Epidemiological aspects of musculoskeletal pain in the upper body

Analyzing common and recurrent binary outcomes

Anna Grimby-Ekman
Institute of Medicine at Sahlgrenska Academy
University of Gothenburg

2010

UNIVERSITY OF GOTHENBURG

ISBN 978-91-628-8184-9

Cover illustration from Medicines Talks, No. 15, 2005. Published by National Prescribing Service (NPS), that is funded by the Australian Government Department of Health and Ageing.

Epidemiological aspects of musculoskeletal pain in the upper body
Analyzing common and recurrent binary outcomes

Anna Grimby-Ekman
Occupational and Environmental Medicine, Department of Public Health and Community Medicine, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden

Abstract

The overall aim of this thesis is to gain epidemiological knowledge about musculoskeletal pain in the upper body in light physical work, in relation to gender, psychosocial factors, and computer use; and to compare different methods for analyzing common and recurrent binary outcomes. Two study groups were investigated using questionnaire data: (a) computer users in the Swedish workforce and (b) a cohort of university students. Regression models used were ordinary logistic models, a Cox model (for calculating prevalence ratios), marginal logistic models (GEE), random intercept logistic models (GLMM), Markov logistic models and a Poisson model. Effect measures used were odds ratio, risk ratio and risk difference.

Musculoskeletal pain in the upper body was more prevalent among women than among men, even among young adults. Risk factors among computer users in the workforce were high work demands, and using the computer most of the work day (women). Protective factors were work control and to learn and develop at work, and for women support from superiors. In the university cohort stress, high work/study demands and computer use break pattern were identified as risk factors for neck pain. Stress was a risk factor associated both with developing and ongoing neck pain, and had an impact on both the group average risk and the subject specific risk of neck pain. Computer use break pattern had an impact on the group average risk for neck pain, but on the subject specific risk only for women. Among women stress and computer use break pattern interacted. The effect of presence of both factors exceeded the additive effect of each. Simple questions, about present neck pain and neck pain period past year, captured features of pain, such as general health, sleep disturbance, stress, and general performance. Neck pain period past year did not reflect more serious pain compared to present neck pain. The choice of statistical model should be based on whether a group average risk or a subject specific risk is of clinical relevance. Women and men differed more in the absolute effect measures than in the relative, regarding neck pain. The causality between risk factors and neck pain may differ between women and men.

Keywords: musculoskeletal, pain, neck, repeated measurements, logistic model, odds ratio (OR), risk ratio (RR), risk difference (RD), biological interaction

ISBN 978-91-628-8184-9

List of papers

This thesis is based on the following papers:

I Ekman[*], A., Andersson, A., Hagberg, M., Hjelm, E. W. (2000) Gender differences in musculoskeletal health of computer and mouse users in the Swedish workforce. Occupational Medicine, 50:608–13

II Grimby-Ekman, A., Andersson, E. M., Hagberg, M. (2009) Analyzing musculoskeletal neck pain, measured as present pain and periods of pain, with three different regression models: a cohort study. BMC Musculoskeletal Disorders, 10:73

III Grimby-Ekman, A., Hagberg, M. The validity of asking about the presence of musculoskeletal neck pain in epidemiological questionnaires. *Manuscript*

IV Grimby-Ekman, A., Andersson, E. M., Hagberg, M., Neck pain and perceived stress. Analyzing a longitudinal binary outcome in a cohort study. *Manuscript*

Reprints were made with permission from the publisher of Paper I.

[*] Note the subsequent change in the author's last name, from Ekman to Grimby-Ekman.

Contents

1 Introduction

Musculoskeletal pain is common in general populations in industrialized countries (Buckle and Jason Devereux 2002; Walker-Bone, Palmer et al. 2004; Haldeman, Carroll et al. 2010) and is one of the most common causes for long-term sick leave (Hansson and Jensen 2004; Waddell 2006). According to the International Association for the Study of Pain the economic burden of musculoskeletal pain is second only to that of cardiovascular disease (International Association for the Study of Pain 2010)

When computers were introduced at workplaces and work tasks were computerized in the decade from 1990 to 2000, it was theorized that the introduction of computers at workplaces would lead to a reduction in hazardous physical exposures (e.g., through introduction of machines to do the heavy work and computers to control them), and hence lead to reduced prevalence of musculoskeletal pain due to work.

A clear increase in musculoskeletal pain in the neck and upper limbs has, however, been seen in the workforce in many countries. New exposures have been introduced in old high risk jobs, but also, new risks have been introduced due to monotonous, repetitive computer work. Hence, the need for research in the area of musculoskeletal pain in the neck and upper limbs in a workforce with light physical exposure has been highlighted, and a large number of studies in groups of computer users have been done (Jensen, Borg et al. 1998; Wahlstrom, Svensson et al. 2000; Marcus, Gerr et al. 2002).

Increasing knowledge in the area of musculoskeletal pain in the neck and upper limbs is needed to understand why the prevalence of musculoskeletal pain is high in a workforce with light physical exposure and why the prevalence seems to differ between women and men. The fact that musculoskeletal pain is a highly prevalent and recurrent outcome has implications for the analysis needed to gain this knowledge. In longitudinal studies, the analysis requires use of regression models that take into account the dependence between repeated measurements. For binary outcomes, e.g., ‘pain’ and ‘no pain’, these models are continuously improved and discussed in the literature (Yu, Morgenstern et al. 2003; Lee and Neider 2004; Molenberghs and Verbeke 2004; Yu and Wang 2008). Methodological issues concerning the use of different effect measures and different regression models will be investigated in this thesis.

1.1 Musculoskeletal pain in the upper body

1.1.1 Pain

Pain has been defined as follows by the International Association for the Study of Pain (IASP):

> *Pain is an unpleasant and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.*
>
> (Bonica 1979)

Dimensions of the pain experience are composed of a sensory part, an affective part, and a cognitive part. These dimensions emphasize the complexity of pain in a psychosocial framework (Melzack 1999). The above mentioned dimensions of pain also support the findings regarding central sensitization and dysfunctional inhibition, denoted as dysfunctional central pain modulation (Woolf and Doubell 1994; Woolf and Salter 2000; Lidbeck 2002). Pain can be discussed under several headings with emphasis on its origin, for example, physiological, inflammatory and neuropathic pain (Woolf 1987), or nociceptive pain, peripheral or central neurodysfunctional pain, idiopathic pain (unknown pain mechanism), and psychological pain (Lidbeck 2002).

1.1.2 Musculoskeletal pain

A generally accepted definition for the term "musculoskeletal pain" is difficult to find. Several closely related, but not equivalent, terms describing the conditions involved have been used in the literature, including "musculoskeletal pain", "musculoskeletal disorders", "musculoskeletal symptoms", and "musculoskeletal conditions". An important distinction between "pain" and "symptoms", "disorders" and "conditions" is that pain does not include symptoms such as numbness or tingling. Writing about the Bone and Joint Decade 2000-2010 Woolf and Pfleger describe musculoskeletal conditions as follows as (Woolf and Pfleger 2003):

> *Musculoskeletal conditions are a diverse group with regard to pathophysiology but are linked anatomically and by their association with pain and impaired physical function. They encompass a spectrum of conditions, from those of acute onset and short duration to lifelong disorders, including osteoarthritis, rheumatoid arthritis, osteoporosis, and low back pain.*

Musculoskeletal pain in itself is not a disease, but if it is long lasting and if it negatively affects health it becomes a healthcare issue. Musculoskeletal pain is prevalent in most populations, but all perceived pain does not impact everyday life for the individual. A few short periods of musculoskeletal pain during a lifetime are not normally viewed as a disease. Musculoskeletal pain is normally the body's warning signal when there is risk of tissue damage or when such damage has occurred. Pain can signal that there is a need for recovery of tissue. Therefore, pain needs to be studied in a larger context, together with health and quality of life (QoL). In this thesis work, musculoskeletal pain is seen as a healthcare problem when it is frequently reoccurring, leads to sick leave, or in other ways reduces the capacity, or negatively affects the life, of the individual (Woolf and Pfleger 2003; Turk, Dworkin et al. 2008).

1.1.3 Definition of "musculoskeletal pain" in this thesis

Musculoskeletal pain is in this thesis viewed to be of public health or occupational health interest when it leads to reduced wellbeing, activity limitations, or participation restrictions. Thus, musculoskeletal pain is seen in the "socio-psycho-physiological framework of health and illness" (Rugulies, Aust et al. 2004), which highlights that factors affecting health can be identified at several different levels, e.g., relating to social and economic structures of society; workplaces and families; individual behaviors and physiological processes within an individual.
"Musculoskeletal pain" in this thesis is defined as pain perceived to be related to the musculoskeletal system. The present thesis is on musculoskeletal pain in the neck and upper limbs, which is also referred to as "musculoskeletal pain in the upper body". Paper I includes pain in both the neck and the upper limbs. In Papers II, III, and IV, only neck pain is discussed.

1.1.4 Prevalence of musculoskeletal pain

Work-related neck/shoulder pain has been reported by 25% of workers in 15 European countries (Bongers, Ijmker et al. 2006). In the general population, 15% have been reported to experience chronic neck pain (>3 months) at some point in their lives; and 11–14% of the working population annually experience activity limitations due to neck pain (Haldeman, Carroll et al. 2008). The high prevalence of musculoskeletal pain and the burden of pain to individuals and society is discussed in many studies (Buckle and Jason Devereux 2002; Woolf and Pfleger 2003; Bongers, Ijmker et al. 2006).
Musculoskeletal symptoms are a major cause of sick leave in developed countries, and they are major medical causes of long-term absence from work (Woolf and Pfleger 2003). In the Swedish general working population (16-64 years old), the prevalence of pain in upper parts of the back or neck at least 1 day per week was reported to be 41% among women and 27% among men (Swedish Work Environment Authority 2008). The corresponding prevalence of pain in the shoulders or arms was 37% among women and 26% among men (Swedish Work Environment Authority 2008). Even among young adults (16-29 years old) in the Swedish workforce, the prevalence of musculoskeletal pain is high. Among young women, the

prevalence of pain in the upper parts of the back or neck was reported to be 41% and among young men, 26%; and the prevalence of pain in the shoulders or arms was reported to be 33% among young women and 23% among young men (Swedish Work Environment Authority 2008).

1.1.5 Possible mechanisms of musculoskeletal pain

Musculoskeletal pain is not due to one single mechanism and neither is the subcategory of musculoskeletal pain in the neck and upper limbs. An exposure can injure different structures depending on the individual and other factors, such as working technique and environmental factors.

This thesis investigates musculoskeletal pain in the neck and upper limbs when physical exposure is at a low level, and in combination with psychosocial factors. Several theoretical models of pathological pathways are proposed for this context. The Cinderella model, in its modified form, assumes that the low threshold motor units, first recruited during low-level contractions, are the units that rest the least (Westgaard and De Luca 2001). Another hypothesis concerns the blood vessel-nociceptor interactions of the connective tissue of the muscle (Knardahl 2002). Other proposed mechanisms are mainly based on theories about disturbed cellular respiration and elevated levels of pain-generating substances in muscles. Hence, impaired local muscle circulation or metabolism can be part of the pathophysiology, even if the reasons for these to occur may differ between the models (Johansson and Sojka 1991; Knardahl 2002; Visser and van Dieen 2006; Larsson, Sogaard et al. 2007; Strom, Roe et al. 2009). These pathways could lead to ischemia, i.e., a shortage of oxygen, glucose, and other blood-borne fuels, that is known to induce sensitization and activation of muscle nociceptors (Mense 1992).

1.1.6 Assessment of musculoskeletal pain in epidemiological studies

Assessment of pain is difficult as pain is subjective and multidimensional (Guzman, Hurwitz et al. 2008; Turk, Dworkin et al. 2008). The presence of pain and the perception of pain can only be described and reported by the individual. Musculoskeletal pain has implications for many aspects of daily life, and questionnaires have been developed to assess these different dimensions (e.g., the von Korff chronic pain scale, the Pain Disability Index, and instruments of kinesiophobia and fear of pain) (Tait, Chibnall et al. 1990; Von Korff, Ormel et al. 1992; McNeil and Rainwater 1998; Roelofs, Goubert et al. 2004; Lee, Chiu et al. 2006). The visual analogue scale (VAS), verbal descriptor scales (VDSs), the McGill Pain Questionnaire (MPQ), and similar scales and questionnaires have been developed for the assessment of perceived pain intensity, and quality and activity limitations.

In epidemiological cohort or surveillance studies, where musculoskeletal pain is only one health aspect among many others investigated, the multidimensional aspects of pain have to be captured in only a few variables. Therefore, multi-item instruments for pain assessments are not plausible in the epidemiological survey setting, as in this thesis. One questionnaire with a limited number of pain questions, which is commonly used in the epidemiological survey setting, is the Nordic Questionnaire (NQ) (Kuorinka, Jonsson et al. 1987).
There are some studies examining the validity of self-reported musculoskeletal pain, assessed with the NQ. In these, diagnosis was used as the gold standard. These studies shows mostly that the NQ has high sensitivity, but low specificity (Bjorksten, Boquist et al. 1999; Palmer, Smith et al. 1999), except for one study in which the NQ also had high specificity (Ohlsson, Attewell et al. 1994). Sensitivity should be high, but since severe pain can stem from many causes, other than the specific diagnoses investigated in these studies, low specificity is neither surprising nor a useful measure of validity in this context. However, good predictive validity was found for the NQ regarding number of pain sites and association with disability pensioning (Kamaleri, Natvig et al. 2009).

1.2 Risk and health factors for musculoskeletal pain

Factors that have been shown to be hazardous for pain in the neck and upper limbs are age, gender, smoking, frequent heavy lifting, repetitive work, vibrations, working with your arms above shoulder height, extensive computer work, and precision work, but also psychosocial exposures such as high demands and conflicts. Protective or health factors are physical activity, break taking, perceived reward for efforts, a sense of coherence, a sense of control over the work, and social support (Ariens, van Mechelen et al. 2001; Wahlstrom 2005; Bongers, Ijmker et al. 2006; Griffiths, Mackey et al. 2007; Larsson, Sogaard et al. 2007; Cote, van der Velde et al. 2008; Haldeman, Carroll et al. 2008; Hogg-Johnson, van der Velde et al. 2008). Up to now, there have been few longitudinal studies investigating pain in the neck or upper limbs. Even fewer studies investigate the combination of exposures in relation to pain in the neck and upper limbs (Hogg-Johnson, van der Velde et al. 2008). The consequences for neck pain of different duration, frequencies, and intensity of exposures are for most exposures not clear and need further research.

1.2.1 Light physical work

In this thesis, pain in the neck and upper limbs is investigated in groups with light physical work/studies, e.g., light manual work, sedentary work, office work or white-collar work. Academic studies are in this thesis considered light physical work. Risk factors connected to light physical work are, for example, extensive computer use and, monotonous and repetitive work. Sauter and Swanson in 1996 proposed a model, specific to work with visual display terminal/office technology, of possible pathways from exposure to light physical work leading to musculoskeletal outcomes (Sauter and Swanson 1996). The model includes physical ergonomic and psychosocial exposure as well as biomechanical and psychological mechanisms. Another important aspect of light physical work is prolonged sitting, which is associated with static muscle activity in the neck, shoulders, and spinal area (Griffiths, Mackey et al. 2007). In the Swedish workforce, the occupational groups with a large percentage of time of sitting were mainly in light physical work. They have been reported to be managers, professionals, technicians, clerks, services workers, and shop sales workers, apart from workers in transportation (Swedish Work Environment Authority 2008).

1.2.2 Psychosocial exposure

Psychosocial factors include a broad group of exposures. In earlier literature about work-related musculoskeletal pain, psychosocial exposures were only considered as confounders when work-related musculoskeletal disorders were investigated, but they are now considered to be possible important risk factors (Feuerstein, Shaw et al. 2004). Psychosocial factors at work could be described as work organizational, psychological, and social factors, e.g., as in the work-related demand, control and support model (Karasek and Theorell 1990).

There are several cross-sectional studies showing the association between the demand, control and support model, and health (Eller, Netterstrom et al. 2009; Hausser, Mojzisch et al. 2010; Rau, Morling et al. 2010). The effect of work-related psychosocial factors on pain in the neck and upper limbs are now supported in longitudinal studies according to a review from 2006, even if the relationship is neither strong nor specific (Bongers, Ijmker et al. 2006). According to another review from 2007, there is also evidence of an association between neck-shoulder disorders and psychosocial factors (Larsson, Sogaard et al. 2007). In the review by Bongers et al. (2002) consistent associations were reported between upper extremity problems and high job stress and non-work-related stress (Bongers, Kremer et al. 2002). However, more longitudinal studies are needed for stronger evidence. Deeney and O’Sullivan (2009) in their review concluded that there is growing evidence of psychosocial risk factors that increase the risk for and severity of musculoskeletal disorders (Deeney and O'Sullivan 2009). They also comment on the lack of knowledge about combined effects of psychosocial and physical risk factors. Huang et al. (2002) present a summary of conceptual models linking psychosocial factors and occupational stress to work-related disorders in the upper body (hand, wrist, arm, elbow, shoulder, and/or neck regions) (Huang, Feuerstein et al. 2002). Common for all of the presented models are suggested relations between organizational factors and musculoskeletal outcomes, where organizational factors are suggested to have an impact on ergonomic exposure, biomechanical load, and stress responses (physiological, psychological, and behavioral).

One model mentioned in the above review is the epidemiological model described by Bongers et al. (1993), which illustrate the pathways largely found in the other models mentioned by Haung et al. (1993) (Figure 1). The model was initially presented as an illustration of the context of a systematic evidence-based literature review of epidemiological studies on the etiology of musculoskeletal pain, in combination with concepts from research on stress and health, and of chronic pain (Bongers, de Winter et al. 1993). Note that the epidemiological model and the model suggested by Sauter and Swanson (1996) overlap to a large extent.

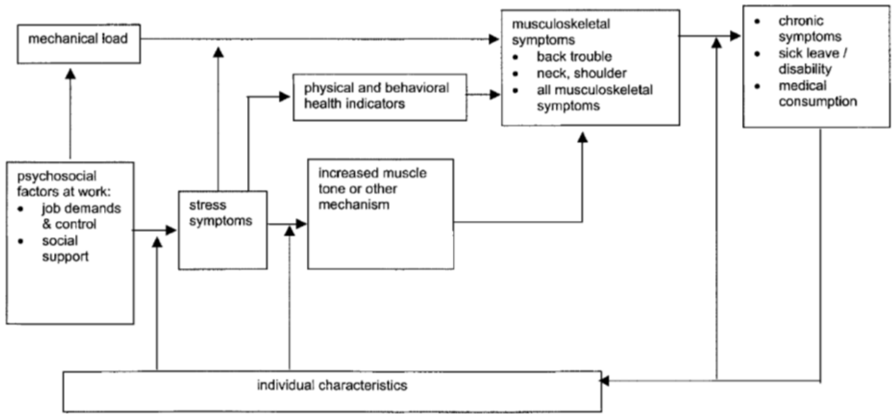

Figure 1. Epidemiological model of musculoskeletal disorders (Bongers, de Winter et al. 1993). Reprinted with permission.

Specific contributions to the psychosocial mechanisms mentioned in the different models are: (a) a pathway from psychosocial factors at work to musculoskeletal symptoms, with stress symptoms as a mediating factor (Bongers, de Winter et al. 1993), (b) psychophysiological mechanisms (stress hormones and blood flow) (Carayon, Smith et al. 1999), (c) non-work demands (Melin and Lundberg 1997), (d) workstyle (Feuerstein, Nicholas et al. 2005), and (e) effects of work organization on symptom perception (Sauter and Swanson 1996).

The possible pathways, from psychological and social factors at work to musculoskeletal disorders, were summarized by Knardahl (2005) as: direct effects on physiological mechanisms, e.g., local muscle circulation and levels of hormones, effects on work style leading to increased biomechanical load, effects on awareness and reporting of musculoskeletal symptoms and affect perceptions, for example on the consequence of pain (Knardahl 2005). Finally, Knardahl suggested that psychosocial risk factors could affect tolerance to other exposures.

The balance theory model of job design and stress proposes that stress is a result of imbalance between various elements of work (Smith and Sainfort 1989; Carayon, Smith et al. 1999), but this reasoning could be extended to also include interactions with home life factors (Dellve, Lagerstrom et al. 2003). Many of the risk and health factors, e.g., computer use, perceived demands, and social support, for musculoskeletal pain in the neck and upper limbs could be argued to come from both work and home life, and from the interaction and lack of balance between those two sources (Melin and Lundberg 1997).

"Psychosocial exposure" and "psychosocial factors" in this thesis work mainly refers to the demand-control-social support model (Karasek and Theorell 1990) and perceived stress (Elo, Leppanen et al. 2003).

1.2.3 Perceived stress

The concept of stress is extensive and represents several aspects. These could be defined as separate entities and includes stress stimuli (external exposure), stress experience (perception of the stimuli), general stress response (physiological response, allostasis), and perceived stress (feedback from the stress response) (Ursin and Eriksen 2004; McEwen 2008). Note that stress stimuli is not by definition a threat to health, but possible ill-health can arise as a product of the stress stimuli, the environmental conditions, and the individual appraisal and coping. This can result in allostatic overload as a result of imbalance between the stimuli and the recovery (McEwen 2008). Perceived stress is suggested to be a mediator between psychosocial exposures and neck pain (Kjellberg and Wadman 2007). Hence, perceived stress could be regarded as a consequence of psychosocial exposures, and is therefore often included in the psychosocial concept. In the present thesis, the concept of stress investigated is "perceived stress", indicating that it is the individual, perceived consequence of stress that is assessed.

1.2.4 Women and men

As mentioned above, musculoskeletal pain is usually more common among women than among men (Karlqvist, Hagberg et al. 1996; Woolf and Pfleger 2003; Strazdins and Bammer 2004; Swedish Work Environment Authority 2008). In Strazdins and Bammer (2004) the higher prevalence of musculoskeletal symptoms in women among a group of white-collar workers was associated with differences in both work and home demands between women and men (Strazdins and Bammer 2004). In one study among blue-collar workers, women had higher prevalence of musculoskeletal symptoms than men, and the women also spent more time on household work and less time on relaxation and exercise than men. Especially parenthood increased the occurrence of musculoskeletal disorders among women (Nordander, Ohlsson et al. 2008). In the same study the exposure for women and men were similar regarding working postures and perceived psychosocial work environment. Women and men did not differed regarding the absolute level of muscular activity used, but women used a higher percentage of their maximum level of muscular activity.

Among Computer-Assisted Design operators, where women and men had identical work tasks, the prevalence of neck pain was higher for women than for men (Karlqvist, Hagberg et al. 1996).
There are several studies indicating biological differences in pain perception between women and men (Wiesenfeld-Hallin 2005). Women are shown to have lower pressure pain thresholds than men (Chesterton, Barlas et al. 2003), and greater response to chemically induced muscle pain (Cairns, Hu et al. 2001). Women also possibly have a more pronounced age-related delay in wound healing, due to reduced estrogen levels with increasing age (Ashcroft and Ashworth 2003).
The effect of fatigue on the spread of pain is suggested to be greater among women than among men, according to an experimental study on mice, where the effect depended on intact ovaries (Sluka and Rasmussen 2010). This may explain why women more commonly develop referred or widespread pain.

To summarize, the higher prevalence of musculoskeletal symptoms among women compared to men may be due to several different factors. Exposure to risk factors may differ between women and men; the impact on pain could possibly be higher for specific risk factors; conditions outside work may be unevenly distributed in a way that is unfavorable for women; biological differences between women and men may influence the impact and perception of pain. Women have higher relative muscular activity and less muscular rest, higher pain thresholds, greater response to chemically induced pain and higher effect of fatigue on the spread of pain and influence of hormones.

Messing and Stellman (2006) point out the importance of studying women’s occupational health and, therefore, the need to define the terms “sex” and “gender” (Messing and Mager Stellman 2006). They state that “sex” may capture genetically based sensitivity to health determinants, while “gender” expresses social forces that could have an impact on exposures and responses to health determinants, e.g., differences in domestic demands and differences in how women and men are treated in society. For example, one study showed that women and men studied were given different medical treatment by physicians (Hamberg, Risberg et al. 2002).
Several papers discuss the need for separate analysis for women and men (Messing and Silverstein 2009; Messing, Tissot et al. 2009; Silverstein, Fan et al. 2009). When gender is included as a determinant in a regression model, a risk factor that has effect on pain only for one gender may be overlooked, as could effect modification by gender. In the present thesis work, the wordings “women and men”, and “gender” are used, and the wordings are here assumed to include both the possible biological and social dimension, as we do not know all the processes of biology, society, and other possible interactions involved.

1.3 Epidemiological effect measures

One aim in occupational medicine research is to identify work-related risk factors and exposures, for causal knowledge and as a basis for interventions with preventive and/or rehabilitation goals. The estimation and identification of risk factors is also important as evidence base for regulation and policies regarding occupational health and insurance.

Three measures of exposure effect are odds ratio (OR), risk ratio (RR) and risk difference (RD), which will here be briefly presented and discussed.
The risk for an individual or the cumulative incidence in a group, to have musculoskeletal pain, is here defined as $P(Y=1|X=x)$ if exposed to x.

The OR, comparing the odds of musculoskeletal pain under the exposure (X=1) to the odds in the reference category (X=0), is defined as

$$OR(1,0)=\frac{P(Y=1|X=1)/\{1-P(Y=1|X=1)\}}{P(Y=1|X=0)/\{1-P(Y=1|X=0)\}}$$

The RR, comparing the risk of musculoskeletal pain under the exposure (X=1) to the risk in the reference category (X=0) is defined as

$$RR(1,0)=\frac{P(Y=1|X=1)}{P(Y=1|X=0)}$$

and the corresponding RD is

$$RD(1,0)=P(Y=1|X=1)-P(Y=1|X=0)$$

In the literature, it has frequently been discussed which effect measure, and which terminology of epidemiological effect measures to use in cross-sectional studies (Lee and Chia 1994; Stromberg 1994; Stromberg 1995). Some authors argue for the use of prevalence ratio (PR), or RR and prevalence difference (PD) or RD (Miettinen and Cook 1981; Greenland 1987; Axelson, Fredriksson et al. 1994; Nurminen 1995; Zocchetti, Consonni et al. 1997; Davies, Crombie et al. 1998; Thompson, Myers et al. 1998) rather than OR.

Another choice to consider regarding effect measures is between using a relative and using an absolute effect measure. In practice, this means mainly making a choice between OR and RR, on the one hand, and RD, on the other. If the risk of pain increases by the same amount in both the reference group and the exposed group, then the relative effect of exposure will decrease while the absolute effect RD is unchanged; and the difference between the OR and RR will be increased. From this, we conclude that the disadvantages of using OR mainly have to do with interpretation of the effect measure and application of the results (Greenland 1987). It has been argued that RD is suitable especially in clinical trials, and in public health settings where the aim is to evaluate the magnitude of the positive effect of a protective action or when a risk factor is removed (Lee, Tan et al. 2009).

In this thesis work some different terms are used regarding epidemiological effect measures. In a parallel meaning to, for example, the term “risk ratio, RR”, the term “proportion ratio, PR” was used in Paper II, and the term “prevalence ratio, PR” was used for cross-sectional data in Paper I.

1.4 Identifying risk and health factors

As mentioned previously, musculoskeletal pain is caused by many factors, possibly acting both separately and in interaction with each other. Estimating the effect measure for only one exposure at a time can give biased results, especially if confounders are also present. Methods based on multiple regression models are able to deal with these problems by adjustment of several confounding covariates. In addition to this, they allow assessment of effect modification between factors.

1.4.1 Methods for cross-sectional studies

As pain in this thesis is a binary response, logistic regression is a relevant method when identifying and estimating the effect of risk or health factors. The variable pain could be denoted as Y, and could be equal to either 1 (pain) or 0 (no pain).

Ordinary logistic regression model
In logistic regression for cross-sectional data, the follow model is used (Agresti 2002)

$$p(x) = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \quad ,$$

and logit transformation is defined as

$$\log \frac{p(x)}{1+p(x)} = \alpha + \beta x$$

where $p(x)$ is the probability of Y=1 when X=x. That is $p(x) = E[Y|x]$. Note that in logistic regression the binomial distribution describes the distribution of the errors of the model. Consequently, the binomial distribution will be the basis for the statistical analysis. This model belongs to the family of generalized linear models and when these use the logit as a link function they are called “logit models”.

Using the model above, the probability of pain is not a linear function of the *X*-variables. Note that the slope of the probability curve is steepest when $p(x) = 0.5$ (Agresti 2002). The tangent at a value x (the rate of change in the probability at point *x*) is equal to $\beta p(x)(1 - p(x))$, where $p(x)$= probability of pain, conditional on $X = x$. Hence, the highest rate of change in probability occurs at $p(x) = 0.5$, and is equal to $\beta / 4$. However, the rate of change in odds is constant, which means that when the OR is known, the value of the maximum rate of change in the probability is known. There are several other possible models to use for binary outcomes (*Y*) in addition to the logistic models introduced above. Two commonly mentioned models will briefly be mentioned below.

Linear probability model
The regression model

$$p(x) = \alpha + \beta x$$

is called a "linear probability model". This is a generalized linear model with a binomial random component and with the identity as link function. A disadvantage of this model is that the estimates of $p(x)$ can fall outside the possible range of values for probabilities. That is, the estimates can take values outside the range 0 to 1. Therefore, this model can usually only be fitted to a restricted range of *x* values, if used at all. The advantage of the model is the easy interpretation of results.

Log-binomial model
The regression model

$$\log(p(x)) = \alpha + \beta x$$

is called a "log-linear model" or, more specifically, a "log-binomial model". This is a generalized linear model with a binomial random component and with the log as link function (Skov, Deddens et al. 1998). The advantage of the log-binomial model is the simple monotonic link function, which makes it easy to present results in terms of probabilities. For this reason, it is becoming popular to use for PR and RR estimation.
One disadvantage of this model is problems with convergence. Some authors are warning against uncritical use of the log-binomial model (Blizzard and Hosmer 2006). Another disadvantage is that the estimate of $p(x)$ can take values >1, which are not valid values for probabilities. This is argued not to be a problem, at least not for cross-sectional data, as the valid parameter space is restricted to $\beta X < 0$ (Skov, Deddens et al. 1998). However, when the estimates are on or near the boundaries of the valid parameter space the estimation algorithm will not converge. Convergence problems are most frequent when the model includes continuous or polychotomous covariates, or if the outcome is common (Yu and Wang 2008). In a review article, it

has also been emphasized that even if they converge, the estimates from a log-binomial model are not guaranteed to be close to maximum likelihood estimates (Lee, Tan et al. 2009). In addition to the already mentioned problems, unlike logistic models, log-binomial models, with recoded outcomes (Y=1 recoded into Y=0) will not generate inverted RRs (Localio, Margolis et al. 2007).

1.4.2 Methods for repeated measurements studies

In the situation described above, there was only one observation for each individual. In Papers II-IV in this thesis work, there were several observations for each individual. Therefore, different models to those previously described will be needed to handle these correlated data. Two such logistic models, the marginal logistic and random intercept logistic model, used in this thesis will be described below.

Marginal logistic regression model

With a marginal model (i.e., the generalized estimation equations, GEE, model), we here refer to a model with no random effects, but with a fixed intercept for all individuals in the population:

$$\text{logit}\left[P(Y_{it} = 1)\right] = \alpha + \beta_1 x_{1,it} + ... + \beta_p x_{p,it},$$

where Y_{it} takes the value 1 (neck pain) or 0 (no neck pain), index p refers to the explanatory variable, index i refers to the individual, and index t refers to the time point.

The marginal logistic model (GEE) takes into account the repeated measurement structure of the data by modeling the correlation structure. No particular multivariate distribution needs to be specified, but the distribution is assumed to belong to the exponential family. The estimation of the model parameters is made using quasi-likelihood equations, which are known as GEEs. The GEE method gives consistent parameter estimates even if the correlation structure is misspecified (Agresti 2002). As the generalized quasi-likelihood does not require the multivariate joint distribution, the full likelihood function is not specified. Hence, likelihood-based methods for test of fit, comparing models, parameter tests, and confidence interval (CI) estimation are not available. For the marginal (GEE) model, the inference is based on the generalized score test (Rotnitzky and Jewell 1990; Boos 1992), since a Wald statistic using empirically based standard errors (SEs) can tend to underestimate the true errors unless the sample size is quite large (Agresti 2002). In the present thesis work, the Wald test was only used to indicate whether a parameter included in a model should be excluded; but the exclusion was then checked with the score test comparing the two models. This model building procedure will be described in more detail in the Statistical analyses section.

Alternative estimating methods to GEE, in the case of a marginal model, are the maximum likelihood (ML) approach and the method of weighted least squares (WLS). Methods based on ML fitting are not practical when the number of explanatory variables increases, as the number of multinomial probabilities to estimate then increases dramatically (Agresti 2002). Some limitations of the method of WLS are that it requires large sample sizes, categorical explanatory variables, and contingency tables that are small and not sparse (Agresti 2002).

Random intercept logistic regression model

A random intercept logistic model (the generalized linear mixed model, GLMM) is here referred to as a model with a random individual intercept and hence the estimation of subject specific effects (Agresti 2002):

$$\operatorname{logit}\left[P\left(Y_{it}=1\right)\right]=\alpha+u_i+\beta_1 x_{1,it}+\ldots+\beta_p x_{p,it}$$

where Y_{it} takes the value 1 (neck pain) or 0 (no neck pain), index *p* refers to the explanatory variable, index *i* refers to the individual, index *t* refers to the time point, and u_i is the individual random intercept assumed to be independent and normally distributed with mean zero and variance σ_u^2. The relative effect of x_p, measured as OR, is estimated by $\exp\left(\beta_p\right)$. Conditional on a specific individual *i*, the random intercept logistic model (GLMM) resembles an ordinary logistic model, and hence maximum likelihood estimates are available for the parameters (Agresti 2002).

If we assume we have two levels of exposure and want to look at the absolute effect of the exposure, then taking the median of individual absolute effects (individual probability differences between the two exposure levels) is equal to the difference between the median probability for one exposure level (median over all individual probabilities at this exposure level) and the median probability for the other exposure level (median over all individual probabilities at this exposure level).

Model fit

The Hosmer-Lemeshow test for model fit has been suggested for dichotomous outcome (Hosmer, Hosmer et al. 1997). For the marginal (GEE) model, the test works only under specific circumstances, one of which is a small intra-cluster correlation (Hosmer and Lemeshow 2000). For the random intercept model, the test can be used but requires that individual predicted values are calculated and this includes an estimated value of the random effect term for each individual (Hosmer and Lemeshow 2000). A check for model fit can also be based on diagnostic plots of the residuals. However, the residuals from binary data must be interpreted with caution; for example, when the fitted values are small, they can be uninformative and lose relevance (Agresti 2002). As discussed by Agresti among others, further research is needed into model checking and diagnostics tools, both for the random intercept model (the generalized linear mixed model, GLMM) and for the marginal model (GEE) (Agresti 2002).

Marginal versus random intercept model

As can be seen in Figure 2 the “slopes” of the individual risk curves are steeper than that of the population average risk curve (at least around p = 0.5), all obtained from the random intercept logistic model (GLMM). The effect measure OR of an exposure, achieved from a random intercept logistic model (GLMM), will show a stronger effect (either OR>>1 or OR<<1), than that from a marginal logistic model (GEE) with a fixed intercept (Agresti 2002). According to Molenberghs and Verbeke (2004) the parameters from a marginal logistic model (GEE) are always smaller than the parameters from a random effects logistic model (GLMM) (Molenberghs and Verbeke 2004).

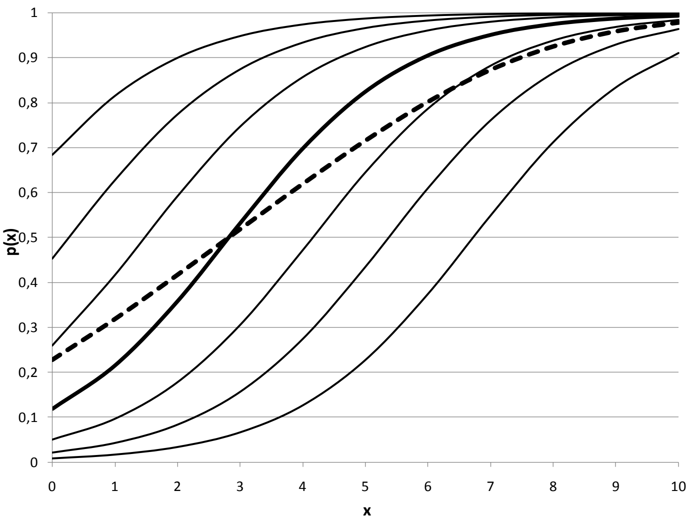

Figure 2. The risk or probability, p, as a function of an explanatory variable, *x*. Subject specific risk curves based on a logistic random intercept model (GLMM) (solid lines), compared to the population average risk curve (dotted line) as obtained by integrating out the GLMM. The population average risk curve obtained from a marginal logistic model (GEE) is close, but not identical, to the dotted line.

The following approximate relationship holds between the slope parameters in the marginal logistic model (GEE) and the slope parameters in the random effects logistic model (GLMM):

$$\hat{\beta}_{RE} \Big/ \hat{\beta}_{M} = \sqrt{c^2 \sigma_u{}^2 + 1} > 1 \quad ,$$

where σ_u^2 is the variance of the random intercepts, and

$$c^2 = 16\sqrt{3} \Big/ 15\pi$$, (Molenberghs and Verbeke 2004).

From the random intercept model, the estimated absolute effect of an exposure is an estimate of the absolute effect from the median curve. The random component represents the variation between individuals. On the probability scale, this represents the variation in the location on the x-axis of the individual probability curve. In a marginal model, the probability is a representation of the population average probability. That is, for each x_j-value a mean is calculated over the individual probabilities, keeping all other explanatory variables fixed. Note however, that if the random intercept variance is large, then the variation between the individual locations of the probability curves is extensive and it is not possible to clearly state what effect a specific decrease in exposure will have for a specific individual. This implies that further research could possibly identify additional risks or health factors explaining the between-individual variance. The effect of a decrease in the exposure could still be presented at the group level. If the variation of the random intercept is small this means that the individual probability curves are close to the population average probability curve.
The estimated absolute effect of stress from the marginal model is a representation of the mean of the individual absolute effects. The marginal absolute effect of stress could also be interpreted as the absolute effect of stress on the prevalence of neck pain in the population.

In addition to the marginal logistic model and the random intercept logistic model for repeated measurements, the logistic Markov transitional model and log-linear Poisson model should be mentioned. These two models only use some of the information in the repeated measurements. In the present thesis the Markov transitional model only uses two time points, and the log-linear Poisson model summarizes the outcomes from all time points in the single outcome "number of years with pain".

Markov transitional logistic regression model
As above, in the Markov transitional logistic model, Y_{it} is the binary response for an individual *i* at time *t*. In a first-order Markov transitional model the probability of the outcome at time *t* is conditioned on the outcome at time *t-1* (Yu, Morgenstern et al. 2003). Therefore, the first-order Markov transitional model is, if logistic regression is used, only an ordinary logistic regression model with the outcome at time *t-1* as an additional explanatory variable.

Poisson log-linear regression model for counts
A fundamentally different way to handle the repeated measurements over time is to use the Poisson regression model for counts. With this method, a variable representing the number of cases for each individual over the repeated times of measurement is modeled. A binary response variable, *Y*, for example pain, is assessed at several time points for each individual. The sum of *Y* over all time points

will represent counts, number of reported pain during the study period. Let μ denote the expected value of Y, $E[Y]$. The model is then

$$\log(\mu) = \sum_p \beta_p x_p \quad ,$$

where index p refers to the explanatory variables and Y is assumed to have a Poisson distribution.

1.4.3 Sample size and power

In logistic regression, as in all statistical analysis, the issue of sample size and power needs to be addressed (Nemes, Jonasson et al. 2009). In logistic regression, the concern is more about the number of outcome events (here, the smaller number of the binary outcome) than about the total sample size of a study. Several papers recommend a minimum of about ten events per explanatory variable (or number of parameters, to estimate category variables) to be included in the model (Peduzzi, Concato et al. 1996). Hence, if the outcome is extremely rare (or extremely common) a larger sample size is needed to achieve this sufficient number of events. Even if the number of events is of importance in logistic regression, sample size is still an issue. For samples with fewer than 100 observations the use of maximum likelihood estimates (MLEs) are not reliable, but the MLEs should be adequate for samples above 500 (Nemes, Jonasson et al. 2009). However, the sample size is highly dependent upon the specific study.

1.5 Estimating epidemiological effect measures

1.5.1 Calculating adjusted risk ratios and risk differences

In the literature, the choice of regression model for estimating RR or RD is discussed, as well as how to calculate their CIs. There are suggestions on methods to calculate RR or RD, based on the results of different models (Greenland 2004), such as logistic regression (Flanders and Rhodes 1987; Localio, Margolis et al. 2007) or log-linear models, log-binomial model and Poisson model (Skov, Deddens et al. 1998; Thompson, Myers et al. 1998; Spiegelman and Hertzmark 2005). For estimating RR, Cox's proportional hazard model has also been suggested (Axelson 1994; Thompson, Myers et al. 1998).

Arguments raised against the use of Cox's proportional hazard are that error estimates are too large (especially for common diseases) and hence, the method tends to be conservative (Thompson, Myers et al. 1998). The use of robust sandwich variance is a possible solution to statistically compensate for the variance-inflation problem, but is under debate (Lee, Tan et al. 2009). An additional disadvantage of the Cox's proportional hazard model is that it produces estimates of RR only, and not of the prevalence or probability. Therefore it cannot be used to calculate RD. The main argument against the logistic model as a base for estimation of RR or RD seems to be that the method is indirect and the calculation of standardized estimates and CIs is a little more complicated than for the alternative models. Authors usually do not mention the logistic regression as a method to calculate RR or RD, and many times the logistic model is used as a synonym to a method only for achieving OR. Though it has been suggested that hypothesis testing should be done with test statistics directly related to ORs, and that RRs and their CIs should be calculated to present the magnitude of the exposure effect (Flanders and Rhodes 1987).

1.5.2 Estimating effect measures from a logistic regression

Odds Ratio

The regression parameters estimated in the logistic regression are directly related to the OR, and hence, the P-values and SEs of the regression coefficients are directly related to inference about the OR. The OR for the effect of an explanatory variable is constant over different combinations of other explanatory variables, when the model does not include product terms. This means that only one value needs to be presented to show the effect of an exposure, even if other explanatory variable are included in the model. This is an advantage of using OR compare to those of using the effect measures of RR and RD, but there are other, interpretational, disadvantages with OR. From a logistic model the OR is the most commonly calculated effect measure and is easily calculated from the estimated regression parameters (see below).

$$OR(1,0) = \exp(\alpha + \beta_1 + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}) / \exp(\alpha + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}) = \exp(\beta_1)$$

In the case of repeated measurements the estimated OR could either be a population average OR, or a subject specific OR for a median individual, depending on whether the estimate comes from a marginal or a random intercept logistic model. Note that the population average OR is not equal to the mean of the individual OR:s, based on the subject specific risks (Greenland 1987).

Risk Ratio

An alternative measure to OR is RR. From a logistic regression,

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \log(odds(x)) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots$$

Therefore the following is true for a risk ratio comparing the risk for two levels, 0 and 1, of an exposure *x*.

$$RR(1,0) = \frac{p(1)}{p(0)} = \frac{odds(1) \big/ (1 + odds(1))}{odds(0) \big/ (1 + odds(0))}$$

If the risk is small, i.e. approximately ≤ 0.10, it follows that the odds approximately equals the probability, and hence

$$RR \approx \frac{odds(1)}{odds(0)}$$

In the application of musculoskeletal pain the prevalence is many times 20% or 30%. Then $1 - p \approx 1$ does not hold and the above approximation is not valid. Instead, from a logistic regression, the RR for the effect of a binary x_1 could be calculated as follows:

$$RR(1,0) = (1 + \exp(-\{\alpha + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}\}) / (1 + \exp(-\{\alpha + \beta_1 + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}\})$$

Note that in contrast to OR, RR for an exposure x_1 depends on the value of the other explanatory variables (x_2, ..., x_p). This complicates the calculation of the RR if the model is complex. The RR above is an estimate of the population average in the case of a marginal model, while in the random intercept model, it is an estimate of a subject specific RR of a median individual. Note that the population average RR is not equal to the mean of the individual RR:s, based on the subject specific risks (Greenland 1987).

Risk Difference

Instead of a ratio, calculations of difference can be used, e.g., the RD. The effect of a binary exposure variable x_1 can be estimated from the logistic regression as

$$RD(1,0) = \left(1 \big/ (1 + \exp(-\{\alpha + \beta_1 + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}\})\right) - \left(1 \big/ (1 + \exp(-\{\alpha + \beta_2 x_{2,it} \ldots + \beta_p x_{p,it}\})\right)$$

Note that the RD, like the RR, is dependent of the value of the other explanatory variables. Also here the interpretation depends on whether a marginal or a random intercept logistic model is being used. For the RD, the population average absolute effect is the mean of the individual absolute effects, based on the subject specific risks in the population.

When groups with different reference levels are compared, the ratio can differ even if the absolute effect is the same (e.g., 0.4/0.2 = 2 and 0.8/0.6 = 1.33).

The actual risk level is easily estimated from a logistic regression. In a model with exposure groups defined by the variables X_1 and X_2, the risk of neck pain can be estimated as

$$P(Y=1) = 1/(1+\exp(-\{\hat{\alpha}+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2\})),$$

with a 95% CI estimated as

$$1/(1+\exp(-\{\hat{\alpha}+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2 \pm 1.96 * \sqrt{Var(\hat{\alpha}+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2)}\})).$$

1.6 Statistical interaction versus biological interaction

The word "interaction" can be used to refer to two distinctly different phenomena: statistical interaction and biological interaction. "Statistical interaction" here refers to the departure from an additive statistical regression model by including a product term, with the goal to build a model that better fits the data. In a logistic regression for example, this means departure from an additive model on the logic scale. Statistical interaction is an association not necessarily causal, and is scale-dependent. As discussed previously in this thesis, the effect of an exposure or factor can be represented by several different epidemiological effect measures, such as OR, RR, or RD. Heterogeneity of an effect is called "effect-measure modification". This is equal to departure from additivity of effects on the chosen effect scale.

Assume that we have factor A and factor B. Then there is effect-measure modification of RR if

$$\frac{P_{A=1,B=1}}{P_{A=0,B=1}} \neq \frac{P_{A=1,B=0}}{P_{A=0,B=0}},$$

and of RD if

$$P_{A=1,B=1} - P_{A=0,B=1} \neq P_{A=1,B=0} - P_{A=0,B=0}.$$

Therefore, when effect-measure modification is present on one effect-measure scale, e.g., OR scale, this always implies absence of effect-measure modification on another scale, for example the RR scale (Rothman and Greenland 1998).

"Biological interaction", on the other hand, refers to a situation where two (or more) risk factors for a disease are involved in the same sufficient cause, according to the sufficient cause model (Rothman 2002). This means that there is at least one pathway from the risk factors to the disease in which both these factors are required.

Therefore, biological interaction implies that the risk of disease due to both factor A and B exceeds the sum of the risks of disease due to each of the factors:

$$\left(P_{A=1,B=1} - P_{A=0,B=0}\right) = \left(P_{A=0,B=1} - P_{A=0,B=0}\right) + \left(P_{A=1,B=0} - P_{A=0,B=0}\right)$$

If no sufficient cause for the disease includes both of the factors, then these factors are said to be independent and no biological interaction is present between them. Biological interaction is equivalent to effect-measure modification of RD. Therefore, biological interaction is evaluated as departure from additivity of absolute effect measures (Rothman, Greenland et al. 2008).

2 Aims

The overall aim of this thesis work is to gain epidemiological knowledge about musculoskeletal pain in the upper body in light physical work in relation to gender, psychosocial factors, and computer use; and to compare different methods for analyzing and interpreting common and recurrent binary outcomes.

Specific aims are

1) to investigate
   a) the influence of gender on the risk of musculoskeletal pain in the neck or upper limbs (Papers I, II, IV).
   b) whether musculoskeletal pain in the neck or upper limbs is associated with psychosocial factors, computer use, and lifestyle (Papers I, II, IV).

2) to evaluate the validity of pain assessments of present neck pain, neck pain period past year and duration of present neck pain in relation to aspects of health and decreased general performance (Paper III).

3) to evaluate whether results regarding gender, perceived stress, and computer use differ depending on whether
   a) a group average (marginal) model or a subject specific (random intercept) model is used (Paper IV).
   b) the effect measure of choice is an OR, an RR or an RD (Paper IV).

3 Methods and material

3.1 Study samples

This thesis work was based on two separate sources of data, *The Work environment survey 1995* (Swedish Work Environment Authority 1995) from Statistics Sweden and the cohort *Health 24 Years* (Herloff, Ahlborg et al. 2003). In Table 1, a general overview is given of the study design, participants, and sex and age of participants in the four papers included in this thesis. In Paper I, the number of computer users was 2044, but of these, 340 had missing values on the outcome. Hence, the data used included 1704 observations (870 women, 834 men).

Table 1. General overview of the data sources.

Paper	Type of study	Number of time points	Study group	N	Proportion of women, %	Baseline age, years
I	Cross-sectional	1	Workforce	1704	50	16-64
II	Longitudinal	3	Students	1204	52	19-25
III	Longitudinal	5	Students	1200	52	19-25
IV	Longitudinal	5	Students	1200	52	19-25

In Paper II, only three time points were used, baseline, and 1-year and 2-year follow-up, as the data on the 3 and 4-year follow-ups were not available at the time of the study.
Data on four of the participants included in Paper II were deleted, and not used in Papers III and IV. Two of these participants were excluded due to misclassification as university students when they actually were upper secondary education students; data on the other two participants were excluded due to double registrations.

3.1.1 Statistics Sweden data

Paper I uses a cross-sectional study sample that was based on data from the Work Environment Survey conducted in 1995 by Statistics Sweden (SCB) commissioned by the National Board of Occupational Safety and Health. About 14000 individuals in the Swedish workforce were asked to participate. The survey material was representative of the Swedish workforce, aged 16-64 years, and consisted of interviews by phone, and questionnaires. The sample was drawn from those answering the Work Force Survey during October, November, and December and employed at the time of the interview.
In Paper I, the subgroup of interest was computer users, and hence, the total workforce was classified into computer users and non-computer users. “Computer

user” was defined as a worker who (a) used a personal computer (PC), computer terminal, or equivalent device for work; (b) used computerized equipment for 50% or more of the work day; and (c) used a computer mouse. In Paper I, the computer users are also called “visual-display terminal (VDT)” workers.
The percentage of non-respondents in the SCB survey was 24%. The non-response to individual questions (partial missing) was between 1% and 3%.

3.1.2 University cohort

Papers II-IV are all based on a cohort originally focusing on information and communication technology (ICT) use in relation to health. The cohort was recruited in 2002 and consisted of university and college students enrolled in medical and information technology (IT)-related studies and of upper secondary educational students. In Papers II-IV, only the university and college students were included and hence, below only descriptions and figures relevant to this group will be presented.
An invitation letter was sent to all students in medical and IT-related studies, aged 19-25, according to university and college enrollment lists in five cities in western and southern Sweden, (Gothenburg, Lund, Linköping, Borås, and Skövde). The invitation letter described the cohort and offered free tickets to the cinema as an incentive to participate. Students could agree to participate either by mail or by online registration; and were then, in a second letter, given an individual username and password for the Web-based questionnaire.
The university cohort was approved by the Regional Ethical Review Board situated at the University of Gothenburg, Gothenburg, Sweden.
The baseline response rate was 69% and the number of respondents to the questionnaire was 1200 (627 women, 573 men). Note, however, that in Paper II, the number of respondents was 1204 for the reasons explained above. For the same reasons, the following figures will relate to the more correct study sample of Papers III and IV.

The 2-year follow-up was only available to those answering the questionnaire at the 1-year follow-up, in contrast to all other follow-ups where the questionnaire was available to all participants joining the cohort at baseline. This may explain the relatively low response rate at the 2-year follow-up seen in Figure 3.

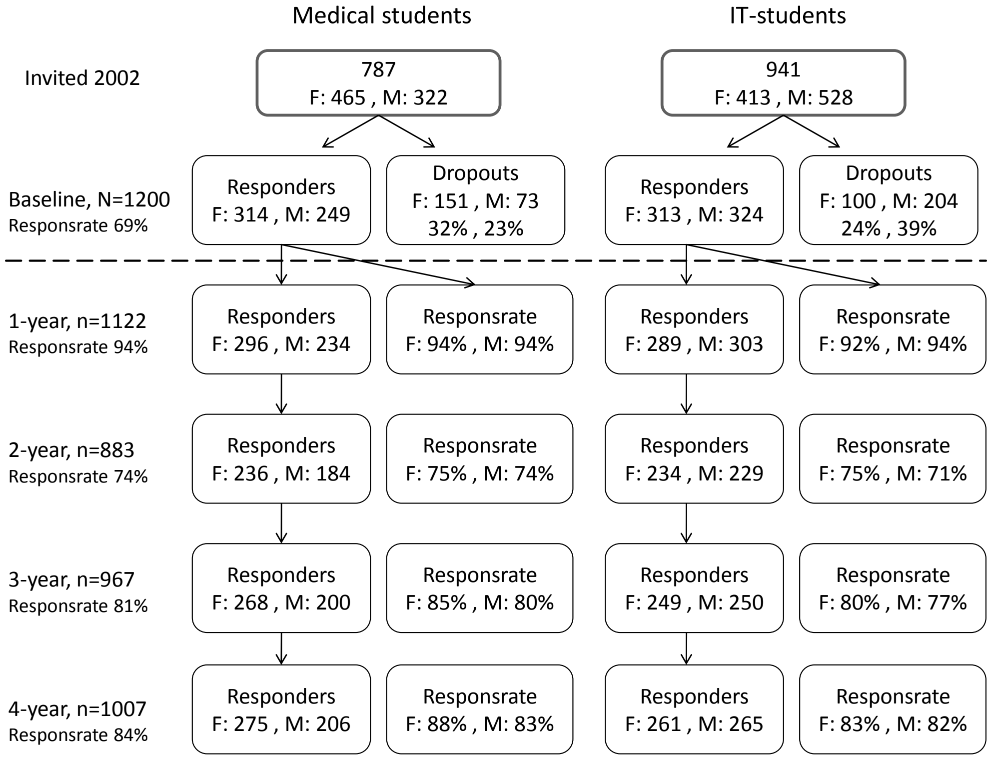

Figure 3. Participant flowchart showing the time points of data collection. The response rates are in relation to baseline.

The two educational groups of students, medical and IT, are equal in most variables examined (Table 2). Note, however, that a smaller proportion of the IT-students ate breakfast regularly and a larger proportion smoked compared to the medical students. The medical students, on the other hand, had more hours of studies per week. The proportion of female IT students experiencing high demands and stress was higher than that of female medical students. The variable that differed the most, between the two groups, was computer use pattern, which is not surprising.

From Table 2, raw data on RD and RR, comparing women and men regarding the prevalence of neck pain, can be calculated for each of the two educational groups. Among IT students, the women had a higher prevalence (%) than men for neck pain (RD = 11, RR = 2), but among medical students, the women had an even higher prevalence (%) compared to men (RD = 21, RR = 3).

Table 2. Descriptive baseline data for women and men of the university cohort, divided into the two educational groups.

	Medical students		IT students	
	Women N=314	Men N=249	Women N=313	Men N=324
Age, years (mean, SD)	23 (1.5)	23 (1.4)	23 (1.5)	22 (1.6)
BMI (mean,SD)	21 (2.6)	23 (2.2)	22 (2.5)	23 (2.7)
Breakfast eaten at least 5 days/week (%)	90	85	80	74
Physical activity, hours/week (mean, SD)	4 (4.0)	4 (3.7)	3 (3.9)	4 (4.8)
Smoking (%)	3	4	8	7
Snuff use and not smoking (%)	2	9	2	12
Asthma (%)	8	8	9	6
Having children (%)	0[a]	2	2	4
Married or living with a partner (%)	30	21	31	21
Not speaking Swedish as mother tongue (%)	12	11	15	7
Gainful employment, h/week (mean, SD)	2 (5.5)	3 (10.2)	4 (7.6)	5 (10.7)
Scheduled studies, h/week (mean, SD)	22 (12.7)	21 (11.9)	15 (9.7)	13 (9.8)
Unscheduled studies, h/week (mean, SD)	14 (10.8)	14 (11.5)	17 (12.4)	15 (12.8)
Present neck pain (%)	21	10	33	12
Decreased general performance, among those with neck pain (%)	28	29	34	31
Duration of neck pain >7 days, among those with neck pain (%)	71	75	66	69
Neck pain period (%)	31	11	35	16
Stress (mood) (mean, SD) [b]	3.3 (0.99)	3.1 (0.88)	3.8 (1.05)	3.2 (1.09)
Energy (mood) (mean, SD) [b]	4.0 (0.81)	3.9 (0.84)	4.0 (0.77)	3.8 (0.86)
Stress (mean, SD) [c]	64	49	67	45
High work/study demands *Not too high (%)*	56	73	43	70
Not affecting home life (%)	29	17	39	22
Affecting home life (%)	15	10	18	8
High home life demands (%)	4	5	3	7
Good relationship with superiors (%)	98	97	92	89
Good relationship with colleagues (%)	95	94	94	92
Computer use pattern *With breaks (%)*	89	75	43	27
1 period/week without a break (%)	5	15	19	15
>1 period/week without a break (%)	5	10	38	58

[a] Only one woman. [b] The mood scale according to (Kjellberg and Iwanowski 1989). [c] Percieved stress according to (Elo, Leppanen et al. 2003). SD = standard deviation.

3.2 Variables used in this thesis

The variable musculoskeletal pain in the neck and upper limbs (Paper I) is a dichotomization (Table 3) of a question in the Work Environment Survey conducted by the SCB 1995 (Swedish Work Environment Authority 1995). The phrasing of the question was,

> *After work, do you experience pain in any of the following places? Upper parts of your back or neck. …*

Table 3. Description of musculoskeletal pain variables in the four studies.

Concept	Variable name	Description	Paper I	Paper II	Paper III	Paper IV
Musculoskeletal pain in neck or upper limb	Musculoskeletal neck and upper limb symptoms	Pain, after work, at least 1 day per week in upper back, neck, shoulders, arms, wrists, or hands	•			
Musculoskeletal neck pain	Pain at present (II) Present neck pain (III) Neck pain (IV)	Present pain/ache in upper back or neck		•	•	•
Musculoskeletal neck pain	A period of pain (II) Pain period past year (III)	Period of pain/ache: during the past year, lasting more than 7 days in the upper part of your back/neck		•	•	
Pain duration	Present neck pain duration	Number of days with current pain/ache			•	

The neck pain variables used in Papers II-IV, regarding present neck pain (Table 3), only differ in the variable name and are all based on the same question:

> *Do you suffer from any of the following AT PRESENT?*
> *Pain/ache from the upper back/neck.*

Both the variables concerning a period of neck pain were based on the question:

> *Have you, during THE PAST YEAR, suffered from any of the following for more than 7 days running?*
> *Pain/ache from the upper back/neck.*

The question in the NQ (Kuorinka, Jonsson et al. 1987) closest to these phrasings is:

> *Have you at any time during the* ***last 12 months*** *had trouble (pain, ache, discomfort) in: neck …*

At the time of the study, it was not possible to include figures in the web-based questionnaire and hence, only words were used to define the location upper back or neck.

Note that in the present thesis work, pain from this area is denoted “neck pain” according to the definition of the Task Force on Neck Pain and Its Associated Disorders (Guzman, Hurwitz et al. 2008). The focused pain location was according to Figure 4, and this has good agreement with the recommended definition of neck pain, according to the work of the Task Force on Neck Pain and Its Associated Disorders (Guzman, Hurwitz et al. 2008).

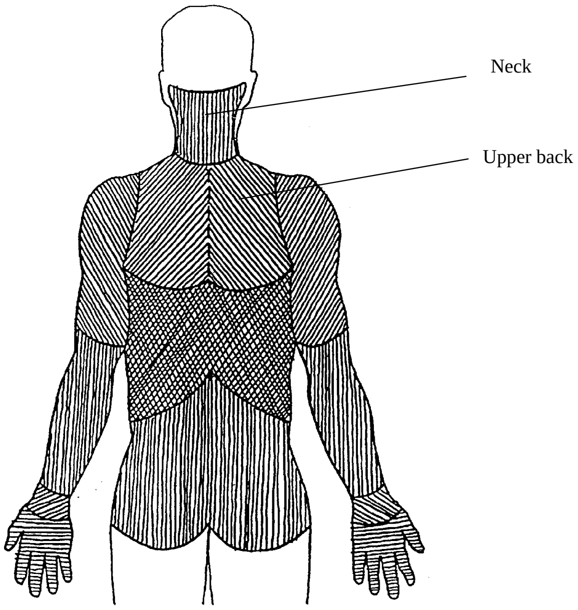

Figure 4. Defining the focus pain location in Papers II-IV.

Medical examinations of a sub-sample of 42 participants, from the baseline of the university cohort, included pain drawings from which presence of neck pain could be defined. From these drawings, the agreement between neck pain, according to Task Force on Neck Pain and Its Associated Disorders, and pain in the upper back and neck (Figure 4) was 93% (95% CI 81.0 ; 97.5). The three participants not in agreement with the definition of the pain area were defined as having pain in the upper back, according to Figure 4, but the marked region was not wholly included in the area defined by the Task Force on Neck Pain and Its Associated Disorders. None of the 42 participants had a problem answering the question due to lack of a picture. That is, all pain drawings showing areas outside (and not even partly included in) the area defined by the Task Force on Neck Pain and Its Associated Disorders were clearly separated from this, e.g., low back, hands or knees.

Brief descriptions of all other variables used in this thesis work will be presented in the Table 4 and Table 5 below. The perceived stress variable used in Papers II and IV, (Table 4) is a modification of a question developed by Elo Leppanen et al., which is a question of good validity (Elo, Leppanen et al. 2003).

Table 4. Description of variables used in this thesis: Stress, energy, computer use and psychosocial factors.

Concept	Description	Paper I	Paper II	Paper III	Paper IV
	Stress and energy				
Perceived stress	Feeling tense, restless, or anxious or unable to sleep at night due to constant thoughts about problems. Experienced during the last 12 months, for more than 7 consecutive days		•		•
Stress	Stress (mood) over the last 7 days			•	
Energy	Energy (mood) over the last 7 days			•	
	Computer use				
Duration of PC work (%)	Percentage of the work day that you are working with computer equipment	•			
Duration of PC work (hours)	Hours using PC last 7 days		•		
Computer use pattern	Computer use without breaks (working on a computer for ≥4 hours continuously without a break more than once last 7 days)		•		•
	Work/study demands, control and social support				
Work/study demands (hours)	Hours of scheduled and unscheduled work/studies over the last 7 days		•		
Work demands	Too much to do at work	•			
Work/study demands	Too high demands that negatively affect home and family life Too high demands that do not negatively affect home and family life		•		
Home life demands	Home/family demands negatively affect studies/work		•		
Work control	Involved in planning your work	•			
Social support	Support from superiors	•			
Social support	Good relation with superiors		•		
Social support	Good relation with colleagues or study mates		•		
Learn and develop at work	Learn and develop in occupation	•			

Table 5 Description of variables used in this thesis: Life style, health and general performance.

	BMI, life style, health and general performance				
Concept	Description	Paper			
		I	II	III	IV
Body mass index	BMI≤25, BMI>25		•		
Breakfast eaten regularly	Eating breakfast >5 days per week		•		
Smoking	Smoking daily or almost daily during last 7 days		•		
Snuff use	Used snuff daily or almost daily during last 7 days, but does not smoke		•		
Physical activity	Hours per week last 7 days		•		
Opportunity to progress your career	Learn and develop in occupation	•			
Asthma	Diagnosed asthma		•		
Health	General self-rated health			•	
Sleep disturbance	Difficulties falling asleep, repeated waking and difficulties falling asleep, not thoroughly rested at waking over the past 6 months			•	
Decreased general performance	Decreased general performance due to ache/pain in muscles or joints over the past 30 days			•	

In a longitudinal study, the explanatory variables can be of two kinds: constant over time (e.g., gender) or time dependent (e.g., age that increases monotonically with time, or perceived stress that varies unrestrictedly across time). To investigate short- and long-term effects, as was the aim in Paper II, the explanatory variable has to vary over the time points.

Time (Papers II-IV)

In a longitudinal study, several different measures of time could be relevant. Examples of time variables are calendar time (here given as year), time in cohort, or age. Age could reflect an exposure that generally increases or decreases with age. Age at baseline could here represent a pseudo-measure of the number of episodes of pain occurring before the start of the cohort. These episodes are hypothesized to be a risk factor for later pain. That is, higher age may imply higher number of earlier episodes of pain. Note, however, that in the present cohort, the age at baseline was low and the age range among participants was no more than 6 years (ages 19-25). As the study group consisted of young adults, only a few participants of whom had children, and who had not worked for a long time, a possible effect of age on pain occurrence would probably not represent unobserved exposure due to family life or typical work exposures. Therefore, in this thesis work, it was assumed that the age effect was minor. Calendar time can be used to reflect a trend in society. The time

variable in this work has possible values of 0, 1, 2, 3, and 4, which is equivalent to calendar years 2002, 2003, 2004, 2005, and 2006.
In the present data, from the university cohort, the time span, 5 years, is probably too short for calendar time to represent trends in society. Years in the cohort could for example capture an effect of being in the cohort.

3.3 Statistical analyses

All analyses were performed using the statistical package SAS (SAS Institute, Cary, NC, USA). All analyses were performed separately for men and women (Messing, Tissot et al. 2009; Silverstein, Fan et al. 2009), except for the analyses in Paper II that were instead adjusted for gender. Therefore, additional analyses, with separate analysis for women and men, are presented in the Results section for Paper II.
All P-values were two-sided and considered statistical significance if less than 0.05. For continuous variables mullticollinearity was said to be indicaed if the correlation coefficient was ≥0.65. For categorical variables, binary or ordinal with three categories, multicollinearity was said to be present if there was a high positive association (percentage agreement >80% Papers I, II and IV and >85% Paper III) between explanatory variables or a high negative association (percentage agreement <20% Papers I, II and IV and Paper III <15%) between the explanatory variables.

Paper I was a cross-sectional study using ordinary logistic regression (PROC LOGISTIC) for identifying associations between the outcome musculoskeletal pain and the explanatory factors. OR with 95% CIs were estimated. Cox regression (PROC PHREG), with constant time equal to 1, was used to calculate PRs as an effect measures, with 95% CI (Skov, Deddens et al. 1998). In addition to the original results in Paper I analyses adjusting for occupational group are presented in this thesis, and as a complement to the original results from Paper I PDs comparing women and men are also calculated.

Before starting to fit a model to the data, it is necessary to describe the characteristics of the outcome and explanatory variables at hand. Of course, this is necessary even in the previous case of cross-sectional data, but in the case of repeated measurements (over time), some specific characteristics are of importance: variation over time in outcome and explanatory variables. In Paper II descriptive statistics were calculated for the explanatory variables, describing variability within individuals over the 3 years. For continuous variables the mean and variance of the within individual variation was calculated, and for categorical variables the percentage of participants in same category all 3 years, was calculated. In Paper III, the baseline description of the study group was presented in terms of the prevalence of pain. The baseline distribution of present pain duration was described in terms of the median value (md) for the number of days with pain, and the 1st and 3rd quartiles (Q1 and Q3), together with the mean number of days to indicate skewness. The baseline distribution of each of the outcome variables was presented as its mean and SD, together with the maximum and minimum observed values; exceptions were general health and decreased general performance, which were presented as prevalence and 95% CI. Variation in reported neck pain over time was presented as the proportion of respondents reporting pain at all time points and as the proportion of respondents reporting pain none of the five time points. In Paper IV the baseline characteristics of the categorical variables stress and computer use pattern were presented as proportions; and in addition, variation over time was presented as the proportion of persons remaining in same response category at all time points.

In Paper II, three methods for regression analysis were used: a marginal logistic regression model (generalized estimating equation, GEE) with the outcome "pain at present"; a Poisson regression model with the outcome "number of years with pain"; and an ordinary logistic model (here called "Markov transitional model") with the outcome "a period of pain". All analyses were performed using the procedure PROC GENMOD in SAS. In all regression models, Wald-type 3 P-values were used to assess the effect of each factor on the outcome variables.
The first step of the analysis was univariate models, where one explanatory variable at a time, together with gender, was included. In the second step of the analysis all those explanatory variables with P-values $\leq$0.2 were included in a multiple regression model. The choice of this limit, rather than P$\leq$0.05, as an inclusion criterion for explanatory variables was made in order not to miss an explanatory variable with a possible association with musculoskeletal pain (Hosmer and Lemeshow 2000).
The OR and corresponding P-values derived from logistic regressions were used to test the association between different explanatory variables and the outcomes (Flanders and Rhodes 1987). In addition PRs were calculated. The calculations were based on the parameter estimates in the simple logistic regression (Flanders and Rhodes 1987; Localio, Margolis et al. 2007). These calculations produces one estimate of PR for each possible combination of other explanatory variables included in the regression, and were therefore not calculated in the case of the multiple regression models. Here, we calculated one PR for women and one PR for men and the adjusted PR is then the mean of these two PRs (McNutt, Wu et al. 2003).

In order to handle the longitudinal design in Paper III, mixed models (PROC MIXED) were used to analyse the relationships between pain assessments and the outcomes sleep disturbance, stress and energy. For the outcomes general health and decreased general performance a random intercept logistic regression (the generalized linear mixed model, GLMM) was used for the longitudinal design. No multicollinearity was found in the explanatory variables.

In paper IV, two logistic models for repeated measures were used, namely marginal logistic regression (GEE) and the random intercept logistic model (GLMM), to examine the association between neck pain and the explanatory variables stress and computer use pattern,. Inference for the random intercept model was based on the likelihood ratio (LR) test. No multicollinearity was present among the explanatory variables used in this study. When using the random intercept logistic model, risks and effect measures RRs and RDs presented in this paper are calculated based on the median risk curve and are estimates for the median individual. The contribution of different exposures to the risk of neck pain is graphically presented in the Result section for Paper IV, to illustrate absolute effects and possible biological interactions (Rothman, Greenland et al. 2008). Confidence limits for estimates of RRs and estimates RDs were based on the Delta method (Billingsley 1986).

The model building procedure in Paper III and IV was based on recommendations by Hosmer and Lemenshow (Hosmer and Lemeshow 2000). The models were constructed as follows: Where P-values were ≤0.2 (Paper III) and P≤0.25 (Paper IV) the variable was included (Mickey and Greenland 1989) in a univariate model the explanatory variables were included. They were excluded from the multiple model if P>0.1, but only if the P-value from the test comparing the nested models was <0.05; and the parameter estimates, for the other variables, did not change by more than 15% (Paper III) or more than 20% (Paper IV). A product term was included if the model included more than one main effect. The product term was excluded if P>0.05 (Hosmer and Lemeshow 2000). For main effects included in the model, all two-way product terms were analyzed according to a procedure corresponding to the one above, with the limits for the P-values being 0.15 (instead of 0.25), and 0.1. The above described model building procedure was used also for inclusion of lagged variables (explanatory variables from one year earlier, t-1). It should be noted that P-values of parameter estimates in the final model are not adjusted for multiple tests and due to the model building procedure should be interpreted with caution.

4 Results

4.1 Musculoskeletal pain among computer users in the workforce

4.1.1 Paper I

The prevalence of musculoskeletal pain in the neck and upper limbs among computer users was 36% for the women and 16% for the men, and the unadjusted PR = 2.2 (95% CI 1.89, 2.64) comparing women to men. Among computer users, women compared to men had higher prevalence of musculoskeletal pain in the neck and upper limbs in all occupational groups, PR>1 (Table 6). The variation in PR between occupational groups may be due to a difference in prevalence level. If the prevalence increases the PR decreases, given that the difference between prevalence for men and women was constant. This does not explain the differences in PR between the occupations seen in Table 6.

Table 6. Prevalence of pain in the neck and upper limbs, PRs and PDs among computer users in different occupational groups.

Occupational groups	N	Cases	Women, %	Men, %	PR	PD
Legislators	52	13	32	20	1.6	11.8
Professionals	573	132	34	15	2.3	19.0
Technicians and associated professionals	541	114	29	16	1.8	13.2
Clerks	398	152	41	8[a]	5.1	33.2
Service workers and shop sales workers	44	13	38	13[a]	2.8	24.6
Craft and related trades workers	33	10	80[a]	21	3.7	58.6
Elementary occupations	40	11	40[a]	26	1.6	14.3

[a] <5 cases.

The proportion of the workforce working with PCs today has increased compared with 1995 when the study was performed. The proportion working with PCs was 51% (women: 47%, men: 54%) in 1995 and 71% (women: 73%, men: 70%) in 2007 (Swedish Work Environment Authority 2008). The proportion working with a PC at least half of the work day was 22% (women: 24%, men: 20%) in 1995, and 38% (women: 38%, men: 37%) in 2007 (Swedish Work Environment Authority 2008). Consequently, computer use is now present in many different occupations and in many different work tasks and hence, the variable does not show a well-defined group. Today computer use may be viewed as an important risk factor in studies of work exposures, but it must also be considered, at least for some groups, as an important leisure time exposure.

For women, identified health factors were “involved in planning work”, ”support from superiors” and ”learn and develop in job”; risk factors were ”too much to do” and ”duration of PC work 100%”. For men, the only clearly identified health factor was ”learn and develop in job”; and the only risk factor identified was age. Note that the variables regarding duration of computer work assess the relative time (% of work day) and not absolute time (hours/work day). For example, among individuals working half-time the variable ”duration of PC work 100%” represents roughly 4 hours computer work per day, while among individuals working full-time the same variable represents roughly 8 hours computer work per day.

4.1.2 Adjusting for occupational group

An additional logistic regression was performed, adjusting for occupational group. A categorical variable for occupation group, with one category for each of the three largest occupational groups, (Table 7), and a reference category of the rest of the occupational groups was used. This did not change the associations between symptoms and the risk/health factors, and hence, occupational group does not seem to be a confounder.

Table 7. Distributions of the binary explanatory variables in the three largest occupational groups.

Variable	Women			Men		
	Professionals n=270, %	Technicians n=237, %	Clerks n=431, %	Professionals n=405, %	Technicians n=392, %	Clerks n=47, %
Too much to do at work	65	59	53	65	67	58
Involved in planning work	86	77	69	92	86	77
Support from superiors	71	77	76	61	69	58
Learn and develop in job	91	87	80	96	96	85
Duration of PC work 75%	30	25	31	27	32	27
Duration of PC work 100%	30	35	34	31	24	25

4.2 Risk factors for neck pain in the university cohort

4.2.1 Paper II

The prevalence of neck pain among the young adult women was 28% (pain at present), 22% (developing pain) and 56% (ongoing pain); and the prevalence of neck pain among the young adult men was 12% (pain at present), 9% (developing pain) and 41% (ongoing pain). The prevalence of present neck pain in this group of young adults was in the same range or slightly smaller than the prevalence of neck and upper limb pain seen among computer users in Paper I, in the occupational groups called ”technicians and associated professionals” and ”professionals”.
The explanatory variables which varied over the 3 years were work/study time, physical activity (to some extent), high work/study demands, computer use pattern and perceived stress. It was therefore possible to test these variables for short-term effect on musculoskeletal pain.

The analysis of the multiple marginal logistic model (GEE) for present neck pain showed statistically significant differences between women and men (Table 8). Identified risk factors were smoking, high work/study demands, computer use pattern (≥2 periods per week without breaks) and perceived stress. Breakfast eaten regularly was a protective factor.

Table 8. **Present neck pain:** analyzed with marginal models (generalized estimating equation, GEE). All odds ratios (ORs) are adjusted for gender, using men as the reference category. The total number of respondents varied from 737 to 885 because of incomplete data in some cases. PR = proportion ratio.

Explanatory Baseline variables	Exposed cases	Simple model		Multiple model			PR
		OR	P-value	OR	P-value	95% CI	
Women/men	160	2.9	<0.001	2.6	<0.001	2.02 ; 3.41	2.4
Overweight	22	1.1	0.463				1.1
Breakfast eaten regularly	168	0.64	0.001	0.77	0.058	0.593 ; 0.995	0.71
Snuff use	5	0.68	0.116	0.61	0.061	0.344 ; 1.09	0.72
Smoking	23	2.2	<0.001	2.0	0.003	1.35 ; 3.00	1.8
Physical activity		0.98	0.157	0.99	0.318	0.967 ; 1.01	0.99
High work/study demands (ref: not too high)	104		<0.001		0.004		
Not affecting home life	71	1.4		1.3		1.06 ; 1.61	1.3
Affecting home life	45	1.8		1.5		1.16 ; 1.99	1.6
High home life demands	12	1.5	0.020	1.2	0.243	0.868 ; 1.77	1.4
Work/study time		1.0	0.266				1.0
Good relationship with superiors	206	0.78	0.139	0.89	0.507	0.624 ; 1.26	0.82
Good relationship with colleagues	207	0.76	0.175	0.92	0.700	0.614 ; 1.39	0.81
Computer use pattern (ref: 0)	120		<0.001		0.012		
One 4 h period without a break	24	1.0		1.0		0.748 ; 1.34	1.0
At least two 4 h periods without a break	76	1.6		1.4		1.11 ; 1.71	1.4
Asthma	23	1.2	0.360				1.1
Perceived stress	163	1.8	<0.001	1.6	<0.001	1.34 ; 2.01	1.6

The analysis of the multiple model for developing musculoskeletal neck pain showed a relative difference between women and men, Table 9. Identified risk factors were computer use pattern (≥2 periods per week without breaks) and perceived stress. Asthma seemed to be a risk factor for developing pain, but the result was not statistically significant.

Table 9. Developing neck pain: Pain period at 1-year follow-up, analyzed with Markov transitional models, for those without pain period at baseline. All odds ratios (ORs) are adjusted for gender (using men as the reference category). The total number of respondents varied from 267 to 326 because of incomplete data in some cases. Estimates were not calculated for explanatory variables with fewer than five exposed cases. PR = proportion ratio.

Explanatory baseline variables	Expose cases	Simple model		Multiple model			PR
		OR	P-value	OR	P-value	95% CI	
Women/men	86	3.0	<0.001	3.1	<0.001	2.00 ; 4.82	2.5
Overweight	10	0.80	0.522				0.82
Breakfast eaten regularly	103	0.81	0.423				0.84
Snuff use	5	0.85	0.737				-
Smoking	7	1.2	0.685				1.1
Physical activity		0.99	0.599				1.0
High work/study demands (ref: not too high)	62		0.126		0.296		
Not affecting home life	48	1.4		1.4		0.913 ; 2.20	1.4
Affecting home life	16	1.5		1.1		0.596 ; 2.21	1.3
High home life demands	8	2.2	0.063	2.2	0.080	0.912 ; 5.07	1.9
Work/study time		1.0	0.889				1.0
Good relationship with superiors	116	0.67	0.280				0.72
Good relationship with colleagues	114	0.56	0.098	0.72	0.373	0.354 ; 1.48	0.62
Computer use pattern (ref: 0)	64		0.016		0.021		
One 4 h period without a break	20	1.6		1.7		0.941 ; 2.94	1.5
At least two 4 h periods without a break	42	1.9		1.8		1.16 ; 2.89	1.7
Asthma	13	2.0	0.046	2.0	0.052	0.996 ; 3.91	1.7
Perceived stress	84	1.9	0.002	1.7	0.011	1.13 ; 2.63	1.7

In the multiple model for ongoing neck pain, no risk or health factors were reliably identified, not even gender. Perceived stress, asthma, and women (compared to men) may be associated with ongoing neck pain (Table 10), but were not found to be statistically significant.

Table 10. **Ongoing neck pain:** Pain period at 1-year follow-up, analyzed with Markov transitional models, for those with pain at baseline. All odds ratios (ORs) are adjusted for gender, using men as the reference category. The total number of respondents varied from 214 to 264 because of incomplete data in some cases. Estimates were not calculated for explanatory variables with fewer than five exposed cases. PR = proportion ratio.

Explanatory baseline variables	Expose cases	Simple model		Multiple model			PR
		OR	P-value	OR	P-value	95% CI	
Women/men	106	1.9	0.027	1.6	0.157	0.837 ; 3.01	1.4
Overweight	14	0.87	0.848				0.96
Breakfast eaten regularly	114	1.5	0.202				1.2
Snuff use	4	-	-				-
Smoking	12	1.4	0.231				1.1
Physical activity		0.98	0.684				0.99
High work/study demands (ref: not too high)	64		0.606				
Not affecting home life	40	1.2					1.1
Affecting home life	32	1.3					1.2
High home life demands	5	0.41	0.104	0.33	0.079	0.095 ; 1.14	0.58
Work/study time		0.98	0.095	0.98	0.093	0.962 ; 1.00	0.99
Good relationship with superiors	126	0.49	0.206				0.73
Good relationship with colleagues	129	1.4	0.545				1.2
Computer use pattern (ref: 0)	81		0.568				
One 4 h period without a break	18	0.98					0.99
At least two 4 h periods without a break	37	0.75					0.86
Asthma	17	1.8	0.198	1.9	0.205	0.708 ; 5.00	1.3
Perceived stress	108	1.9	0.029	1.8	0.064	0.965 ; 3.50	1.4

4.2.2 Separate analysis for women and men

In addition to the original results in Paper II, separate analyses for the marginal model and the model for developing neck pain were performed to investigate whether identified risk and health factors differed between women and men (Table 11 and Table 12). For ongoing neck pain, this additional analysis was not performed, due to few cases when splitting the data.

Table 11. **Present neck pain:** Women and men in separate marginal logistic models (generalized estimating equation, GEE), using all three time points. Explanatory variables included in these models were included in respectively total model. Unexplained risk is when all explanatory variables (factors) are equal to 0. Single factor risk = all factors, except one, are set to zero. OR = odds ratio.

	Marginal model for neck pain					
Explanatory baseline variables	WOMEN Unexplained risk = 0.32			MEN Unexplained risk = 0.11		
	OR	P-value	Single factor risk	OR	P-value	Single factor risk
Breakfast eaten regularly	0.95	0.7456	0.30	0.53	0.0067	0.06
Smoking	2.4	0.0016	0.75	1.7	0.2428	0.19
Physical activity	1.0	0.9707	0.32	0.98	0.2183	0.11
High work/study demands (ref: not too high)		0.2396			0.0008	
Not affecting home life	1.2		0.38	1.7		0.20
Affecting home life	1.3		0.40	2.5		0.29
High home life demands	1.3	0.2984	0.41	1.2	0.5697	0.13
Good relationship with superiors	0.71	0.1247	0.23	1.6	0.1864	0.18
Good relationship with colleagues	0.95	0.8464	0.30	0.69	0.2988	0.08
Computer use pattern (ref: 0)		0.0682			0.1846	
One 4 h period without a break	1.1		0.36	0.78		0.09
At least two 4 h periods without a break	1.4		0.44	1.2		0.14
Perceived stress	1.6	0.0002	0.51	1.6	0.0133	0.19

The separate analyses for women and men for the marginal logistic model gave similar results as the total analysis regarding the direction of the associations. Some exceptions were that breakfast eaten regularly was now a protective factor for men, smoking was now a clear risk factor only for women, high work/study demands was now a clear risk factor only for men, and computer use pattern was not statistically significant for either men or women.

The difference in impact of risk and health factors on neck pain shown in Table 11 is only possible to judge based on the relative effect measure OR. The risk levels are calculated only for the specific combination of all explanatory variables set to zero, except for one factor. Therefore, Table 11 should mainly be viewed as providing information about what factors seem to have an effect (or not) on neck pain.

Table 12. Developing neck pain: Women and men in separate Markov transitional models. Analysis of pain period at 1-year follow-up, for those without pain at baseline. Unexplained risk is when all explanatory variables (factors) are equal to 0. Single factor risk = all factors, except one, are set to zero. OR = odds ratio.

Model for developing neck pain						
Explanatory baseline variables	**WOMEN** Unexplained risk=0.07			**MEN** Unexplained risk=0.12		
	OR	P-value	Single factor risk	OR	P-value	Single factor risk
High work/study demands (ref: not too high)		0.1131			0.8974	
Not affecting home life	1.8		0.12	0.94		0.12
Affecting home life	1.4		0.09	0.74		0.09
High home life demands	1.9	0.3737	0.13	1.9	0.2669	0.24
Good relationship with colleagues	1.9	0.2756	0.12	0.27	0.0059	0.03
Computer use pattern (ref: 0)		0.2383			0.0852	
One 4 h period without a break	1.4		0.10	2.1		0.26
At least two 4 h periods without a break	1.6		0.11	2.4		0.30
Asthma	1.2	0.7182	0.08	4.0	0.0046	0.50
Perceived stress	1.8	0.0314	0.12	1.5	0.2176	0.19

The separate analyses for women and men for the model for developing pain gave effects in the same directions as the total analysis, but now computer use pattern was not statistically significant and perceived stress was only a statistically significant risk factor for women. For men, factors now statistically significant were good relationship with colleagues (protective) and asthma (risk).

4.2.3 Adjusting for educational course

If including the variable educational course (medicine vs. IT) in the model for developing neck pain (only variables with $P<0.20$ were included in the final model) the results presented as OR (95% CI) would be as follows:

Women/men	OR=3.0	(2.00 ; 4.62)
Educational course (medicine/IT)	OR=0.5	(0.361 ; 0.807)
High home life demands (yes/no)	OR=2.2	(0.954 ; 5.12)
Asthma (yes/no)	OR=2.0	(1.00 ; 3.91)
Perceived stress (yes/no)	OR=1.9	(1.26 ; 2.87)

Women still had a higher risk of developing pain compared to men. Medical students seemed to have a lower risk than IT students to develop neck pain. High home life demands, asthma, and perceived stress were risk factors, with figures equal to the original results in Paper II. Including the variable educational course (medicine vs. IT) in the model for ongoing neck pain gave no indication of a clear difference between medical and IT students (OR=0.9 95% CI 0.492 ; 1.53).

4.3 Validation of simple neck pain questions (Paper III)

Of those reporting present neck pain in the upper part of the back/neck, about half (women: 57%; men: 46%) also reported present neck pain in the lower back or arms/hands. Of those reporting no present neck pain in the upper part of the back/neck, only 13% of women and 11% of men reported pain in any of the other locations.
Among women, 4% reported *present pain* in all five years, and 42% reported no *present pain* in any of the five years. Among men, the corresponding figures were 1%, and 72% respectively. Among women, 4% reported a *pain period past year* in all five years, and 38% reported no *pain period past year* in any of the five years. Among men, the corresponding figures were 1%, and 67% respectively. Among those reporting *present pain* in response to the baseline questionnaire, the distribution statistics of the number of days with continuous pain reported by women were: mean = 652, md = 60, Q1 = 5 and Q3 = 730; and by men were: mean = 631, md = 135, Q1 = 4 and Q3 = 912.

For women the analysis comparing *present pain* and *pain period past year* showed that stress was less for those with only present neck pain compared to those with only neck pain period, (diff = -0.24 95% CI -0.403 ; -0.073). Among men, the results showed differences in sleep disturbance for those with only present neck pain compared to those with only neck pain period, (diff = 0.35 95% CI 0.025 ; 0.683).

Subjects' perceptions of their neck pain in relation to decreased general performance (DP), in terms of the three categories "no pain", "pain without DP", and "pain with DP", were used as outcome variable. For both women and men present neck pain captured neck pain that affects general health, stress, and sleep disturbance, as it was associated with those variables (Table 13). For women, energy was also associated with present pain. A dose-response relationship was found between the three categories of pain and the outcome variables (Tables 13), except for energy among men, which was not associated with any of the categorical pain variables.

Table 13 Outcome variables in combination with decreased performance (DP), analyzed with random intercept linear models, except for general health that was analyzed with a random intercept logistic model. Parameter estimates are presented together with standard error (SE). Subjects: N_{women} = 600-627, N_{men} = 548-573, observations: N_{women} = 1478-2653, N_{men} = 1683-2382. For general health, analyzed with a logistic regression, the number of events were n_{women} = 614, n_{men} = 841.

	No present pain, present pain without DP, present pain with DP				
Outcome variables		WOMEN		MEN	
		Estimates	SE	Estimates	SE
General health (t)	Intercept	-0.9086	0.13	-1.0573	0.12
	Present neck pain (t) without DP	-0.6118	0.21	-0.9645	0.27
	Present neck pain (t) with DP	-1.9119	0.40	-1.3103	0.44
	Present neck pain (t-1) without DP	-0.6118	0.22	-	-
	Present neck pain (t-1) with DP	-1.0436	0.35	-	-
	Inter-individual variance	4.18		5.02	
Sleep disturbance (t)	Intercept	4.4	0.06	4.1	0.06
	Present neck pain (t) without DP	0.2	0.08	0.2	0.11
	Present neck pain (t) with DP	0.7	0.11	0.6	0.17
	Present neck pain (t-1) without DP	-	-	-	-
	Present neck pain (t-1) with DP	-	-	-	-
	Inter-individual variance	1.42		1.65	
	Intra-individual variance	1.32		1.15	
Stress (t)	Intercept	3.4	0.04	3.2	0.03
	Present neck pain (t) without DP	0.3	0.06	0.2	0.08
	Present neck pain (t) with DP	0.5	0.09	0.5	0.12
	Present neck pain (t-1) without DP	0.1	0.06	0.02	0.08
	Present neck pain (t-1) with DP	0.2	0.08	0.4	0.12
	Inter-individual variance	0.33		0.33	
	Intra-individual variance	0.67		0.52	
Energy (t)	Intercept	4.1	0.02	3.9	0.02
	Present neck pain (t) without DP	-0.06	0.04	-	-
	Present neck pain (t) with DP	-0.1	0.05	-	-
	Present neck pain (t-1) without DP	-		-	-
	Present neck pain (t-1) with DP	-		-	-
	Inter-individual variance	0.20		0.20	
	Intra-individual variance	0.40		0.40	

When analyzing duration of present neck pain, only the subgroup with present neck pain at baseline was included. The variable was grouped into four categories: pain for 1-7 days; pain for 8-90 days; pain for 91-365 days; and pain for >365 days (Figure 5). For both women and men, duration of present neck pain was associated with decreased performance (women: P<0.001, men: P<0.001). For men, duration of present neck pain was also associated with general health (P<0.024), sleep disturbance (P<0.035) and energy (P<0.008). For women, there was no statistically significant association with general health (P<0.195), sleep disturbance (P<0.095), stress (P<0.700) or energy (P<0.830).

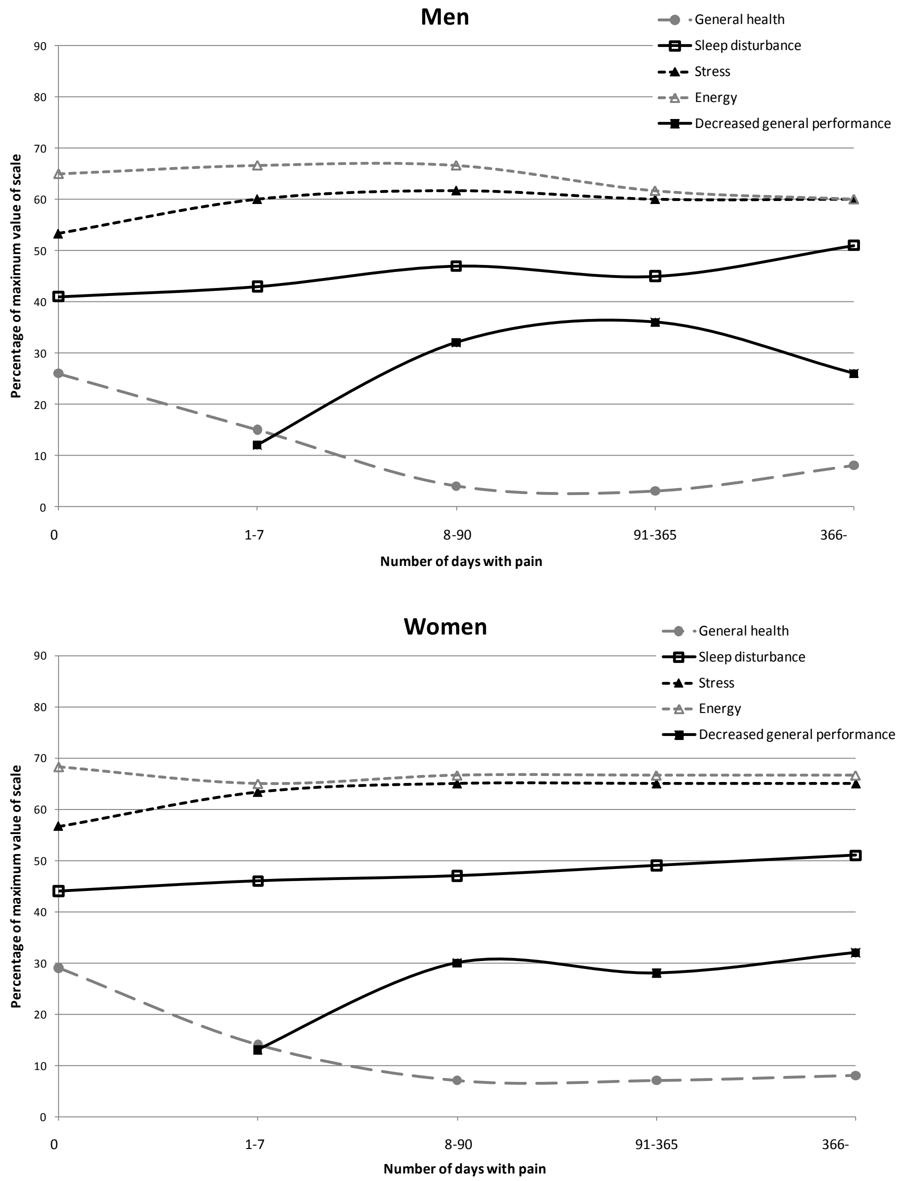

Figure 5. Duration of present neck pain. Results from longitudinal data analyzed using random intercept linear models, except for decreased performance, which was analyzed using a random intercept logistic model. Subjects: N_{women} = 331, N_{men} =151; observations: N_{women} = 574-675, N_{men} = 220-271. The outcome when no pain is present is also presented, except for decrease in general performance, but it was not included in the analysis.

4.4 Analyzing a repeated binary outcome

4.4.1 Paper IV

At baseline, 65% of the women and 47% of the men perceived stress. Among the women, 63% changed their reporting of stress over the five time points; among the men, the corresponding figure was 60%. Among the women, the results at baseline of computer work for >4 hours without a break were as follows: 66% responded zero times/week, 12% reported once/week, and 22% reported twice or more/week. For the men, the corresponding figures at baseline were 48%, 15%, and 37%. Among the women, 55% changed their computer use pattern over the five time points. Among the men, the corresponding figure was 63%.

Marginal logistic model

The marginal logistic model (GEE) analysis showed that, for women, the variables stress, lagged stress, and the binary version of computer use pattern were positively associated with neck pain (Table 14). For the men, the only variable associated with neck pain was stress (Table 14).

Table 14. Neck pain analyzed with marginal logistic model (generalized estimating equation, GEE). Parameter estimates, with 95% confidence intervals (CIs), for the final model for women and men.

	WOMEN		MEN	
No. of observations	1932		2437	
No. of events	498		283	
	β	95% CI	β	95% CI
Intercept	-1.62	-1.850 ; -1.393	-2.27	-2.489 ; -2.056
Stress	0.46	0.251 ; 0.674	0.47	0.207; 0.734
Lagged stress[a]	0.32	0.116 ; 0.529		
Computer use pattern[b]	0.45	0.217 ; 0.681		

[a] Stress from the previous year. Stress and lagged stress: no, yes. [b] Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.

The effect measures in Table 16 are based on the parameter estimates in Table 14. For the women, the effect measures indicate that stress had a greater impact on pain compared to the computer use pattern. Only presenting the effect measures (Table 16) means loss of information regarding the actual level of the risk of neck pain. These levels of the risk of neck pain, as functions of the different contributing factors stress and computer use pattern, are presented in Figure 6. There is an indication of possible biological interaction (RD effect-measure modification) for the combination of stress and computer use pattern for women (Figure 6).

Table 15. Risk difference (RD), risk ratio (RR), and odds ratio (OR), for the risk of neck pain, estimates based on a marginal logistic model (generalized estimating equation, GEE).
Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2437, number of events = 283.

		Effect measures					
		WOMEN			MEN		
Evaluating effect of -	Condition	RD	RR	OR	RD	RR	OR
Stress[a]	Computer use with breaks	0.14 0.087 ; 0.188	1.8 1.75 ; 1.92	2.2[c] 1.65 ; 2.91	0.05 0.017 ; 0.080	1.5 1.48 ; 1.55	1.6[c] 1.23 ; 2.08
	Computer use without breaks	0.17 0.100 ; 0.236	1.7 1.58 ; 1.85				
Computer use pattern[b]	No stress	0.07 0.009 ; 0.134	1.4 1.31 ; 1.57	1.6[c] 1.24 ; 1.98			
	Stress[a]	0.10 0.034 ; 0.170	1.3 1.16 ; 1.54				

[a]For the women, stress represented both stress (time *t*) and lagged stress (time *t-1*). For the men stress represented only stress at time *t*. [b]Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week. [c]Note that, in the absence of a product term in the logistic regression model, the OR for one explanatory variable is independent of the levels of other explanatory variables.

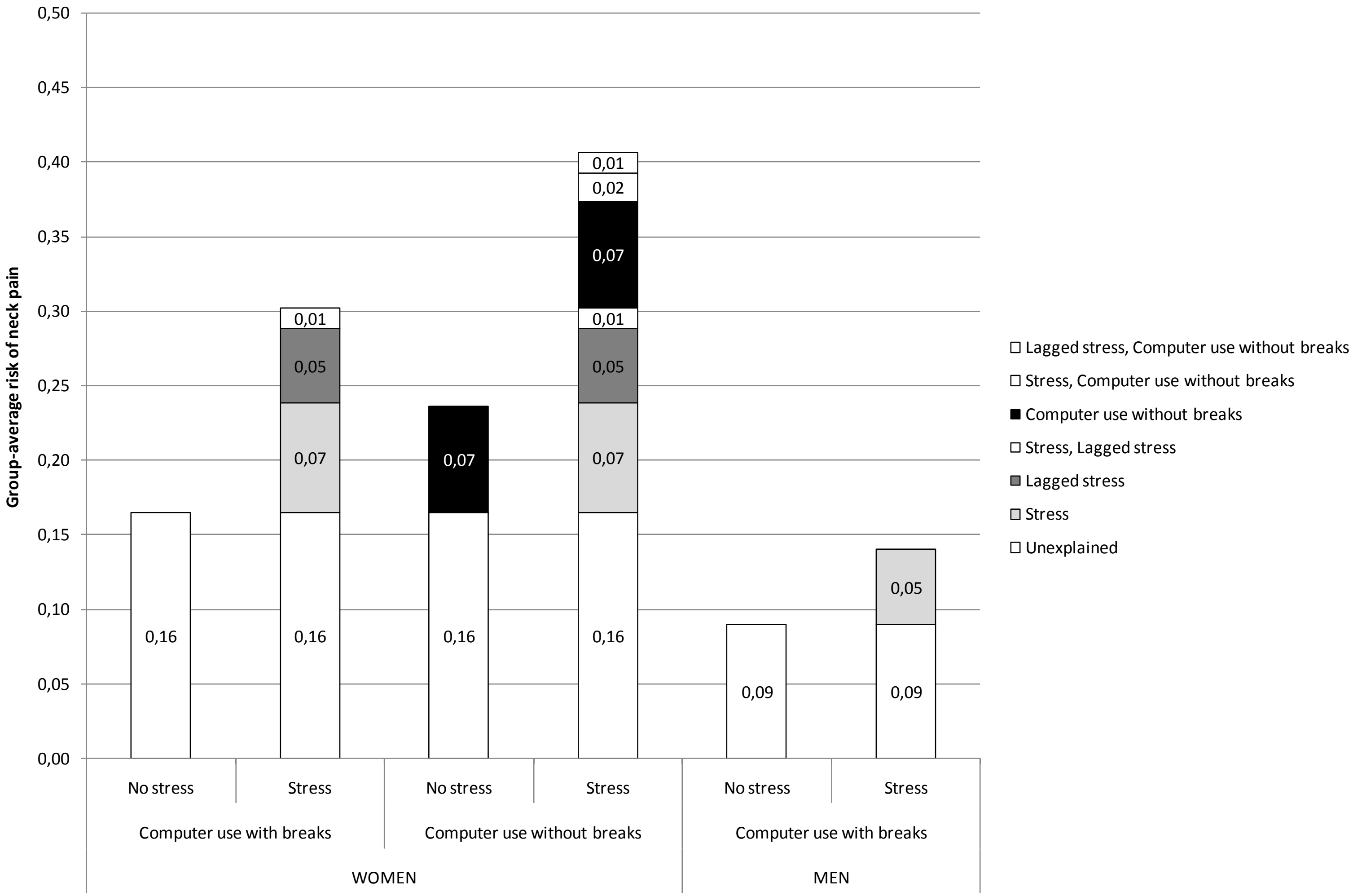

Figure 6. Group average risks for neck pain for women and men, based on the marginal logistic model (generalized estimating equation, GEE). Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2437, number of events = 283.

Random intercept logistic model

The analysis using the random intercept logistic model (GLMM) showed that, for the women, the variables stress, lagged stress, the binary version of computer use pattern, and time were positively associated with neck pain (Table 16). For the men, stress, computer use, and time were associated with neck pain (Table 16).

Table 16. Neck pain analyzed with random intercept logistic model (generalized linear mixed model, GLMM). Parameter estimates, with 95% CIs for the final model for women and for men.

	WOMEN		MEN	
No. of observations	1932		2422	
No. of events	498		283	
	β	95% CI	β	95% CI
Intercept	-1.9	-2.35 ; -1.46	-4.1	-4.66, -3.54
Stress	0.60	0.294 ; 0.897	0.76	0.380 ; 1.136
Lagged stress[a]	0.41	0.102 ; 0.722		
Computer use pattern[b]	0.64	0.294 ; 0.978	0.46	0.059 ; 0.856
Time	-0.15	-0.261 ; -0.038	0.10	-0.012 ; 0.205
Random intercept variance	2.5	1.62 ; 3.33	4.9	3.12 ; 6.75

[a] Stress from the previous year. Stress and lagged stress: no, yes. [b] Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.

Table 17, Table 18 and Figure 7 are based on a simplified model where the variable time was not included. This did not considerably change the parameter estimates. For the women, simplification resulted in the estimates -2.2852 (intercept), 0.6316 (stress), 0.4594 (lagged stress), 0.6352 (computer use without breaks), and the random intercept 2.4293. For the men, the simplification (excluding time) resulted in the estimates -3.8896 (intercept), 0.7317 (stress), 0.4449 (computer use without breaks), and the random intercept 4.8847. For the women, according to the absolute effect measure RD, the effect of stress was largest for those with a computer use pattern without breaks, Table 18. Note, however, that according to the RR, the effect of stress was smallest for those with a computer use pattern without breaks. For the men, the absolute effect of stress was small (RD = 0.02, RR = 0.03), but the relative effect of stress seemed of important size (RR = 2.0, OR = 2.1). Hence, the interpretation differs depending on whether an absolute or a relative effect measure was used.

Table 17. Subject specific risk difference (RD), risk ratio (RR), and odds ratio (OR), for the risk of neck pain, estimates based on a random intercept logistic model (generalized linear mixed model, GLMM).
Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2422, number of events = 283.

		Effect measure					
		WOMEN			MEN		
Evaluating effect of	Condition	RD	RR	OR	RD	RR	OR
Stress[a]	Computer use with breaks	0.14 0.086 ; 0.191	2.5 1.63 ; 3.37	2.9[c] 1.75 ; 4.14	0.02 0.006 ; 0.035	2.0 1.29 ; 2.78	2.1[c] 1.30 ; 2.86
	Computer use without	0.20 0.127 ; 0.274	2.2 1.52 ; 2.96		0.03 0.010 ; 0.052	2.0 1.29 ; 2.73	
Computer use pattern[b]	No stress	0.07 0.024 ; 0.114	1.7 1.23 ; 2.25	1.9[c] 1.25 ; 2.53	0.01 0 ; 0.022)	1.5 0.947 ; 2.14	1.6[c] 0.94 ; 2.18
	Stress[a]	0.13 0.057 ; 0.205	1.4 1.07 ; 1.17		0.02 0.0002 ; 0.043	1.5 0.951 ; 2.10	

[a]For the women, stress represented both stress (time *t*) and lagged stress (time *t-1*). For men stress represented only stress at time *t*. [b] Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.
[c] In the absence of a product term in the logistic regression model, the OR for one explanatory variable is independent of the levels of other explanatory variable

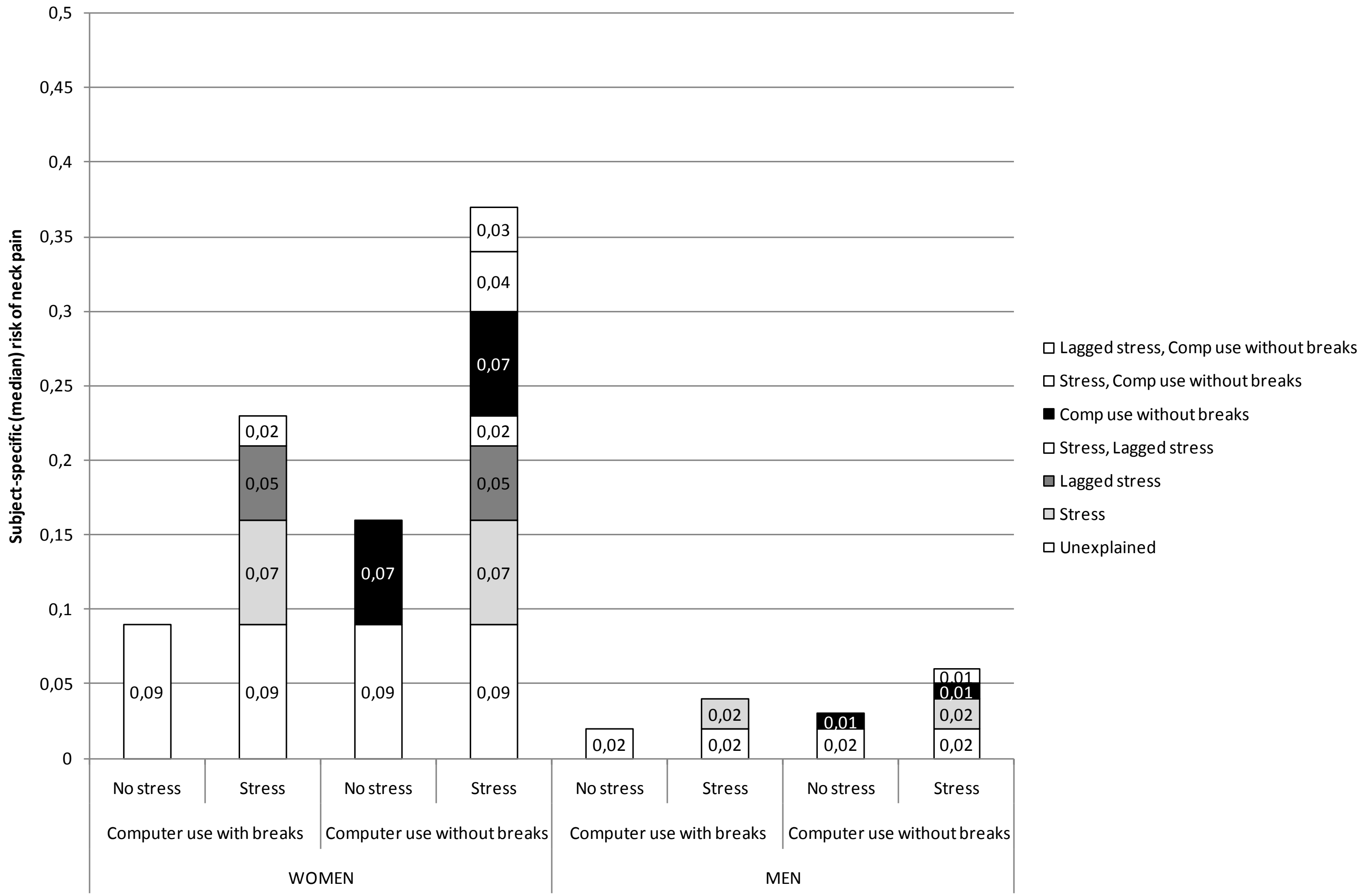

Figure 7. Subject specific risks of neck pain for women and men, based on the random intercept logistic model (generalized linear mixed model, GLMM). Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2422, number of events = 283.

As mentioned above, only presenting the effect measures means loss of information, and therefore the risk of neck pain as a function of the different contributing factors stress and computer use pattern is presented in Figure 7. There is an indication of possible biological interaction (RD effect measure modification) for the combination of stress and computer use pattern for women (Figure 7).

To illustrate the meaning of the random intercept the probabilities for neck pain, in different exposure groups, were calculated at the 25, 50 and 75 percentile level (Table 18).

Table 18. Subject specific risk of neck pain. Estimates based on the simplified (variable time excluded) random intercept logistic model (generalized linear mixed model, GLMM). Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2422, number of events = 283.

		WOMEN *P* (neck pain)		
Stress[a]	Computer use pattern[b]	25th percentile	50th percentile	75th percentile
No	With breaks	0.03	0.09	0.23
Yes	With breaks	0.09	0.23	0.46
No	Without breaks	0.06	0.16	0.35
Yes	Without breaks	0.17	0.36	0.62
		MEN *P* (neck pain)		
Stress[a]	Computer use pattern[b]	25th percentile	50th percentile	75th percentile
No	With breaks	0.005	0.020	0.083
Yes	With breaks	0.009	0.041	0.159
No	Without breaks	0.007	0.031	0.124
Yes	Without breaks	0.015	0.062	0.228

[a] For the women, stress represented both present stress and lagged stress. For the men, stress represented only present stress. P(neck pain): probability of neck . [b] Computer use pattern: with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.

For example, at least 25% of the young men were estimated to have an almost null risk for neck pain (<0.02) for any combination of the exposures. At the same time, 25% of the young men (75% quartile) were estimated to have a risk of neck pain of at least 0.23 if both stress and computer use pattern without breaks were present. Of the young women, at least 25% were estimated to have a risk of neck pain as high as 0.62 or higher if both factors were present.

4.4.2 Adjusting for previous neck pain

In addition to the original analysis in Paper IV, neck pain reported the previous year was included in both the marginal logistic model and the random intercept logistic model. Note that this is not equivalent to the analyses of the Markov transitional model used in Paper II and hence not interpretable as estimating risks of developing neck pain or having ongoing neck pain.

Marginal logistic model

When including neck pain reported the previous year in the model for women, lagged stress was not statistically significant and was not retained in the model. For the men, the results were similar to those without neck pain reported the previous year (Table 19).

Table 19. Neck pain: analyzed with marginal logistic model (generalized estimating equation, GEE). Parameter estimates, with 95% CIs for the final model for the women and for the men.

	WOMEN		MEN	
No. of observations	1932		2292	
No. of events	498		197	
	β	95% CI	β	95% CI
Intercept	-2.30	-2.512 ; -2.083	-2.71	-2.983 ; -2.443
Stress[a]	0.47	0.229 ; 0.713	0.51	0.186 ; 0.828
Computer use pattern[b]	0.35	0.086 ; 0.611		
Neck pain, previous year	2.55	2.297 ; 2.796	1.94	1.542 ; 2.337

[a] Stress: no, yes. P(neck pain): probability of neck . [b] Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.

Calculations based on the parameter estimates, Table 20, gave that the unexplained (previous neck pain = no; stress = no; computer use = with breaks) group average risk of neck pain, adjusted for previous neck pain, was 0.09 for the women and 0.06 for the men. Similarly, the group average risk for neck pain for those with previous neck pain was estimated to be 0.56 for the women and 0.32 for the men.

Table 20. Neck pain: analyzed with marginal logistic model (generalized estimating equation, GEE). Risk difference (RD) for the possible combinations of exposures. Women: Number of observations = 1932, number of events = 498. Men: Number of observations = 2292, number of events = 197.

		WOMEN		MEN	
Evaluating effect of	Condition	No previous neck pain RD	Previous neck pain RD	No previous neck pain RD	Previous neck pain RD
Stress[a]	Computer use with breaks	0.05 0.020 ; 0.074	0.11 0.040 ; 0.182	0.04 0.009 ; 0.065	0.12 0 ; 0.243
	Computer use without breaks	0.06 0.019 ; 0.103	0.10 0.034 ; 0.165		
Computer use pattern[b]	No stress	0.03 0 ; 0.068	0.08 0 ; 0.171		
	Stress	0.05 0.005 ; 0.089	0.07 0.006 ; 0.137		

[a] Stress: no, yes. P(neck pain): probability of neck . [b] Computer use pattern: computer use with breaks = never working on a computer for ≥4 hours continuously without a break, or working at most once/week without a break; computer use without breaks = working on a computer for ≥4 hours continuously without a break more than once/week.

For the women, the results, adjusted for previous neck pain, were similar to the original results when previous pain was present. When there was no previous pain the effect of stress and computer use pattern seemed smaller than in the original results (Table 21). For the men, the opposite way applied (Table 21). Their results were similar when there had been no previous pain; and when previous pain was present the effect of stress seemed larger in the adjusted analysis compared to the original analysis.

Random intercept logistic model

For both the women and the men the between-individual variance decreased considerably when previous neck pain was included in the model. For the men, it was estimated to be 2.3 in the adjusted analysis, compared to 4.9 in the original analysis. For the women, the estimated between-individual variance decreased from 2.5 to about zero, but for the women, this analysis was problematic as the Hessian matrix had at least one negative eigenvalue. The adjusted results for the women may therefore not be reliable and are not presented here. The unexplained (previous neck pain = 0; stress = 0; computer use = 0 or 1) subject specific risk of neck pain, adjusted for previous neck pain, was 0.03 for the men. The subject specific risk of neck pain for those with previous neck pain was estimated to be 0.12 for the men. For the men, the effect of stress was estimated to be 0.03 (no previous pain) and 0.09 (previous pain), which is similar to the original results.

5 Discussion

5.1 Musculoskeletal pain in the upper body

The prevalence of musculoskeletal pain in the upper body was high both among computer users (women: 36%; men: 16%) and among the group of university students (women: 28%; men: 12%). In the university cohort, the high prevalence was to some extent unexpected as the participants were young. They were therefore not expected to have accumulated exposure from risk factors or to have accumulated a long history of pain. Neither were they exposed to demands of family life connected to parenthood. High prevalence of neck pain in young age groups has been found previously, both in a group of college students (aged <35 years), with prevalence of neck pain of 40% and of pain in the upper back of about 14% (Noack-Cooper, Sommerich et al. 2009) and in the Swedish workforce aged 16-29 years, with prevalence of pain in the upper parts of the back or neck, at least 1 day per week, of 41% (women) and 24% (men) (Swedish Work Environment Authority 2008).

5.1.1 Women and men

Musculoskeletal pain in the upper body is more prevalent in women than in men, even among young adults. The women in the university cohort developed more musculoskeletal neck pain compared to the men, but in relative terms, the genders had almost equal risk of ongoing pain. Due to the higher risk of ongoing pain in general the absolute difference between the genders may not be as equal as it seems in relative terms. Some support for this explanation is given by unadjusted effect measures at 1-year follow-up; comparing women and men regarding the prevalence for ongoing pain gives PR = 1.4 and PD = 15, and for developing pain PR = 2.4 and PD = 12. If this is also true for confounder adjusted effect measures, cannot be answered by the analyses used in Paper II. Methods are needed that can calculate a single value for RD, but that is adjusted for several covariates.
In a review by the Task Force on Neck Pain and Its Associated Disorders, it was concluded that the majority of evidence showed higher prevalence of neck pain among women compared to men, though most of the reported prevalences were unadjusted (Hogg-Johnson, van der Velde et al. 2008). Differences between women and men in musculoskeletal symptoms were supported by a study of Computer-Assisted Design operators with identical work tasks, in which estimated PR’s (women/men) ranged from 1.4 for symptoms in low back, to 3.4 for symptoms in elbow (Karlqvist, Hagberg et al. 1996).

It could be argued that among young adults not yet having children and a complex household, women and men should have more equal prevalence of musculoskeletal symptoms, as the domestic demands and the time for relaxation and recovery would

not be as unequal as later in life (Strazdins and Bammer 2004; Nordander, Ohlsson et al. 2008). In this thesis, it was therefore hypothesized that, in a group of young university students, women and men would be more equal regarding neck pain. However, in the group of young university student, the prevalence of pain was higher among the women than among the men. Higher prevalence for young women has previously been reported in studies of adolescents (Keogh and Eccleston 2006), but not among school children (Murphy, Buckle et al. 2007).
Clear explanations for the gender differences were not found in this thesis, but some indications for further studies have been highlighted. Women and men may have different factors contributing to the development of musculoskeletal neck pain, as seen in the results from the additional analysis in Paper II. That women and men have different pain coping strategies has been indicated in a study of long-lasting pain among adolescents (Keogh and Eccleston 2006).

A small difference in prevalence of musculoskeletal disorders between women and men could be hypothesized to exist, due to biological differences (Cairns, Hu et al. 2001; Chesterton, Barlas et al. 2003; Wiesenfeld-Hallin Z. 2005; Nordander, Ohlsson et al. 2008).

5.1.2 Psychosocial factors, computer use, and lifestyle

The results of this thesis, of identified psychosocial factors, computer use, and lifestyle factors as risk and protective factors for musculoskeletal pain in the upper body (especially neck pain) could be explained by suggested pathways based on the integration of the epidemiological model (Figure 1) suggested by Bongers et al. (Bongers, de Winter et al. 1993; Bongers, Kremer et al. 2002) and the biopsychosocial model suggested by Melin and Lundberg (Melin and Lundberg 1997).

The results showed that high work/study demands was a risk factor, while work control and support from superiors were protective factors, for musculoskeletal pain in the neck and upper limbs, results that are supported by the psychosocial factors at work in the epidemiological model. High work/study demands negatively affecting home life could represent demands too high to allow for sufficient recovery after work, as suggested in the biopsychosocial model. The variable learning and developing at work was a protective factor, which could be viewed as a positive aspect of the consequence of the work organization, and as such supported by the epidemiological model.

To use computers during most of the work day (the female computer users in the workforce) and to use computer continuously for >4 hours without a break (university cohort) were risk factors for musculoskeletal pain in the upper body. Computer use pattern without breaks can be regarded as a negative health behavior

or a negative work style (Feuerstein, Nicholas et al. 2005). This behavior is probably due to several factors such as time pressure, high work demands, and a lack of understanding of the health consequences. Therefore, the computer use pattern could be seen as being intermediate to psychosocial factors and mechanical load, in the context of the epidemiologic model. Computer use pattern in this thesis was found to be a short-term risk factor for neck pain, which may indicate that it represents a physical exposure resulting in biomechanical work-load. This work-load is present only when the exposure is present, which was supported by the result showing that computer use pattern reported 1-year earlier did not seem to be associated with present pain. Computer use pattern was also found to be a risk factor for developing neck pain, but not a risk factor for ongoing neck pain. This is consistent with the reasoning of computer use pattern leading to a biomechanical work-load only present when the exposure, computer use pattern without breaks, is present.
The risk factor of working with a computer most of the work day may largely be a work organizational factor rather than involving a behavior, and is also assumed to lead to musculoskeletal pain in the upper body through pathways involving mechanical load.

Among the young adult women in this thesis work, perceived stress was more prevalent and had a higher impact on neck pain compared to the young adult men. Perceived stress was a risk factor associated with neck pain both for developing pain and for ongoing pain; and had an impact on both the group average risk (women and men) and the subject specific risk (women) of neck pain. Perceived stress and its impact on neck pain could be explained by the epidemiological model, as it is viewed as a mediator between psychosocial factors and musculoskeletal neck pain. In one study on female computer users it was shown that stress (mood) acted as a mediator between work demands and neck/shoulder symptoms (Larsman, Sandsjo et al. 2006). Among women, stress and computer use break pattern were found to interact, producing an effect of both factors together that exceeded the additive effect of each. Findings about interrelations, between psychosocial factors and physical load (external), are supported in previous studies (Bongers, Ijmker et al. 2006; Johnston, Jull et al. 2010).
In addition, the results of perceived stress affecting neck pain both close in time and as a long-term effect could be explained according to the biopsychosocial model (Melin and Lundberg 1997), which suggests that in the post-work activity, full recovery of the stress response may not occur quickly and is dependent on, for example, household work and child care. Perceived stress is not an exposure in the same sense as computer use pattern, and stress is not turned “on or off” (Lundberg 2002). It is connected to physiological stress reactions affecting sensitization and dysfunctional inhibition (Melzack 1999; Lundberg 2002). Perceived stress could lead to nociception, e.g., through tense trapezius muscles, and an important part of its

impact on pain may also be through the physiological response processes and sensitization (Lundberg 2002; Bongers, Ijmker et al. 2006).

Identified life-style factors having an impact on neck pain were smoking and eating breakfast regularly, but life style is not incorporated in the epidemiological model by Bongers et al (Bongers, de Winter et al. 1993). The effect of life style on musculoskeletal neck pain could be explained in terms of the theory proposed by Melin and Lundberg (Melin and Lundberg 1997), according to which the post-work activities could either prolong or speed up the recovery from work-related stress response.

The above findings, regarding the identified psychosocial factors, computer use, lifestyle factors, and the direction of these associations are consistent with systematic reviews concerning pain in the upper body (Ariens, van Mechelen et al. 2001; Bongers, Kremer et al. 2002; Bongers, Ijmker et al. 2006; Haldeman, Carroll et al. 2010) and with results in resent studies (Harcombe, McBride et al. 2010; Johnston, Jull et al. 2010).

In this thesis, it is suggested that the concept of allostatic overload is seen as an internal imbalance between internal stimuli and recovery, leading to ill health (McEwen 2008). The concept from the balance theory model of job design and stress, mentioned in the introduction, is seen as an external imbalance between external stimuli and recovery, leading to musculoskeletal symptoms (Smith and Sainfort 1989; Carayon, Smith et al. 1999). It follows that neck pain could be seen as the effect of imbalance at these different levels, including both work/studies and home-life. Computer use without breaks being a more important risk factor for neck pain than computer use duration may be interpreted as supporting the importance of recovery in relation to external stimuli. The results regarding perceived stress, as an important risk factor for neck pain, supports the concept of allostatic overload between stimuli and recovery.

The combined epidemiological and biopsychosocial model of Bongers et al. and Melin and Lundberg gives a good understanding and frame-work for the findings in this thesis interpretable along pathways from exposures to neck pain. Using the theory of allostatic overload and the balance theory model of job design and stress in combination gives a usable framework for explaining and understanding the results about light physical exposures, perceived stress and the role of recovery.

5.2 Validity of simple neck pain questions

Presence of neck pain, determined from the answer to a simple neck pain question, is related to lower levels of perceived general health, sleep disturbance, stress, and decreased general performance, at least among young university students. At the same time, it appears that a large proportion of the neck pain reported among the young university students did not seriously affect their health and perceived general performance. In the literature, it is discussed if people are reporting discomfort rather than pain, especially in modern society (Hadler and Carey 1998). This is contradicted in a study comparing the prevalence of musculoskeletal complaints in a native population living under primitive conditions with a representative sample of the Norwegian population (Eriksen, Hellesnes et al. 2004). From this the conclusion is drawn that the high prevalence of for example musculoskeletal complaints is not specific for industrialized societies.

As mentioned in the Introduction, most validations of self-reported pain questions, based on the NQ, used specific clinical diagnosis as a gold standard. This is a confusing validity as the perception of the patient is the gold standard, according to the widely accepted pain definition (IASP) stated earlier. Therefore, validation of pain assessments should include comparisons with other measures of the intended characteristics of pain. According to the work of the Task Force on Neck Pain and Its Associated Disorders diagnostic procedures for non-traumatic neck pain have not been proven valid or useful, but they conclude that self-assessment questionnaires can be useful in management and prognosis, if reliable and valid (Haldeman, Carroll et al. 2008).

One suggested scale of neck pain, with four grades, has been defined by the Task Force on Neck Pain and Its Associated Disorders (Guzman, Hurwitz et al. 2008). The focus in this scale is on physiological findings (impairment in body function or body structure) rather than a grading of possible disability due to pain. Both this proposed scale and the scales mentioned previously in the Introduction (e.g., PDI and MPQ) are suitable for research or clinical settings different from the one used in this thesis, as they are not down-sized as needed in the discussed situation.

In the development of down-sized pain instruments or few-item pain questionnaires for epidemiological population studies, it would be helpful to have good knowledge of lay people’s view of pain and, specifically, musculoskeletal pain. Few studies give this information, but some confirm the multi-dimensionality of pain; and lay people’s view on the impact of pain on life and pain management (Johansson, Hamberg et al. 1999). Through experiences of pain in our lives we personally learn the concept of pain. Therefore, the experience of pain is an interaction between biology and culture (Johansson, Hamberg et al. 1999). Pain can also be of widely differing intensity and seriousness, and this was seen even in the young population, of university students. In a broader population the effect of pain on everyday life could vary even more.

5.3 Regression models for binary outcomes

From Papers II and IV, it could be argued that by using different regression models, different aspects of neck pain patterns could be addressed and the risk factors' impact on pain pattern could be identified. The Markov transitional logistic model can provide answers to questions about factors affecting the development and reoccurrence of neck pain. The subject specific model (random intercept logistic GLMM) gave a detailed understanding of the impact of risk factors on musculoskeletal neck pain, compared to the group average model (marginal logistic GEE). The group average model (marginal logistic GEE) and Markov transitional logistic model used in Paper II have previously been used with longitudinal data in studies of musculoskeletal pain, but the additional knowledge achieved by using both of these models was not thoroughly discussed previously (Viikari-Juntura, Martikainen et al. 2001; Yu, Morgenstern et al. 2003).
In Paper IV, where the two models marginal logistic model (GEE) and a random intercept logistic model (GLMM) were used, the results for men differed between the marginal and the random intercept model, indicating higher impact at a group level. This emphasizes the need to consider which of the two models is appropriate in specific studies and the importance of clear statement of whether results are group averages or subject specific, when presenting results. In addition, the results shows that depending on the relation between the amount of between-individual variation and the absolute effect of the risk factor, the group level analysis can hide important information about the effect of the risk factor on the outcome. The marginal logistic model is often suggested as an appropriate model when focusing on population or group averages (Rothman, Greenland et al. 2008).

If between-individual variance is large further research could be aimed at identifying additional risk or health factors explaining this variance. For example, including previous pain in the random intercept models substantially decreased the between-individual variance. The effect of decrease in the exposure could still be presented at the group level. If the variation of the random intercept is small, this means that the individual probability curves are close to the population average probability curve.

5.4 Epidemiological effect measures

The relative effect measures (OR, RR) did not differ as much between women and men as the absolute effect measure (RD) did. The interpretations of the results for women compared to men diverged depending on which effect measure was used.

The choice of effect measure has been extensively discussed in the literature (Nurminen 1995; Zocchetti, Consonni et al. 1997; Thompson, Myers et al. 1998; Localio, Margolis et al. 2007), but to the authors' knowledge, it has not been discussed for longitudinal studies. The present paper therefore contributes by discussing the choice of effect measure in a longitudinal study with common and recurrent outcomes. When choosing between effect measures, it is important to determine whether an absolute or relative measure is more clinically relevant in the specific situation. Risks divided into contributions from different exposures, may be more appropriate effect measures than ratios in applications like the present, where it may be more desirable to reduce the number of events among those with high prevalence than among those with low prevalence. Based on the results of this thesis it is recommended to including both the effect measure of choice and the actual probabilities for different exposure combinations in presentation of results.

The biological interaction, when defined according to the sufficient cause model, can only be evaluated with absolute effect measures on the probability scale (Rothman and Greenland 1998). In the random intercept model no product terms were significant. The absence of a product term in the regression model (statistical interaction on a logit scale) does not imply absence of interaction in all effect measures. In the present study we saw an interaction between stress and computer use pattern in the absolute effect measure (probability difference), but not in the relative effect measure OR.

5.5 Strengths and limitations

The variables in the present thesis are all self-reported, and hence, there is a risk of reporting bias. This bias can arise for example, from individuals with symptoms being more likely to remember high exposures than individuals without symptoms. This type of bias could be decreased to a certain degree if the outcome is related to explanatory variables reported both at the same time as the outcome, and at an earlier time point, as in Papers III and IV. In a study of white-collar government employees Stradezins et al. (Strazdins and Bammer 2004) saw no evidence of reporting bias. Self-reported workplace exposures were not influenced by whether the individual had musculoskeletal symptoms or not. The type of self-reported neck pain outcome used in the present thesis is noted to be sensitive to effects of seasonal variation at different follow-ups (Takala, Viikari-Juntura et al. 1992). In Papers II-IV, data collection took place at the same time of year to avoid this seasonal effect.

The term “pain/ache” used in this thesis describes a broader concept than does the word “pain” on its own. Gaston-Johansson (1985) concludes from one linguistic study and one questionnaire study that pain is described as dynamic, transient, sudden, usually discontinuous, and distinctly located experience, while ache is characterized as continuous, slow, grinding and involving large body surfaces (Gaston-Johansson 1984; Gaston-Johansson and Allwood 1988). Both these concepts are termed as “pain”, but when only using the word “pain” a respondent may answer to a question narrower in definition than intended. Therefore, if this is a common

way among lay people of describing and making a distinction between “pain” and “ache”, asking about pain/ache captures both these characteristics of pain.
The assessments of musculoskeletal pain used in this thesis mainly lack information about the disability or burden of the pain. One exception is Paper III, where one variable used is the presence of neck pain in combination with decreased general performance.

One limitation of this thesis was the restricted choice of variables. In particular in Paper III there were insufficient data available for specifically assessing Quality-of-Life and well-being. The pain assessments could therefore only be validated through associations with questions about perceived general health, sleep disturbance, stress/energy and perceived decrease in general performance.

In Paper I those defined as “computer users” used the computerized equipment for a minimum of half their working hours, so there is also reason to be concerned about how the non-computer work, which occupied their remaining work hours, affected them. Analyses including additional variables describing body postures during work, such as ”bent forward”, ”twisted” and ”hands abreast of, or above the shoulder height” did not change the results. The estimated ratios usually differed in the second or third decimal, compared to when these variables were not included. In a few exceptions the estimated ratios differed by one unit in the first decimal.

Stress could be a mediator between psychosocial work conditions and musculoskeletal pain in the neck (Wadman C and Kjellberg A 2007), but in the present thesis work, perceived stress is not assumed to be a total mediator. This kind of analysis is not included in the present thesis, but this is not a major issue as long as the possible mediation of perceived stress is remembered when interpreting the results. For example, the results in Paper II regarding the association between demands and neck pain have to be interpreted as the direct impact of demands on neck pain, not including the possible impact of the mediator perceived stress.

In this thesis work, most of the analyses are based on logistic regression models. One alternative could have been to use log-binomial regression models, but this was not chosen mainly due to problems with convergence. Reasons for the experienced problems with convergence were probably that the outcome “pain” is common; and in some models both continuous and polychotomous covariates were included (Yu and Wang 2008). The log-binomial model, contrary to the logistic model, is also problematic when estimating probabilities close to 0 or 1 (Yu and Wang 2008), which was the case in some situations in the present thesis.

In none of the papers the analysis is truly causal. In Paper I this is obvious as the design is cross-sectional. In Paper II and IV the design is longitudinal, but the analyses are investigating temporal effects (one of Hills necessary, but not sufficient causal criteria), and these are not necessarily truly causal effects. In the longitudinal studies one possible approach for causal inference is based on the counter factual model, using inverse-probability weighting, as discussed by e.g. (Rothman, Greenland et al. 2008).

One of the aims of Paper II was to classify factors as having long-term or short-term effects. In order to allow interpretation of the explanatory variables as long-term or short-term effects, the stability of these variables over time was investigated. High work demands, computer use pattern, and perceived stress varied over the years, and could therefore be investigated regarding short-term and long-term. Factors that were not changeable (gender), or that could vary but did not in the present material (regularly eating breakfast, smoking) could not be investigated concerning long- or short-term effects.

In Paper III, generalizations from the results should be made with caution, because the respondents, being young university students enrolled in academic studies, formed a group with quite specific characteristics. The reported neck pain in the present young age group may be less severe than, and may not affect life in the same way as, reported pain in an older age group. This could possibly explain the not so clear differences in the outcomes between pain and no pain, and may point to the necessity of an additional question as discussed previously. A ”yes” from a not so affected individual is not possible to distinguish from a ”yes” from a more affected individual, which may explain the high prevalence of pain in young age groups. The generalization of results to other groups of young adults may be possible. The homogeneity of the university cohort allowed for investigating psychosocial factors and computer use patterns with less confounding by other factors, such as high physical work load, and high demands from complex home and family life. Furthermore, studies using population-based material and material with other outcome variables, such as mental well-being and number of sick-leave days, would increase the knowledge of the validity of simple pain questions used in epidemiological studies, and would increase the quality of conclusions drawn from such studies.

In Papers II-IV, one outcome was to have had neck pain for more than 7 consecutive days during the last 12 months. There has been some debate as to whether a 12-month period may be too long period about which it is reasonable to ask questions about pain. Some studies (Orhede 1994; Brauer, Thomsen et al. 2003) that have compared the results of surveys conducted 3 months and 12 months after experiencing pain, recommend the shorter period. In this thesis this is not viewed as a serious problem as in the validation study, Paper III, it was shown that the results for validity of the 12 months neck pain period was quite similar to the validity of present neck pain.

In Papers II-IV, which are based on longitudinal data, partial missing over time is a issue of concern. The regression methods used in these papers allow for partial missing over time without excluding the individuals concerned. This is a strength compared to methods where individuals with partially missing values are completely excluded from the analysis. Still, partially missing values can be a problem, for example if the response rate decreases considerably over time. There are proposed methods for handling this issue, e.g., imputation (Furlow, Fouladi et al. 2007; Yang, Li et al. 2008). As there are several different imputation methods, possibly not giving the same result, a sensitivity analysis has to be considered. In the present thesis, there were partially missing values, but the response rates over time was high, and no imputation methods have been used.

In Paper IV, estimates were calculated for RR and RD and the different absolute contributions to risk of neck pain. A possible limit is that the 95% CI for these estimates was approximate CIs, based on the Delta method (Billingsley 1986) applied in SAS (CARY, NC, USA). The properties of the Delta method-based 95% CIs need to be evaluated in the specific setting with data of repeated measures of a common outcome.

In Paper IV, only two exposures were considered, namely, stress and computer use pattern. When more exposures or confounders are added to the models, the number of combinations of levels of the explanatory variables rapidly increases. This complicates the presentation and interpretation of results based on RR or RD measures. There are suggested methods for adjusting the RR and RD for several confounders, but they are not directly implementable for repeated measures (Flanders and Rhodes 1987; Beaudeau and Fourichon 1998). In future studies when presenting risk differences and absolute contributions of risk factors for neck pain as proposed in Paper IV, it may be of interest to use inverse probability treatment weighting (IPTW) to adjust for several confounders (Rothman, Greenland et al. 2008).

5.6 Implications

The results showed differences between women and men regarding prevalence of, and risk for, musculoskeletal pain in the upper body. The results also indicated that perceived stress was more prevalent among women and also possibly had a larger impact on women's musculoskeletal neck pain. If verified and expounded in future research, these findings should be taken into consideration when planning preventive work on musculoskeletal disorders, and could also have an effect on decisions regarding workmen's compensation and health insurance.

Due to the findings mentioned above women and men should be analyzed separately in research on musculoskeletal pain in the upper body.

The choice between a marginal logistic model and a random intercept logistic model should be determined by whether group average risks or subject specific risks are relevant to a specific study.
In future studies, care should be taken not to routinely present results only as associations between risk factors and musculoskeletal pain, but to clearly state whether the effect measure chosen is relative or absolute. When presenting results of a logistic regression, the parameter estimates should always be included together with a specific effect measure, which is preferably either a RR or a RD depending on the approach of the specific study.

6 Conclusions

General conclusion
The prevalence of musculoskeletal pain in the upper body was higher in women than in men, even among young adults. Women in the university cohort developed more musculoskeletal neck pain than young men. Among young adult women, perceived stress was more prevalent and had a higher impact on neck pain compared to young adult men. The relative effect measures (OR, RR) did not differ as much between women and men as the absolute effect measure (RD).

Specific conclusions
Women had a higher prevalence of musculoskeletal symptoms in the neck and upper limbs compared to men, even when controlling for occupational group, computer use and psychosocial factors. In addition, women had a higher risk of developing neck pain, even after adjusting for confounders in the group of young adult university and college students, but the genders were closer regarding (relative) risk of ongoing pain.

Risk factors among computer users in the workforce were high work demands; for women, and computer use during most of the work day. Health factors were variables connected to work control and to viewing work as providing career development and education; and support from superiors (women). In the university cohort perceived stress, high work/study demands and computer use break pattern were identified as risk factors. Perceived stress was a risk factor associated with neck pain, both for developing and for ongoing pain; and had an impact on both the group average risk (women and men) and the subject specific risk (women) of neck pain. Among the women stress and computer use break pattern were found to interact, producing an effect where both factors present exceeded the additive effect of each. Perceived stress had a greater impact on neck pain among young adult women compared to men.

Simple questions, about present neck pain and neck pain period past year, captured features of pain such as perceived general health, sleep disturbance, stress, energy (women) and general performance, in the university cohort. Neck pain period past year did not reflect a more serious pain than present pain. Although, duration of present pain appeared to be related to general performance, there was no evidence of a dose-response pattern as hypothesized.

For the young adult men, the marginal model indicated a higher impact of exposures on the group average risks of neck pain, than the random intercept model indicated about subject specific risks. The choice of statistical model should be based on whether a group average risk or a subject specific risk is of clinical relevance.

Results regarding differences between women and men diverged depending on which effect measure was used.

7 Future work

More knowledge is needed on differences in impact of work-related risk factors related to musculoskeletal pain between women and men. Further research is needed to better understand the causes of musculoskeletal pain to develop effective and efficient preventive and rehabilitative methods at group and individual levels. More research is also needed to gain knowledge about the effect on neck pain of a combination of risk factors.

Future work should include development of additional simple questions about neck pain characteristics to increase the specificity of the epidemiological assessment of pain prevalence in questionnaires. Simple questions about neck pain should as a minimum include a question on the prevalence of neck pain and a question on the effect of neck pain on everyday life. Validation of such simple neck pain questions should be done in general populations.

The confidence intervals for RR and RD, here calculated based on the Delta method, could probably be improved using other techniques. One of the reasons researchers hesitates to use RR and RD, instead of OR, could be the lack of easily available statistical methods for evaluating these effect measures.

It would be interesting to investigate possible causal analysis methods, to improve the knowledge about risk factors for neck pain and possible pathways. Such causal methods are, for example, directed acyclic graphs (DAG:s) and counter-factual models (Rothman, Greenland et al. 2008).

8 Sammanfattning

Muskuloskeletal smärta/värk är vanligt förekommande i många länder och en av de vanligaste orsakerna till långtidssjukskrivning . Enligt IASP (International Association for the Study of Pain) är kostnaderna till följd av muskuloskeletal smärta/värk på andra plats efter hjärt-, kärlsjukdom. Det övergripande syftet med denna avhandling är att öka kunskapen om muskuloskeletal smärta i övre delen av kroppen i lätt fysiska arbeten i relation till kön, psykosociala faktorer och datoranvändning, och att jämföra olika statistiska metoder för att analysera vanligt förekommande och återkommande binära utfall.

Två grupper undersöktes utifrån enkätdata: (a) datoranvändare i den svenska arbetande befolkningen, (b) en kohort av universitetsstudenter. I den statistiska analysen användes: logistisk modell, Cox modell (för att beräkna prevalenskvoter), marginal logistisk modell(GEE), logistisk modell med slump-intercept (GLMM), logistisk Markov modell och Poisson modell. Effektmått som användes var oddskvoter, riskkvoter och risk differenser.

Muskuloskeletal smärta/värk i nacke och övre extremiteter var vanligare hos kvinnor än hos män, till och med bland unga vuxna. Riskfaktorer för muskuloskeletal smärta/värk hos datoranvändare i den arbetande befolkningen var höga arbetskrav och att använda datorn större delen av arbetsdagen (kvinnor). Skyddande faktorer var kontroll i arbetet och att ha ett utvecklande arbete. För kvinnor var även socialt stöd en skyddande faktor. I universitetskohorten identifierades arbets/studie-krav och datoranvändning utan pauser som riskfaktorer för nacksmärta. Upplevd stress var en risk faktor både för att utveckla nacksmärta och för att få återkommande nacksmärta. Upplevd stress påverkade förekomsten av nacksmärta både på grupp- och individnivå.

Endast bland kvinnorna påverkade datoranvändning utan pauser nacksmärtan på individnivå, men på gruppnivå påverkade datoranvändningen både kvinnors och mäns nacksmärta. Bland kvinnorna sågs en samverkan mellan de två faktorerna upplevd stress och datoranvändning utan pauser, dvs. vid samtidig närvaro av faktorerna blev effekten på nacksmärta större än summan av de enskilda effekterna. Enkla frågor, om nuvarande nacksmärta och nacksmärta senaste året, fångade aspekter av smärt såsom upplevd generell hälsa, sömnproblem, stress och generell prestation. Nacksmärta senaste året tycktes inte speglade en allvarligare smärta än nuvarande nacksmärta gjorde.

Valet av statistisk modell bör baseras på om risken på gruppnivå eller om risken på individ nivå är mest kliniskt relevant i en specifik studie. Kvinnors och mäns risk för nacksmärta skiljde sig mer åt mätt i absoluta effektmått än mätt i relativa effektmått. Det kausala förhållandet mellan riskfaktorer och nacksmärta kan skilja sig åt mellan kvinnor och män.

9 Acknowledgements

It is a wonderful feeling to have reached this part of the thesis work. Thank you to everyone who made this possible for me, and to everyone who have supported me.

A special thank you to:

Mats Hagberg - my main supervisor, for generously sharing your expertise, for making this journey possible for me, and encouraging me to do this.

Eva M Andersson - co-supervisor for your careful reading and for fruitful statistical discussions.

Sara Thomée – my colleague, for all occasions when you have taken your time to listen to me, and for all our discussions of big and small issues about work, life and children.

Ingela Nyth - for looking after me in a great way and for making jokes and laughs present in our corridor.

Ewa Gustafsson - for being a wonderful colleague and for all your help and good advice.

Jesper Löwe – my former colleague. Thank you for interesting and rewarding discussions, and for broadening my awareness in issues such as gender, work ability and life.

Linda Åhlström, Christina Ahlstrand, Lotta Dellve and Maria Boström - for being great colleagues and all your encouragements.

Adnan Noor - my new statistical colleague. Thank you for your great social support in the last part of this thesis work, and for reminding me that lunch is important even when writing thesis.

Helena Sandén – it was great to share the moments of frustration and of happiness in the final sprint of this thesis work, with someone in the same situation.

Andreas Jonsson, Agneta Lindegård and Emina Hadzibajramovic - for valuable comments on my work and good discussions.

To everyone in the research group “Workability and Ergonomics” who has helped with gathering data, been available for discussions, and for good coffee breaks; for being patient with my notes on the door: “Still working on my thesis. No statistical questions allowed!”

All my colleagues at the Department of Occupational and Environmental Medicine, you are the best colleagues one can have.

Mari Lundberg - for being such a wonderful friend and supporting me and encouraging me to keep on working on this project.

Magdalena, Ann-Sofie, Margareta and Linda, in the “syjunta” - our wonderful moments are very important to me. Sewing or not sewing, but always eating well and sharing the ups and downs of life with each other.

Peter Jonasson at Geson Hylte Tryck - for your positive attitude, your ability to find solutions, and for all your patience with me despite my changes in the last minute.

Ulla, my mother - for letting me grow-up in an environment full of possibilities, encouraging me to try new challenges, showing me that nothing is impossible and for always believing in me.

Gunnar, my father - the most curious person I know and with a never ending thirst for knowledge. Thank you for all your help, ranging from taking care of your grandchildren to valuable discussions about my thesis.

Mats, my father in law - for making me see this thesis in a new light and making me proud of what I do.

Claes - for being my secure and safe rock in life and for all the support you have been given me during these last months: taking care of the family and giving me space to struggle with my work in evenings and on weekends. I love you.

Ellinor, Simon, Johan and Laura - your presence has helped me to see that, even if this project is a major thing for me, it is not the meaning of life. You mean so much to me.

Thank you also to the Swedish Council for Working Life and Social Research for financial support.

10 References

Agresti, A. (2002). Categorical data analysis. New York, Wiley-Interscience.
Ariens, G. A., W. van Mechelen, et al. (2001). "Psychosocial risk factors for neck pain: a systematic review." Am J Ind Med **39**(2): 180-193.
Ashcroft, G. S. and J. J. Ashworth (2003). "Potential role of estrogens in wound healing." Am J Clin Dermatol **4**(11): 737-743.
Axelson, O. (1994). "Some recent developments in occupational epidemiology." Scand J Work Environ Health **20 Spec No**: 9-18.
Axelson, O., M. Fredriksson, et al. (1994). Use of the prevalence ratio v the prevalence odds ratio as a measure of risk in cross sectional studies. **51:** 574-.
Beaudeau, F. and C. Fourichon (1998). "Estimating relative risk of disease from outputs of logistic regression when the disease is not rare." Preventive Veterinary Medicine **36**(4): 243-256.
Billingsley, P. (1986). Probability and Measure. New York, John Wiley & Sons, Inc.
Bjorksten, M. G., B. Boquist, et al. (1999). "The validity of reported musculoskeletal problems. A study of questionnaire answers in relation to diagnosed disorders and perception of pain." Appl Ergon **30**(4): 325-330.
Blizzard, L. and D. W. Hosmer (2006). "Parameter estimation and goodness-of-fit in log binomial regression." Biom J **48**(1): 5-22.
Bongers, P. M., C. R. de Winter, et al. (1993). "Psychosocial factors at work and musculoskeletal disease." Scand J Work Environ Health **19**(5): 297-312.
Bongers, P. M., S. Ijmker, et al. (2006). "Epidemiology of work related neck and upper limb problems: psychosocial and personal risk factors (part I) and effective interventions from a bio behavioural perspective (part II)." J Occup Rehabil **16**(3): 279-302.
Bongers, P. M., A. M. Kremer, et al. (2002). "Are psychosocial factors, risk factors for symptoms and signs of the shoulder, elbow, or hand/wrist?: A review of the epidemiological literature." Am J Ind Med **41**(5): 315-342.
Bonica, J. (1979). "Editorial The need of a taxonomy." Pain **6**(3): 247-252.
Boos, D. D. (1992). "On Generalized Score Tests." The American Statistician **46**(4): 327-333.
Brauer, C., J. F. Thomsen, et al. (2003). "Can we rely on retrospective pain assessments?" Am J Epidemiol **157**(6): 552-557.
Buckle, P. W. and J. Jason Devereux (2002). "The nature of work-related neck and upper limb musculoskeletal disorders." Applied Ergonomics **33**(3): 207-217.
Cairns, B. E., J. W. Hu, et al. (2001). "Sex-related differences in human pain and rat afferent discharge evoked by injection of glutamate into the masseter muscle." J Neurophysiol **86**(2): 782-791.
Carayon, P., M. J. Smith, et al. (1999). "Work organization, job stress, and work-related musculoskeletal disorders." Hum Factors **41**(4): 644-663.
Chesterton, L. S., P. Barlas, et al. (2003). "Gender differences in pressure pain threshold in healthy humans." Pain **101**(3): 259-266.
Cote, P., G. van der Velde, et al. (2008). "The burden and determinants of neck pain in workers: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders." Spine (Phila Pa 1976) **33**(4 Suppl): S60-74.
Davies, H. T. O., I. K. Crombie, et al. (1998). When can odds ratios mislead? **316:** 989-991.
Deeney, C. and L. O'Sullivan (2009). "Work related psychosocial risks and musculoskeletal disorders: potential risk factors, causation and evaluation methods." Work **34**(2): 239-248.

Dellve, L., M. Lagerstrom, et al. (2003). "Work-system risk factors for permanent work disability among home-care workers: a case-control study." Int Arch Occup Environ Health **76**(3): 216-224.
Eller, N. H., B. Netterstrom, et al. (2009). "Work-related psychosocial factors and the development of ischemic heart disease: a systematic review." Cardiol Rev **17**(2): 83-97.
Elo, A. L., A. Leppanen, et al. (2003). "Validity of a single-item measure of stress symptoms." Scand J Work Environ Health **29**(6): 444-451.
Eriksen, H. R., B. Hellesnes, et al. (2004). "Are subjective health complaints a result of modern civilization?" Int J Behav Med **11**(2): 122-125.
Feuerstein, M., R. A. Nicholas, et al. (2005). "Workstyle: Development of a measure of response to work in those with upper extremity pain." Journal of Occupational Rehabilitation **15**(2): 87-104.
Feuerstein, M., W. S. Shaw, et al. (2004). "From confounders to suspected risk factors: psychosocial factors and work-related upper extremity disorders." J Electromyogr Kinesiol **14**(1): 171-178.
Flanders, W. D. and P. H. Rhodes (1987). "Large sample confidence intervals for regression standardized risks, risk ratios, and risk differences." J Chronic Dis **40**(7): 697-704.
Furlow, C. F., R. T. Fouladi, et al. (2007). "A Monte Carlo study of the impact of missing data and differential item functioning on theta estimates from two polytomous Rasch family models." J Appl Meas **8**(4): 388-403.
Gaston-Johansson, F. (1984). "Pain assessment: differences in quality and intensity of the words pain, ache and hurt." Pain **20**(1): 69-76.
Gaston-Johansson, F. and J. Allwood (1988). "Pain assessment: Model construction and analysis of words used to describe pain-like experiences." Semiotica **71**(1-2): 73-92.
Greenland, S. (1987). "Interpretation and choice of effect measures in epidemiologic analyses." Am J Epidemiol **125**(5): 761-768.
Greenland, S. (2004). "Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies." Am J Epidemiol **160**(4): 301-305.
Griffiths, K. L., M. G. Mackey, et al. (2007). "The impact of a computerized work environment on professional occupational groups and behavioural and physiological risk factors for musculoskeletal symptoms: a literature review." J Occup Rehabil **17**(4): 743-765.
Guzman, J., E. L. Hurwitz, et al. (2008). "A new conceptual model of neck pain: linking onset, course, and care: the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders." Spine **33**(4 Suppl): S14-23.
Hadler, N. M. and T. S. Carey (1998). "Low back pain: an intermittent and remittent predicament of life." Ann Rheum Dis **57**(1): 1-2.
Haldeman, S., L. Carroll, et al. (2010). "Findings from the bone and joint decade 2000 to 2010 task force on neck pain and its associated disorders." J Occup Environ Med **52**(4): 424-427.
Haldeman, S., L. Carroll, et al. (2008). The Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: Executive Summary. **33:** S5-S7 10.1097/BRS.1090b1013e3181643f3181640.
Hamberg, K., G. Risberg, et al. (2002). "Gender bias in physicians' management of neck pain: a study of the answers in a Swedish national examination." J Womens Health Gend Based Med **11**(7): 653-666.
Hansson, T. and I. Jensen (2004). "Swedish Council on Technology Assessment in Health Care (SBU). Chapter 6. Sickness absence due to back and neck disorders." Scand J Public Health Suppl **63**: 109-151.

Harcombe, H., D. McBride, et al. (2010). "Physical and psychosocial risk factors for musculoskeletal disorders in New Zealand nurses, postal workers and office workers." Inj Prev **16**(2): 96-100.
Hausser, J. A., A. Mojzisch, et al. (2010). "Ten years on: A review of recent research on the Job Demand-Control (-Support) model and psychological well-being." Work & Stress **24**(1): 1-35.
Herloff, B., G. Ahlborg, et al. (2003). Unga IT-användare. Delrapport 1, Metodbeskrivning för webb-baserad enkätundersökning. Göteborg, Yrkes- och miljömedicin.
Hogg-Johnson, S., G. van der Velde, et al. (2008). "The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders." Spine **33**(4 Suppl): S39-51.
Hosmer, D. W., T. Hosmer, et al. (1997). A COMPARISON OF GOODNESS-OF-FIT TESTS FOR THE LOGISTIC REGRESSION MODEL. **16:** 965-980.
Hosmer, D. W. and S. Lemeshow (2000). Applied logistic regression, John Wiley & Sons.Inc.
Huang, G. D., M. Feuerstein, et al. (2002). "Occupational stress and work-related upper extremity disorders: Concepts and models." **41**(5): 298-314.
International Association for the Study of Pain. (2010). "Musculoskeletal Pain." Retrieved 2010-09-30, 2010, from http://www.iasp-pain.org/AM/AMTemplate.cfm?Section=Home&CONTENTID=9287&TEMPLATE=/CM/ContentDisplay.cfm.
Jensen, C., V. Borg, et al. (1998). "Job demands, muscle activity and musculoskeletal symptoms in relation to work with the computer mouse." Scand J Work Environ Health **24**(5): 418-424.
Johansson, E. E., K. Hamberg, et al. (1999). "The meanings of pain: an exploration of women's descriptions of symptoms." Soc Sci Med **48**(12): 1791-1802.
Johansson, H. and P. Sojka (1991). "Pathophysiological mechanisms involved in genesis and spread of muscular tension in occupational muscle pain and in chronic musculoskeletal pain syndromes: a hypothesis." Med Hypotheses **35**(3): 196-203.
Johnston, V., G. Jull, et al. (2010). "Interactive effects from self-reported physical and psychosocial factors in the workplace on neck pain and disability in female office workers." Ergonomics **53**(4): 502-513.
Kamaleri, Y., B. Natvig, et al. (2009). "Does the number of musculoskeletal pain sites predict work disability? A 14-year prospective study." Eur J Pain **13**(4): 426-430.
Karasek, R. and T. Theorell (1990). Healthy work: stress, productivity and the reconstruction of working life., New York (NY): Basic Books.
Karlqvist, L. K., M. Hagberg, et al. (1996). "Musculoskeletal Symptoms among Computer-assisted Design (CAD) Operators and Evaluation of a Self-assessment Questionnaire." Int J Occup Environ Health **2**(3): 185-194.
Keogh, E. and C. Eccleston (2006). "Sex differences in adolescent chronic pain and pain-related coping." Pain **123**(3): 275-284.
Kjellberg, A. and S. Iwanowski (1989). The stress/energy questionnaire - development of a method to evaluate estimations of mood at work., National Institute of Occupational Health**:** Report.
Kjellberg, A. and C. Wadman (2007). "The role of the affective stress response as a mediator of the effect of psychosocial risk factors on musculoskeletal complaints - Part 1: Assembly workers." International Journal of Industrial Ergonomics **37**(4): 367-374.
Knardahl, S. (2002). Psychophysiological mechanisms of pain in computer work: the blood vessel-nociceptor interaction hypothesis, Routledge. **16:** 179 - 189.
Knardahl, S. (2005). "Psychological and social factors at work: contribution to musculoskeletal disorders and disabilities." G Ital Med Lav Ergon **27**(1): 65-73.
Kuorinka, I., B. Jonsson, et al. (1987). "Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms." Appl Ergon **18**(3): 233-237.

Larsman, P., L. Sandsjo, et al. (2006). "Perceived work demands, felt stress, and musculoskeletal neck/shoulder symptoms among elderly female computer users. The NEW study." Eur J Appl Physiol **96**(2): 127-135.
Larsson, B., K. Sogaard, et al. (2007). "Work related neck-shoulder pain: a review on magnitude, risk factors, biochemical characteristics, clinical picture and preventive interventions." Best Pract Res Clin Rheumatol **21**(3): 447-463.
Lee, J. and K. S. Chia (1994). "Use of the prevalence ratio v the prevalence odds ratio as a measure of risk in cross sectional studies." Occup Environ Med **51**(12): 841.
Lee, J., C. S. Tan, et al. (2009). "A Practical Guide for Multivariate Analysis of Dichotomous Outcomes." Annals Academy of Medicine Singapore **38**(8): 714-719.
Lee, K. C., T. T. Chiu, et al. (2006). "Psychometric properties of the Fear-Avoidance Beliefs Questionnaire in patients with neck pain." Clin Rehabil **20**(10): 909-920.
Lee, Y. and J. A. Neider (2004). "Conditional and Marginal Models: Another View." Statistical Science **19**(2): 219-228.
Lidbeck, J. (2002). "Central hyperexcitability in chronic musculoskeletal pain: a conceptual breakthrough with multiple clinical implications." Pain Res Manag **7**(2): 81-92.
Localio, A. R., D. J. Margolis, et al. (2007). "Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression." J Clin Epidemiol **60**(9): 874-882.
Lundberg, U. (2002). "Psychophysiology of work: stress, gender, endocrine response, and work-related upper extremity disorders." Am J Ind Med **41**(5): 383-392.
Marcus, M., F. Gerr, et al. (2002). "A prospective study of computer users: II. Postural risk factors for musculoskeletal symptoms and disorders." Am J Ind Med **41**(4): 236-249.
McEwen, B. S. (2008). "Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators." Eur J Pharmacol **583**(2-3): 174-185.
McNeil, D. W. and A. J. Rainwater, 3rd (1998). "Development of the Fear of Pain Questionnaire--III." J Behav Med **21**(4): 389-410.
McNutt, L. A., C. Wu, et al. (2003). "Estimating the relative risk in cohort studies and clinical trials of common outcomes." Am J Epidemiol **157**(10): 940-943.
Melin, B. and U. Lundberg (1997). "Abiopsychosocial approach to work-stress and musculoskeletal disorder." J Psychophys **11**: 238-247.
Melzack, R. (1999). "From the gate to the neuromatrix." Pain **Suppl 6**: S121-126.
Mense, S. (1992). Peripheral Mechaniism of Muscle Nociception and Local Muscle Pain, Informa Healthcare. **1:** 133 - 170.
Messing, K. and J. Mager Stellman (2006). "Sex, gender and women's occupational health: the importance of considering mechanism." Environ Res **101**(2): 149-162.
Messing, K. and B. A. Silverstein (2009). "Gender and occupational health." Scand J Work Environ Health **35**(2): 81-83.
Messing, K., F. Tissot, et al. (2009). "Should studies of risk factors for musculoskeletal disorders be stratified by gender? Lessons from the 1998 Quebec Health and Social Survey." Scand J Work Environ Health **35**(2): 96-112.
Mickey, R. M. and S. Greenland (1989). "The impact of confounder selection criteria on effect estimation." Am J Epidemiol **129**(1): 125-137.
Miettinen, O. S. and E. F. Cook (1981). Confounding: Essence and detection. **114:** 593-603.
Molenberghs, G. and G. Verbeke (2004). "Meaningful statistical model formulations for repeated measures." Statistica Sinica **14**: 989-1020
Murphy, S., P. Buckle, et al. (2007). "A cross-sectional study of self-reported back and neck pain among English schoolchildren and associated physical and psychological risk factors." Appl Ergon **38**(6): 797-804.
Nemes, S., J. M. Jonasson, et al. (2009). "Bias in odds ratios by logistic regression modelling and sample size." BMC Med Res Methodol **9**: 56.

Noack-Cooper, K. L., C. M. Sommerich, et al. (2009). "College students and computers: assessment of usage patterns and musculoskeletal discomfort." Work **32**(3): 285-298.
Nordander, C., K. Ohlsson, et al. (2008). "Gender differences in workers with identical repetitive industrial tasks: exposure and musculoskeletal disorders." Int Arch Occup Environ Health **81**(8): 939-947.
Nurminen, M. (1995). "To use or not to use the odds ratio in epidemiologic analyses?" Eur J Epidemiol **11**(4): 365-371.
Ohlsson, K., R. G. Attewell, et al. (1994). "An assessment of neck and upper extremity disorders by questionnaire and clinical examination." Ergonomics **37**(5): 891-897.
Orhede, E. (1994). "Nordic cooperation in research on the work environment." Scand J Work Environ Health **20**(1): 65-66.
Palmer, K., G. Smith, et al. (1999). "Repeatability and validity of an upper limb and neck discomfort questionnaire: the utility of the standardized Nordic questionnaire." Occup Med (Lond) **49**(3): 171-175.
Peduzzi, P., J. Concato, et al. (1996). "A simulation study of the number of events per variable in logistic regression analysis." J Clin Epidemiol **49**(12): 1373-1379.
Rau, R., K. Morling, et al. (2010). "Is there a relationship between major depression and both objectively assessed and perceived demands and control?" Work & Stress **24**(1): 88-106.
Roelofs, J., L. Goubert, et al. (2004). "The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia." European Journal of Pain **8**(5): 495-502.
Rothman, K. (2002). Epidemiology: an introduction. New York, Oxford University Press.
Rothman, K. and S. Greenland (1998). Modern Epidemiology. New York, Lippincott-Raven Publishers.
Rothman, K. J., S. Greenland, et al. (2008). Modern Epidemiology. New York, Lippincott Williams & Wilkins.
Rotnitzky, A. and N. P. Jewell (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. **77:** 485-497.
Rugulies, R., B. Aust, et al. (2004). The epidemiology of health and illness.A socio-psycho-physiological perspective. . The Sage handbook of health psychology. Sutton S, Baum A and J. M. London, United Kingdom, Sage Publishers**:** 27–68.
Sauter, S. and N. G. Swanson (1996). An ecological model of musculoskeletal disorders in office work. Beyond Biomechanics. Psychosocial Aspects of Musculoskeletal Disorders in Office Work. S. D. M. a. S.L.Sauter. London, Taylor & Francis.
Silverstein, B., Z. J. Fan, et al. (2009). "Gender adjustment or stratification in discerning upper extremity musculoskeletal disorder risk?" Scand J Work Environ Health **35**(2): 113-126.
Skov, T., J. Deddens, et al. (1998). "Prevalence proportion ratios: estimation and hypothesis testing." Int J Epidemiol **27**(1): 91-95.
Sluka, K. A. and L. A. Rasmussen (2010). "Fatiguing exercise enhances hyperalgesia to muscle inflammation." Pain **148**(2): 188-197.
Smith, M. J. and P. C. Sainfort (1989). "A balance theory of job design for stress reduction." International Journal of Industrial Ergonomics **4**(1): 67-79.
Spiegelman, D. and E. Hertzmark (2005). "Easy SAS calculations for risk or prevalence ratios and differences." Am J Epidemiol **162**(3): 199-200.
Strazdins, L. and G. Bammer (2004). "Women, work and musculoskeletal health." Soc Sci Med **58**(6): 997-1005.
Strom, V., C. Roe, et al. (2009). "Work-induced pain, trapezius blood flux, and muscle activity in workers with chronic shoulder and neck pain." Pain **144**(1-2): 147-155.

Stromberg, U. (1994). "Prevalence odds ratio v prevalence ratio." Occup Environ Med **51**(2): 143-144.
Stromberg, U. (1995). "Prevalence odds ratio v prevalence ratio--some further comments." Occup Environ Med **52**(2): 143.
Swedish Work Environment Authority (1995). Arbetsmiljön 1995. Arbetsmiljöstatistik Rapport. Stockholm.
Swedish Work Environment Authority (2008). Arbetsmiljön 2007. Arbetsmiljöstatistik Rapport. Stockholm. **2008:4**.
Tait, R. C., J. T. Chibnall, et al. (1990). "The Pain Disability Index: psychometric properties." Pain **40**(2): 171-182.
Takala, E. P., E. Viikari-Juntura, et al. (1992). "Seasonal variation in neck and shoulder symptoms." Scand J Work Environ Health **18**(4): 257-261.
Thompson, M. L., J. E. Myers, et al. (1998). "Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done?" Occup Environ Med **55**(4): 272-277.
Turk, D. C., R. H. Dworkin, et al. (2008). "Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain." Pain **137**(2): 276-285.
Ursin, H. and H. R. Eriksen (2004). "The cognitive activation theory of stress." Psychoneuroendocrinology **29**(5): 567-592.
Waddell, G. (2006). "Preventing incapacity in people with musculoskeletal disorders." Br Med Bull **77-78**: 55-69.
Wadman C and Kjellberg A (2007). "The role of the affective stress response as a mediator for the effect of psychosocial risk factors on musculoskeletal complaints - Part 2: Hospital workers." International Journal of Industrial Ergonomics **37**(5): 395-403.
Wahlstrom, J. (2005). "Ergonomics, musculoskeletal disorders and computer work." Occup Med (Lond) **55**(3): 168-176.
Wahlstrom, J., J. Svensson, et al. (2000). "Differences between work methods and gender in computer mouse use." Scand J Work Environ Health **26**(5): 390-397.
Walker-Bone, K., K. T. Palmer, et al. (2004). "Prevalence and impact of musculoskeletal disorders of the upper limb in the general population." Arthritis Rheum **51**(4): 642-651.
Westgaard, R. H. and C. J. De Luca (2001). "Motor control of low-threshold motor units in the human trapezius muscle." J Neurophysiol **85**(4): 1777-1781.
Wiesenfeld-Hallin, Z. (2005). "Sex differences in pain perception." Gender Medicine **2**(3): 137-145.
Wiesenfeld-Hallin Z. (2005). "Sex differences in pain perception." Gender Medicine **2**(3): 137-145.
Viikari-Juntura, E., R. Martikainen, et al. (2001). "Longitudinal study on work related and individual risk factors affecting radiating neck pain." Occup Environ Med **58**(5): 345-352.
Visser, B. and J. H. van Dieen (2006). "Pathophysiology of upper extremity muscle disorders." J Electromyogr Kinesiol **16**(1): 1-16.
Von Korff, M., J. Ormel, et al. (1992). "Grading the severity of chronic pain." Pain **50**(2): 133-149.
Woolf, A. D. and B. Pfleger (2003). "Burden of major musculoskeletal conditions." Bull World Health Organ **81**(9): 646-656.
Woolf, C. J. (1987). "Physiological, inflammatory and neuropathic pain." Adv Tech Stand Neurosurg **15**: 39-62.
Woolf, C. J. and T. P. Doubell (1994). "The pathophysiology of chronic pain--increased sensitivity to low threshold A beta-fibre inputs." Curr Opin Neurobiol **4**(4): 525-534.
Woolf, C. J. and M. W. Salter (2000). "Neuronal plasticity: increasing the gain in pain." Science **288**(5472): 1765-1769.

Yang, X., J. Li, et al. (2008). "Imputation-based strategies for clinical trial longitudinal data with nonignorable missing values." Stat Med **27**(15): 2826-2849.
Yu, B. and Z. Wang (2008). "Estimating relative risks for common outcome using PROC NLP." Comput Methods Programs Biomed **90**(2): 179-186.
Yu, F., H. Morgenstern, et al. (2003). "Use of a Markov transition model to analyse longitudinal low-back pain data." Stat Methods Med Res **12**(4): 321-331.
Zocchetti, C., D. Consonni, et al. (1997). "Relationship between prevalence rate ratios and odds ratios in cross-sectional studies." International Journal of Epidemiology **26**(1): 220-223.