would like to express my appreciation and gratitude to all, in particular:

My supervisors:
Professor Lars Alfredsson, my main supervisor, for opening the world of epidemiology to me, and for all exciting and inspiring scientific discussions during these years. Thank you for giving me this opportunity and for all your support when I needed it most, especially towards the end of this “trip”. I am deeply impressed by your excellent mind and skills in epidemiology and honoured to have had the possibility to work together with you;

PhD Eva Skillgate; what can I say? This would not have been possible if it wasn’t for you. You have supported me a hundred percent, available twenty-four-seven, and encouraged me to go on even when times were rough. Together we have sorted out all possible and impossible definitions of back and neck pain and discovered the secrets of the prognosis. Your never-ending enthusiasm has kept me on track and forced me to look deeper into the “black box” of back and neck pain research. Thank you for being such a nice person and for believing in me. After these years I’m still confused, but at a much, much higher level. I look forward to future collaboration;

Professors Eva Vingård and Irene Jensen, for sharing your long and valuable experiences in musculoskeletal epidemiology. I have enjoyed our meetings where you guided me through the process of producing good research. Your support and help has been much appreciated, and I am very grateful to have got to know you.

Thank you all, I am deeply honoured and proud to have had you as my supervisors throughout the work.

My mentor:
Lena Nilsson-Wikmar, for being a part of this work from the very beginning, helping me to apply for the position as a PhD student. You have been a tremendous support with your wonderful personality and ability to give good advice when I needed it most. Your experience in how to guide PhD students through the process of producing a thesis has been invaluable. I look forward to working with you at the Division of Physiotherapy.

All the members of the MUSIC research group at IMM:
Lena Holm, also my co-author, for always being willing to discuss epidemiology and for having a mind open to science and life in general. I have enjoyed working with you.

Kari Paanalahti, Lina Palmlöf, Anna Peterson and Eva Rasmussen-Barr, my room-mates at IMM, for all the laughs we have had, and all our more or less clever discussion
about the mysteries of epidemiology and life in general. I truly enjoyed your everyday company and I will miss you when this work is finished.

**My colleagues and friends in the research group at The University of Toronto:**
You made Toronto my second home town;

Professor Pierre Côté, the principal investigator during my work on Study IV. Thank you for inviting me to Toronto and making my visits memorable. I had a great time! I admire you as a scientist and person, and am very honoured to have you as a friend. You made me understand that playing baseball takes a lot more “brain” than playing soccer. I’m very glad, too, to have met your wonderful family;

Professor J David Cassidy, for allowing me to work with your data, and for all encouraging and helpful comments on the work we did together. I hope I sometime in the near future will have the privilege to listen to you playing one of your many mandolins;

Assistant professor Eleanor Boyle, for being a co-author and a guide in the jungle of SAS commands;

Maja Stupar, I enjoyed having you as room-mate at the IMM and Toronto office, and our lunches discussing life as a PhD student and life in general;

*Cesar, Craig, Heather, Poonam, Angela, Vicki and Sophie,* I sincerely hope we’ll meet again.

**My other co-authors:**
Professor Johan Hallqvist for sharing your expertise and taking epidemiological thinking to levels I couldn’t imagine;

Professor Linda J Carroll for accepting to be a part of this work and allowing me to use your Canadian data. I greatly appreciate your friendly and respectful way of helping me along.

All the people responsible, teachers and classmates at The Health Care Science Postgraduate School, especially Jan Ekstrand, Lena von Koch and Petter Gustavsson. It has been a privilege to take part in such well-organized education, including all aspects of health care science. Not to forget the financial support that has made my PhD project possible. Thank you!

*Henrik Källberg,* my “fadder”, friend, “personal trainer” and never-ending source of biostatistical and epidemiological knowledge;

Professor Matteo Bottai, the “Mozart” of biostatistics. Your handling of complex analyses seems like a natural skill. I’m deeply impressed by your work and tremendously grateful for all your help;
Professor, Maria Feychting for always having your door open and for your ability to solve the most complicated problems in epidemiology;

Professors Anders Ahlbom and Ulf de Faire, for your enthusiasm and for being such good role models. Despite many years in research you seem to enjoy every moment;

All of my other past and present dear friends and colleagues in the corridor at IMM for enriching my everyday work with joy, especially Annmarie Wesley, Max Vikström, David Pettersson, Karin Leander, Anita Berglund, Bruna Gigante, Mats Halldin, Qory, Dong Yang, Lena Nise, Marie-Louis Serra, Camilla Bengtsson, Tahere Moradi, Annette Linnersjö, Karin Modig, Sofia Carlsson and Axel Carlsson;

And of course, all other colleagues and friends at the administration, and other IMM departments;

All members, no one mentioned no one forgotten, in the Journal clubs and the Work-in-progress group I participated in. It has been a joy to discuss all angles of epidemiology and taken part in your good work;

Professor Cecilia Magnusson, Michael Lundberg and Peeter Fredlund at the Department of Public Health Sciences for allowing me to use the Stockholm Public Health cohort data, and for all help sorting the data and getting started with the analyses. Cecilia, you “forced” me to understand the concept of “reverse causation”;

All my colleagues and especially my bosses Annette Heijne and Cecilia Friden, at the Division of Physiotherapy, KI, for giving me all possible support and help during my thesis work. I look forward to my future full-time work at the division;

Peter Kaiser, for being such a good friend and source of inspiration;

All my friends outside the academia; wondering what I have been doing during these years, but still giving me a hundred percent support.

The ”Bohman clan”;

And most of all, my family;
Camilla, Arvid, Agnes och Anton. Jag älskar er så mycket! Vad vore livet utan er? "Nothing compares 2U” - Sinead O’Connor

Financial support and a travel scholarship were granted by;
The Health Care Sciences Postgraduate School, Karolinska Institutet.

Tony Bohman, September 2013, Stockholm, Sweden
9 REFERENCES


