On the origin of back pain

Pieter Coenen

The study presented in this thesis was performed at MOVE Research Institute Amsterdam, Faculty of Human Movement Sciences, VU University Amsterdam, Amsterdam, The Netherlands.

VU University Amsterdam participates in Body@Work, Research Center on Physical Activity, Work and Health, which is a joint initiative of VU University Medical Center (Department of Public and Occupational Health, EMGO Institute for Health and Care Research), VU University Amsterdam, and TNO.

Financial support for the printing of this thesis has kindly been provided by:

Body@Work, Research Center on Physical Activity, Work and Health, the
Netherlands Organisation for Applied Scientific Research TNO, Work Solutions,
the EMGO+ Institute for Health and Care Research, and the VU University.

and

Ms Coenen-v Kimmenade

English title: On the origin of back pain
Nederlandse titel: Over het ontstaan van rugklachten

ISBN: 978-94-6191-910-6

Cover design, layout and printing by: Ipskamp Drukkers, Enschede

VRIJE UNIVERSITEIT

On the origin of back pain

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan

de Vrije Universiteit Amsterdam,

op gezag van de rector magnificus

prof.dr. F.A. van der Duyn Schouten,

in het openbaar te verdedigen

ten overstaan van de promotiecommissie

van de Faculteit der Bewegingswetenschappen

op vrijdag 29 november 2013 om 9.45 uur

in de aula van de universiteit,

De Boelelaan 1105

door

Pieter Coenen

geboren te Tilburg

promotoren:	prof.dr. J.H. van Dieën prof.dr.ir. P.M. Bongers
copromotoren:	dr. I. Kingma dr. C.R.L. Boot

...als je goed kijkt, zie je steeds weer

nieuwe oplossingen...

(Jeannet van Ganzewinkel, 2010)

Table Of Contents

Chapter 1

General Introduction

LOW-BACK PAIN IN SOCIETY

In our society, low-back pain (LBP) is one of the most common health problems, thereby causing a large burden, medically as well as economically (Goetzel et al., 2003; Maetzel & Li, 2002). As a specific pathological diagnosis is not made in many cases of LBP (Koes et al., 2006), LBP is often labeled as non-specific. In this thesis, when mentioning LBP, we refer to self-reported LBP, without focusing on diagnoses.

While data differ considerably between studies executed in different countries, lifetime LBP prevalence worldwide is estimated to be approximately 39% while the point prevalence is estimated to be around 19% (Hoy et al., 2012). In the Netherlands, the point prevalence of LBP was estimated to be 26%, depicting that more than a quarter of the Dutch population experiences LBP at any moment in time (Picavet & Schouten, 2003). Besides, LBP has been shown to be recurrent in a majority of patients (Andersson, 1999; Picavet & Schouten, 2003) and it can potentially lead to chronic pain (Kovacs et al., 2005). Because of this high prevalence and its potential to develop into chronicity, LBP can strongly interfere with people's lives as well as with their participation in society. Furthermore, in the working population, LBP has shown to lead to work disability (Eriksen et al., 2004; Matsudaira et al., 2012; Welch et al., 2009), sick leave (Geuskens et al., 2008; van den Heuvel et al., 2004), and early retirement (Costa-Black et al., 2010; Faber et al., 2010; Picavet & Schouten, 2003), indicating a large impact on the working population as well. All the above-mentioned consequences have economic effects that have been highlighted in a recent study estimating the total costs of LBP for Dutch society to be €4.3 Billion in 2007 (which was at that time 0.6% of the gross national product) as a consequence of, among other variables, health costs, production loss, and disability costs (Lambeek et al., 2011). Therefore, it can be concluded that LBP is a major issue in (working) society. In order to better understand the problem of LBP, more knowledge on the causal mechanisms of LBP is needed. This thesis describes a combined epidemiological and biomechanical approach to enhance our understanding of LBP etiology.

WORK-RELATED RISK FACTORS OF LBP

In the past years, epidemiological studies have contributed to our understanding of the etiology of LBP. In certain sectors of industry and in some occupations, the prevalence of LBP is considerably higher than in the general working population (Punnett & Wegman, 2004), indicating some work-relatedness of the etiology of LBP (Lötters et al., 2003). This work-relatedness has become more clear as, besides personal risk factors (e.g. age, smoking habits, physical capacity and body weight; Hamberg-van Reenen et al., 2007; Hooftman et al., 2004; Leboeuf-Yde, 2004; Wai et al., 2008) and (work related) psychosocial risk factors (e.g. stress, social support and job satisfaction, role conflict and job control; Eatough et al., 2012; Hartvigsen et al., 2004; Linton, 2001), the occurrence of LBP has been associated with physical work-related risk factors. Of these physical risk factors, lifting, carrying, pushing, pulling, awkward trunk postures (e.g., flexion and rotation) and whole body vibrations are most frequently reported to be associated with LBP (Chen et al., 2009; da Costa & Vieira, 2010; Griffith et al., 2012; Lis et al., 2007; Lötters et al., 2003). Despite this, other studies have argued that evidence concerning physical risk factors of LBP is weak, possibly as a result of insufficient quality of studies performed thus far (Bakker et al., 2009; Kwon et al., 2011) due to the absence of adequately quantified physical work load in prospective studies. This inconsistency and lack of knowledge has negatively affected the prevention of LBP and has hampered abilities to recommend acceptable levels of biomechanical loads at work (Fallentin et al., 2001). Furthermore, although work-related interventions in attempts to reduce LBP occurrence have frequently been applied (Westgaard & Winkel, 1997), in general, these interventions have not proven to be successful on a large scale (Dempsey, 2007; Verbeek et al., 2011). In part, this may be due to absence or inadequacy of measurements of physical loading.

Failure mechanisms

Despite our lack of knowledge on LBP etiology, several models have been developed to describe the causal chain of the occurrence of LBP (e.g.; Chaffin, 2009; van der Beek & Frings-Dresen, 1998; van Dieën et al., 1999; Wells et al., 2004). All of these models assume mechanical load in the lower back as a result of exposure to physical load at the workplace (i.e., due to the above mentioned risk factors, such as lifting and trunk flexion) to be an important variable in this chain (Figure 1.1). Such mechanical loads (i.e., low-back moments as indicators for mechanical load, or compression and shear forces on the lumbar spine) are in most of these models at, or close to, the end of the causal chain, thereby providing a more direct relationship with spinal failure and consequently with LBP than exposure variables. These mechanical load metrics can therefore provide important insights into the etiology of LBP (Wells et al., 2004). The advantage of the use of mechanical load metrics as opposed to more traditional exposure measures is that different exposures (e.g., lifting, twisting and bending) that can be expressed in three dimensions (i.e., duration, frequency and intensity) affect the same mechanical load (Burdorf, 2010). Besides, the magnitude of exposure variables (i.e., number of lifts or time working in an awkward posture) is not directly related to the magnitude of mechanical load variables. As an example, when lifting a 6kg box, compression forces can be up to 5000N during lifting objects from ground level, but these forces are approximately half this magnitude when the box is lifted from shoulder level (Faber et al., 2009). Moreover, even with no or small loads on the hands, mechanical low-back loading can be substantial, as a result of gravitational forces acting on the upper body and upper extremities as well as due to acceleration of these body segments (van Dieën et al., 2010). Therefore, several exposure variables that can be expressed in terms of frequency, duration and intensity of a lifting task all affect the magnitude of mechanical load on the lower back in a different way (Figure 1.1; Davis & Marras, 2000; Faber et al., 2007; Ferguson et al., 2002; Hoozemans et al., 2008; Marras et al., 1999).

Besides the above, mechanical loads can also take other mediating factors into account. These factors, such as psychosocial factors, personal factors and work-related factors can interact with the abovementioned causal chain in multiple ways (Chaffin, 2009; Wells et al., 2004). As an example, under psychosocial load, workers are more likely to experience more physical strain during work, for instance due to a change in work velocity or work strategy, and this may increase the risk of LBP (Eatough et al., 2012). With regard to personal factors, it has been shown that men may have a higher back load due to a higher torso mass (Hooftman et al., 2004), but also a higher load tolerance than women (Waters et al., 1993). As a final example, the type of job and company are associated with variables like deadlines and workplace culture (Moray, 2000), influencing the way a workers interacts with the environment, which potentially affects the physical load on the worker. From the above, it can be concluded that when measuring mechanical loads rather than crude exposure estimates (i.e., number of lifts, time in a trunk flexed posture), the causality with LBP can be assessed with more accuracy because mediating factors can be taken into account.

Empirical evidence has shown that mechanical load metrics are stronger associated with LBP than exposure estimates (Norman et al., 1998). Therefore, using mechanical load metrics in field settings seems to be important when striving to enhance our understanding on the etiology of LBP (Burdorf, 2010; Wells et al., 2004). However, measurement methods are prone to a trade-off between accuracy and feasibility, in terms of investments in time and costs (Winkel & Mathiassen, 1994). Therefore, in general, with a limited research budget, relatively simple (subjective) observations or self-reports are applied on a larger group of subjects, whereas a more thorough assessment of the work-load often implies that fewer workers can be measured. Moreover, these thorough assessment tools often consist of laboratory-based measurements that are difficult to apply in a field-based setting. This trade-off and the currently available measurement tools will be discussed later in this introduction.

In the studies described in this thesis we assess low back moments at the level of the L5-S1 joint only. Furthermore, we did not separate these moments into shear forces or compression forces. However, we assume that there is a strong correlation of loads among the different levels of the low-back. Furthermore, a strong correlation of low-back moments with shear forces and compression forces has been shown before (van Dieën & Kingma, 2005). Therefore, it is assumed that moments at the level of L5-S1 provide a representative measure of low-back loads in general.

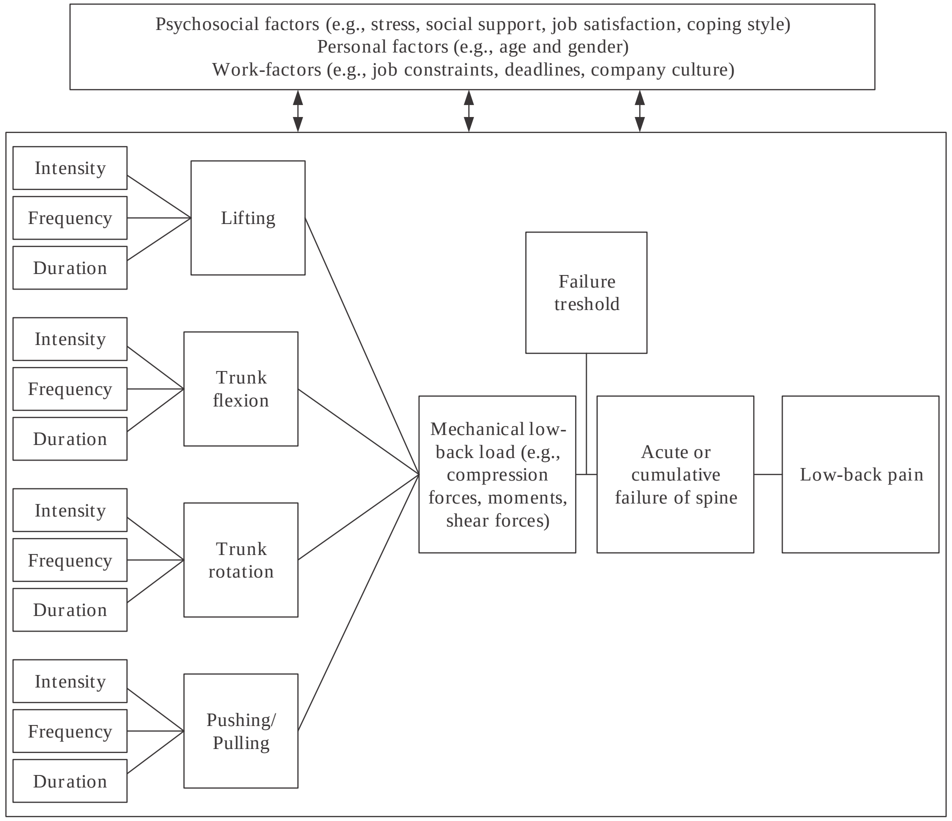

Figure 1.1 | Model representing the association of physical work load and LBP, inspired by other models (Chaffin, 2009; van der Beek & Frings-Dresen, 1998; Wells et al., 2004). Different exposures (i.e., lifting, trunk flexion, trunk rotation and pushing/pulling) that can be expressed in a duration, frequency and intensity are taken into account when measuring a mechanical load (e.g., compression forces on the spine or low-back moments). Subsequently, these mechanical loads can, depending on the failure threshold of the spine, cause failure (either acute or due to cumulative loads) which potentially leads to LBP. These loads furthermore take into account mediating variables of which the effects, for the sake of readability of the figure, are represented in a simplified way. It should be noted that the exposures shown are not independent, as, for example a lifting task usually involves trunk flexion.

The second part of the causal chain for LBP etiology, in which the (mechanical) load eventually leads to the occurrence of LBP, has been discussed in the literature as well (e.g., Adams, 2004; Chaffin, 2009; Marras, 2012). Although a specific cause of LBP is established in only 10% of all LBP cases (Koes et al., 2006), damage to structures of the vertebral column as a result of mechanical loading is a likely cause of LBP (van Dieën et al., 1999; Wang et al., 2012b). In cadaver experiments, damage to several structures of the spinal motion segments (facet joints and inter-vertebral discs, but in most of the cases vertebral endplates) has been shown under several protocols of realistic mechanical loading of spinal motion segments (Adams et al., 1994; Brinckmann et al., 1988; Callaghan & McGill, 2001; Gallagher et al., 2007; Gunning et al., 2001; Howarth & Callaghan, 2012; van Dieën et al., 2006). Moreover, in a retrospective cadaveric study, signs of endplate and disc damage were strongly related to a history of LBP (Wang et al., 2012a, b). From these data, in general, two mechanisms for the occurrence of failure can be derived. The first mechanism assumes damage of spinal structures due to acute high loads, causing instantaneous failure of tissue (Figure 1.2, upper panel). This mechanism is supported by studies reporting instantaneously high loads causing damage to spinal structures (Howarth & Callaghan, 2012), in most cases failure of the spinal endplate (Adams et al., 1994; Brinckmann et al., 1988). However, not just a single supra-maximal compression but also repeated sub-maximal compression can lead to injury. This repeated sub-maximal compression causes similar damage at lower force levels (Brinckmann et al., 1988; Hansson et al., 1987). Therefore, the second mechanism supposes an accumulation of micro-damage, decreasing the tolerance of tissue and eventually leading to failure after sustained or repeated loading (Figure 1.2, lower panel).

The above mentioned in-vitro studies, showing that peak and cumulative loads may cause spinal failure, militate in favor of both the peak as well as the cumulative etiological mechanism. However, this information is based on in-vitro studies, which bring along some limitations. For example, it is known that there is no one-to-one relationship between mechanical damage to the spine and the actual occurrence of LBP (Wang et al., 2012a, b). Besides, cadaver material does recover poorly from loads as biological repair is absent (van der Veen et al., 2005). The abovementioned studies on cumulative loading should therefore be interpreted with caution as damage in these studies might have occurred earlier than during *in-vivo* conditions. On the other hand, the opposite, underestimation of cumulative load effects in *in-vitro* studies, also cannot be excluded. Specifically, alternative explanations for cumulative load effects are not taken into account in *in-vitro* studies. Such alternative explanations for the cumulative etiological mechanism are impaired coordination due to neuromuscular or cardiovascular fatigue after cumulative loading. It has been suggested that this impaired coordination might cause a reduction of the tolerance of the spine due to lack of stability (Granata & Gottipati, 2008; Johanson et al., 2011; Sparto et al., 1997) or alterations in work postures posing higher loads on the spine (Bonato et al., 2003; Dolan & Adams, 1998).

From the above, it can be concluded that epidemiological studies in which peak and cumulative mechanical load and LBP are assessed *in-vivo* in work settings should be considered in order to obtain more information on the etiology of LBP. Marras and colleagues investigated the predictive value of a variety of low-back load parameters for the risk of LBP (2010; 1995). Other studies suggest that cumulative loads acting on the spine may contribute to LBP (Kerr et al., 2001; Kumar, 1990; Neumann et al., 2001a; Norman et al., 1998) as well as to specific lower back pathologies (i.e., lumbar disc disease; Seidler et al., 2009; Seidler et al., 2003). Other studies showed evidence for the association of peak loads and LBP (Kerr et al., 2001; Neumann et al., 2001a; Norman et al., 1998; Punnett et al., 1991). However, the above-mentioned studies describe either cross-sectional studies or prospective studies with low-back loads that are based on crude estimates. Risk associations that are based on prospective studies are more valid for obtaining insight into etiological causalities as the occurrence of LBP follows the exposure to a certain risk factor. Therefore, these designs are more preferable in epidemiological studies (Rothman & Greenland, 2005). However, to the best of our knowledge, information on mechanical loads on the lower back and the occurrences of LBP from such prospective studies is not available. Two important reasons for this void are the lack of field-based measurement techniques to determine mechanical loads on the low-back, and the lack of knowledge on the variability of physical load and thus on the type of measurement allocation (e.g., measuring multiple workers, a few times or a few workers multiple times) needed. Caveats and potential possibilities in measurement strategies of epidemiological studies considering these two factors will therefore be discussed in the following paragraphs.

ASSESSMENT OF PHYSICAL WORK LOAD

An important reason for the inconsistency of information on LBP etiology is that risk associations are highly influenced by the choice of a measurement method (Burdorf, 2010; David, 2005). Measurement methods are prone to a trade-off between accuracy and feasibility (in terms of investments in time and costs) and the available resources determine the precision of a measurement and hence, statistical efficiency (Mathiassen & Bolin, 2011). In order to enhance our knowledge on the etiology of LBP it seems to be relevant, as stated in the previous paragraphs, to assess mechanical load metrics rather than exposure estimates. However, as accurate measurement tools to assess these mechanical loads are often difficult to apply to field situations, concessions with respect to the quality of measurement techniques are often made. As a result, many studies measure exposure variables rather than mechanical load.

Work-related risk factors can be assessed by either self-reports, observations (i.e. subjective risk estimations or structured observations) or direct measurements (e.g. muscle activity measurements, goniometry and measurement of external forces). These methods have been discussed in the literature and advantages and disadvantages have been evaluated (David, 2005; van der Beek & Frings-Dresen, 1998). Self-reports have been used in numerous epidemiological studies and are easily applicable; however, their accuracy has been questioned (Balogh et al., 2004; Punnett, 2004). Therefore, in contrast to epidemiological studies, self-reports of workload are rarely used for evaluation in ergonomic practice (Hansson et al., 2001).

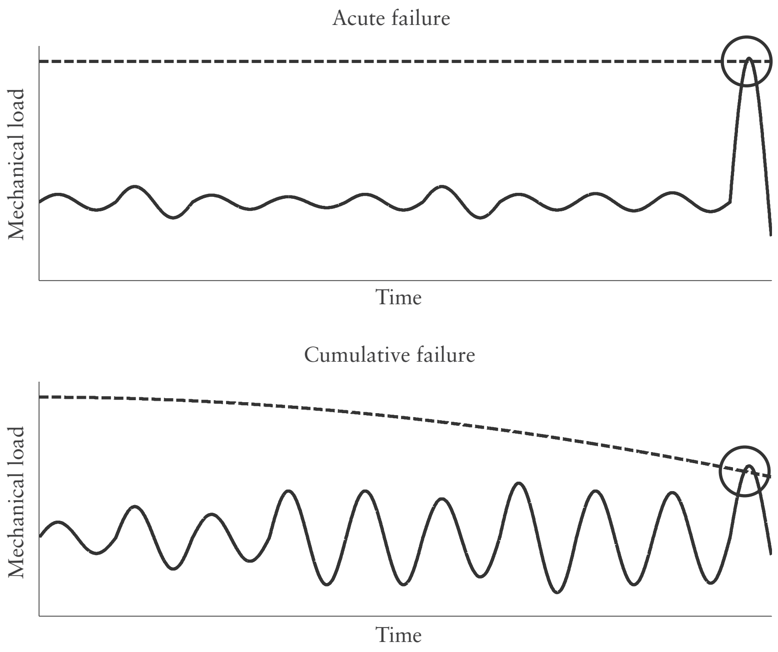

Figure 1.2 | Illustration of spinal motion segment failure due to either a single acute loading event (upper panel) or cyclic loading leading to cumulative fatigue failure (lower panel). In both figures, the mechanical loading pattern is represented with a solid line whereas the failure threshold of the spine is represented with a dashed line. In the event of a large acute loading, a single load can reach the failure threshold. During a cyclic loading pattern the repeated load lowers the threshold, eventually leading to failure at a smaller load level. The figure is inspired by earlier work (Chaffin, 2009; Marras, 2012; McGill, 2009).

Instead, risk estimates by observers are more frequent used. Although these observations have a higher accuracy and validity than self-reports, their accuracy and validity is assumed to be lower than that of direct measurement tools (Spielholz et al., 2001; Takala et al., 2010). Accuracy is limited because risk estimations are often based on crude categorization. Validity is limited because of the difficulty for the assessors to conduct such measurements objectively. Structured observations performed by observers might not be so vulnerable for subjectivity. However, these observations are just as the risk estimations, prone to limited accuracy as they are often based on crude categorization (de Looze et al., 1994a; van Wyk et al., 2009). This inaccuracy has been shown to lead to large errors when used as input in biomechanical models to estimate mechanical low-back load (de Looze et al., 1994b). Ultimately, observational methods often lack a clear quantification of physical load in the dimension of duration, frequency and magnitude (Takala et al., 2010).

The last group of measurement tools is the group of direct measurements. These measurements are assumed to be the most objective and thus the most valid and have the ability to provide mechanical load estimates. For instance, it has been shown that mechanical low-back load can be measured accurately by using inverse dynamic linked-segment models, which combine information from three-dimensional motion tracking procedures and external force measurements (e.g., Kingma et al., 1996; Kingma et al., 2010; Plamondon et al., 1996). However, such measurements are time and money consuming. In addition, they can hardly be used outside the laboratory setting as they strongly interfere with the work performed which highly complicates measurements of realistic occupational situations. Furthermore, it is known that when mock-ups of field situations are made in a laboratory setting, workers tend to execute tasks differently than they would have done in the actual field (Faber et al., 2011). Therefore, laboratory measurements tools are, despite their high accuracy and internal validity not always externally valid.

Accordingly, research has focused on less costly (with respect to time and money) low-back load assessment methods, which can be brought into the work place easily. For example, variables serving as a proxy for mechanical load are often adopted, such as muscle activity measures (Hägg et al., 2000), static position measures (i.e., load distances; Potvin, 1997; van Dieën et al., 2010) or measures obtained from instrumented motion monitors (e.g., Marras et al., 2007; Marras et al., 2010). Despite the fact that some of these estimates are closely correlated to mechanical low-back load (Neumann et al., 2001b), it is believed that mechanical low-back load estimates are needed to properly assess the load on the low-back (Wells et al., 2004). Video-based methods using postural exposure data in biomechanical models to calculate mechanical low-back loads have been shown to be a promising category of techniques (Chang et al., 2010; Norman et al., 1998; Potvin, 1997; Sutherland et al., 2008) to assess low-back load metrics such as static (Neumann et al., 2001b), cumulative (Sutherland et al., 2008) or peak low-back moments (Norman et al., 1998). These methods allow raters with minimal training and minor use of equipment to collect occupational low-back load data with high inter-rater agreement (Cann et al., 2008; Sullivan et al., 2002). However, these methods have rarely been implemented in field-based epidemiological studies. Therefore, improving such measurement tools (in terms of validity, reliability and feasibility) should be considered, which is another focus of this thesis.

VARIABILITY IN PHYSICAL WORK LOAD

Another important aspect to be considered when constructing a measurement strategy for physical risk factors of LBP is the variability (between and within workers) of physical risk factors (either expressed in exposure metrics or in mechanical loads). This variability should be considered in the planning, analysis and interpretation of epidemiologic studies as inadequate distribution of measurements can lead to biased regression results (Tielemans et al., 1998) and to a reduced statistical power (Mathiassen et al., 2002; Mathiassen et al., 2003). Therefore, measurement occasions should be distributed adequately over subjects, time and tasks groups (Loomis & Kromhout, 2004).

Statistical consequences of work load variability between individuals, and within and between days within individuals have been addressed in several studies for various load metrics and occupational settings (e.g., Hansson et al., 2006; Svendsen et al., 2005; Wahlstrom et al., 2010). The effect of sample size on the variance of the load estimate has been discussed, including the number of samples to arrive at a sufficiently reliable load estimate (Allread et al., 2000; Paquet et al., 2005; Svendsen et al., 2005). It has been shown that although the reliability of a measurement improves when more subjects are sampled or when load is measured over multiple occasions, with increasing sample size, the load estimate improves less when measuring more subjects (Mathiassen et al., 2002; Mathiassen et al., 2003). Other studies discuss several options for collecting data from workers (sample allocation). For example, it has been shown that it might be more beneficial to collect data over multiple days from multiple workers rather than to collect data from just a few workers on a single day (Liv et al., 2010; Svendsen et al., 2005). Also the effects of group-based measurement approaches, that are often adopted (e.g., Ariëns et al., 2001; Burdorf & Jansen, 2006), have been described frequently. In these approaches, workers are classified into groups; work load is measured only in a selection of workers within each group, and the (mean) group-based work load of the measured workers is assigned to all subjects in the group. Work load-outcome relationships are then determined using these load estimates together with individual outcome data (LBP) from all subjects. These group-based measurement approaches have proven to be successful for the assessment of workloads during several occupational tasks (Hoozemans et al., 2001; Paquet et al., 2005). Furthermore, stronger associations have been found in a group-based approach compared to an individual-based approach when it comes to associations of physical load to outcomes (Jansen & Burdorf, 2003).

From the above, it can be concluded that there is quite some knowledge available on how to deal with variability in physical work load. Also the effect of sampling strategies and study protocols (e.g., group-based measurement approaches) on the reliability of measured physical risk factors has been discussed thoroughly. However, information on the effect of this variability on statistical power of eventual risk associations is limited. This is therefore an additional focus of the present thesis.

AIMS AND OUTLINE OF THE THESIS

From the previous paragraphs it can be concluded that insufficient knowledge on the linkage of biomechanical loading and the etiology of LBP is available. More specifically, limited information on the effect of mechanical low-back load on LBP has been obtained from prospective epidemiological studies. This is partly because study properties that highly affect the risk associations (e.g., data sampling, the exact load metric used etc.) are insufficiently understood. Another reason is the limited availability of occupational assessment tools that are easily applicable in field based situations. From these hiatuses, four principle aims that will be addressed in this thesis are formulated.

In this thesis we aim to assess:

1. The predictive value for LBP of mechanical loads as compared to (subjective) exposure estimates
2. The effects of methodological issues on the predictive value of low-back load metrics for the occurrence of LBP
3. The applicability of video-based quantification of mechanical low-back load in a field situation

These three aims will be instrumental for our main aim, to gain insight into:

4. The etiology of LBP using mechanical load metrics

These principle aims will be addressed in the chapters of the thesis according to the following outline.

The predictive value for LBP of mechanical loads as compared to (subjective) exposure estimates

As mechanical low-back loads have been assumed to have a higher predictive value than exposures (obtained from self-reports or from observations) for LBP, our initial goal was to test this hypothesis in a prospective study. In Chapter 2, a study is described in which the predictive value of subjective observer assessments for the risk of musculoskeletal pain is evaluated. Results of this study can be used to assess the quality of these subjective metrics. In Chapter 3, a study is described in which, based on video observation, a first attempt was made to obtain a mechanical load metric in a prospective epidemiological study. In this study, mechanical loads were assessed with static calculations of mechanical back load based on crude posture observation categories. The association with LBP of the mechanical load metric studied in this chapter was assessed and was compared to associations of exposure estimates that are generally adopted as exposure risk factors. These studies provide information on the predictive value of mechanical low-back load metrics in comparison to exposure metrics (determined either subjectively or from observations).

The effects of methodological issues on the predictive value of low-back loads for LBP
Previously, cumulative load has often been suggested to be a potential risk factor of LBP. As peak loads are also assumed to be independent risk factors of LBP, the question arises how repeated peaks should be weighted in cumulative load calculations. Chapter 4 describes an analysis based on in-vitro data. In this analysis, the contribution of repetition of peaks in the calculation of cumulative loads was assessed. Data from this study may provide important information for future studies assessing cumulative low-back load as a risk factor of LBP.

The statistical power of studies assessing risk factors of LBP is highly influenced by the measurement strategies used. However, in physical load-outcome associations this influence is poorly understood. Therefore, a simulation study assessing the effect of several measurement strategies on the predictive value of such risk associations has been performed (Chapter 5). Data from this study can provide useful information for the design of future epidemiological studies.

The applicability of video based quantification of mechanical low-back load in a field situation
As described above, measuring mechanical low-back loads in work field settings is a daunting task as current measurement methods often interfere with the work or provide only crude estimates. Therefore, a video analysis method for the assessment of low-back loads in the field was developed. As opposed to the earlier used method of mechanical load calculation based on crude posture observations (Chapter 3), this method consists of a detailed kinematic analysis of manual material handling tasks. This analysis method can potentially be used in ergonomic practice and future epidemiological studies as it copes with abovementioned drawbacks. In order to test the quality of this method, at first, the validity of the method was tested by comparing it to a gold-standard laboratory method. The proposed video analysis method is described in detail in Chapter 6, in which also the outcomes of this validation-test are provided. Also the inter-rater reliability of the video analysis method applied to actual field situations is assessed (Chapter 7). Results of both studies provide information on the applicability of the described method in future research and in ergonomic practice.

The etiology of LBP using mechanical load metrics
The earlier mentioned video analysis method was applied in a large prospective cohort study. Results from this study are described in Chapter 8, providing insight in LBP etiology that can be useful for future prevention of LBP.

In the epilogue (Chapter 9), an overview of the studies described above will be provided. At the end of this chapter, final conclusions of this thesis will be drawn. Furthermore, implications for ergonomics practice and future research will be discussed.

Chapter 2

Work-site musculoskeletal pain risk estimates by trained observers - a prospective cohort study.

P. Coenen
I. Kingma
C.R. Boot
M. Douwes
P.M. Bongers
J.H. van Dieën

Ergonomics, 2012,
55(11):1373-1381.

ABSTRACT

Work-related musculoskeletal pain (MSP) risk assessments by trained observers are often used in ergonomic practice; however, the validity may be questionable. We investigated the predictive value of work-site MSP risk estimates in a prospective cohort study of 1745 workers. Trained observers estimated the risk of MSP (neck, shoulder or low-back pain) using a three-point scale (high, moderate and low risk) after observing a video of randomly selected workers representing a task group. Associations of the estimated risk of pain and reported pain during a three-year follow-up were assessed using logistic regression. Estimated risk of neck and shoulder pain did (odds ratio, OR: 1.45 (95% confidence interval, CI: 1.01–2.08); 1.64 (95% CI: 1.05–2.55)), however, estimated risk of low-back pain did not significantly predict pain (OR: 1.27 (95% CI: 0.91–1.79)). The results show that observers were able to estimate the risk of shoulder and neck pain, whereas they found it difficult to estimate the risk of low-back pain.

INTRODUCTION

Work-related musculoskeletal pain (MSP), which often affects the lower back, neck or shoulder region (Picavet & Schouten, 2003), is a great concern for society (Alexopoulos et al., 2004; Punnett et al., 2005). The high prevalence of MSP is associated with a loss of quality of life and high costs (e.g. medical costs, costs due to work absenteeism and costs due to a reduction of productivity while working during sickness, so-called presenteeism; Bot et al., 2005; Lambeek et al., 2011; Stewart et al., 2003). In addition to personal risk factors (e.g. age, gender; Côté, 2012; Leboeuf-Yde, 2004) and psychosocial risk factors (e.g. work pressure, social support and job satisfaction; Hartvigsen et al., 2004; van den Heuvel et al., 2005), several work-related physical risk factors were found to be associated with MSP. For example, trunk bending and twisting, lifting and whole body vibrations are associated with the occurrence of low-back pain (LBP; Hartvigsen et al., 2001; Tiemessen et al., 2008; van Nieuwenhuyse et al., 2006), whereas repetitive handling, extreme body postures (e.g. upper arm flexion and neck flexion), high forces or a combination of these factors are associated with neck and shoulder pain (Côté et al., 2008; Palmer & Smedley, 2007; van Rijn et al., 2010).

An important issue when assessing physical risk factors for MSP in epidemiological research and ergonomic practice is to choose an appropriate method of measurement (Burdorf, 2010; David, 2005). Work-related risk factors can be assessed by self-reports, observations (i.e. subjective risk estimations or structured observations of exposure variables) and direct measurements (e.g. muscle activity measurements, goniometry and measurement of external forces). Self-reports have been used in numerous epidemiological studies (e.g.; Balogh et al., 2001; Barrero et al., 2009a) and are easily applicable; however, their accuracy has been questioned (Balogh et al., 2004; Punnett & Wegman, 2004). Therefore, in contrast to epidemiological studies, self-reports of workload are rarely used for evaluation in ergonomic practice (Hansson et al., 2001). Instead, subjective risk estimations by observers are frequently used. Although these observations have higher validity than self-reports, their validity is assumed to be lower than obtained by direct measurement (Spielholz et al., 2001; Takala et al., 2010). Regrettably, when moving from self-report to direct measurement, cost and measurement time increase while feasibility decreases (Barrero et al., 2009b; David, 2005). Therefore, when selecting an appropriate measurement method in epidemiologic studies or in ergonomics practice, a trade-off between accuracy and feasibility should be considered.

When constructing a sound measurement strategy, besides choosing an appropriate measurement method, also the way of sampling exposure measurements (e.g. measuring over a single day or over multiple days) has to be chosen (Mathiassen et al., 2003b) and either a group or an individual measurement approach should be adopted (Jansen & Burdorf, 2003). Based on reviews, the predictive validity of measuring methods depends largely on the measurement strategy. For example, no differences in exposure-response associations for neck pain in studies using objective and subjective measurement methods (Fejer et al., 2006) have been reported, suggesting that objective measurements provide only limited additional predictive information, possibly as a result of inadequate or time-limited measurements (Palmer & Smedley, 2007). Structured observations and direct measurements may lack accuracy when using a poor measurement strategy, whilst self-reports and subjective risk estimations can be useful, especially when efficient measurement strategies are needed. Despite the abovementioned suggestions, the predictive validity of subjective risk estimations is unknown. Therefore, in the present study, data from a prospective study were used to investigate whether MSP risk estimates of workers in the workplace by trained observers were predictive for MSP (LBP, neck and shoulder pain). If proven to be valid, such subjective assessments could be useful for risk assessments in ergonomics practice and epidemiological research.

METHODS

Population

Data used in this study are part of the Study on Musculoskeletal disorders, Absenteeism and Health (SMASH) previously described in more detail (Ariëns et al., 2001; Hoogendoorn et al., 2000a). In short, the study is a prospective longitudinal assessment of MSP risk estimation and personal characteristics by trained observers for a cohort of workers at baseline and then by self-administered annual questionnaires during a three-year follow-up. Workers were recruited from 34 companies in the Netherlands representing several industrial and service branches, including metal, computer software, chemical, pharmaceutical, food and wood construction industries, as well as insurance companies, childcare centers, hospitals, distribution companies and road worker organizations. Thus, the study population included workers performing various tasks with a wide range of physical and mental workloads.

At baseline, 1990 of the invited 2048 workers participated in the study. A total of 1802 of the original 1990 participants completed all questionnaires. Forty-six workers were excluded because they were employed in their current job less than one year or worked less than 20 h a week. Eleven workers were excluded because they had had a paid job for a substantial amount of time at a company other than the one from which they were recruited. After exclusion, 1745 workers were eligible to participate in the current study on MSP risk estimations. The MSP risk estimation data were available for 1338 workers (Figure 2.1).

Data collection

At baseline, data were collected on personal factors (e.g. age and gender) by questionnaires and observers made MSP risk estimations as described in more detail in the next paragraph. At baseline and in a subsequent three-year follow-up, MSP prevalence (in the lower back, neck and shoulder regions) was assessed annually using a self-administered Dutch version of the Nordic Questionnaire for assessment of musculoskeletal symptoms (Kuorinka et al., 1987). Subjects were asked to indicate how often they had experienced neck, shoulder or LBP in the last 12 months: never, occasionally, regularly or prolonged. Musculoskeletal pain was defined when workers reported regular or prolonged pain in the 12 months prior to the completion of the questionnaire. Musculoskeletal pain during follow-up was defined as MSP in at least one of the three follow-up questionnaires. This definition of MSP was independent from MSP at baseline.

MSP risk estimation

For the risk estimations, workers were video-recorded at the workplace on four occasions, randomly selected over the course of a single workday. The duration of each video recording was 5–14 min depending on the variability of the worker's task. Observers allocated all workers to a total of 145 groups with similar tasks and physical loads based on the International Standard Classification of Occupations (1968). Videos of one fourth of the workers in each task group were randomly selected and were used for a structured observation protocol in which several kinematic exposure variables (e.g. trunk flexion angles and arm elevation angles) were assessed whilst replaying the video. After observing the video, the observers were asked: 'make an estimation of the risk of shoulder and neck pain and LBP respectively'. This estimated risk was expressed in three categories: low, moderate and high risk of pain. For all task groups, the modal estimated risk of the observed workers in a tasks group was assigned to all workers within that task group. This group approach has been shown to lead to efficient data collection that might even result in higher predictive individual estimates (Jansen & Burdorf, 2003; Spielholz et al., 2001) than individual exposure assessment.

All video observations were conducted by a group of 31 well-trained and experienced research assistants with significant knowledge on human kinesiology, recruited from a group of students of the Faculty of Human Movements Sciences of the VU University Amsterdam. The observers were trained to minimize inter-observer variation and ascertain the repeatability of kinematics using a structured video-observation protocol. However, observers were not specifically trained in making risk estimations.

Statistical analyses

Crude associations between risk estimates of neck pain, shoulder pain and LBP, and the actual reported prevalence of pain during follow-up were assessed using logistic regression analysis. In each analysis, the estimated risk was considered as independent variable (on an ordinal scale categorized as low, moderate or high risk for MSP) and the prevalence of self-reported pain during the three years of follow-up (regardless of MSP at baseline) as a dichotomous dependent variable. Associations of estimated MSP risk (for shoulder pain, LBP and neck pain) and pain during the follow-up were assessed in two ways, resulting in a total of six logistic regression analyses; associations using the lowest risk score as a reference were assessed, as well as the association of risk estimates across the three risk categories. Since observers may have incorporated the effect of confounders (e.g. age and gender) into their MSP risk estimates, we decided not to correct for confounders in the present study. All statistical analyses were performed using SPSS (version 17.0.1).

RESULTS

Population

The 1338 workers for who risk estimates were available had a mean age of 35.6 + 8.8 years and 74% were male. For this group, data on the prevalence of MSP during at least one of the three years of follow-up were available for 1005 workers (neck pain), 1038 workers (LBP) and 840 workers (shoulder pain), which is 75, 78 and 63%, respectively (Table 2.1; Figure 2.1). Specifically, during at least one of the three years of follow-up, 334 (32%) workers reported neck pain, 528 (51%) workers reported LBP and 187 (22%) workers reported shoulder pain.

Associations

Associations of the estimated risk and the reported prevalence of neck and shoulder pain were significant (Table 2.2). Workers with estimated high risk of neck or shoulder pain had a significantly higher reported prevalence of neck and shoulder pain compared to workers with estimated low risk of neck and shoulder pain (odds ratio, OR: 1.45 (95% confidence interval, CI: 1.01–2.08) and 1.64 (95% CI: 1.05–2.55), respectively). Furthermore, there was a significant trend of MSP across the three levels of estimated risk for neck and shoulder pain (OR: 1.20 (95% CI: 1.00–1.43) and 1.28 (95% CI: 1.03–1.59), respectively). In contrast, workers with estimated high risk of LBP did not report a significantly higher prevalence of LBP compared to workers with an estimated low risk of LBP (OR: 1.27 (95% CI: 0.91–1.79)). The risk estimates of LBP across the three risk levels were also not significantly associated with the reported prevalence of LBP (OR: 1.14 (95% CI: 0.96–1.35)).

DISCUSSION

Summary of findings and interpretation

The results of this study show that MSP risk estimates by trained observers were predictive for the occurrence of shoulder and neck pain, but not for LBP. Therefore, these estimates provide an assessment method that is crude, but useful for neck and shoulder pain risk assessment in ergonomics practice and in epidemiological studies.

Self-reports are often applied in epidemiologic studies while in ergonomic practice, subjective risk estimates by observers are more frequent. The subjective risk estimates are relatively cheap and easy to apply. However, it has been suggested that these estimates may be inaccurate because of the crude categorical scales (e.g., low, medium, high) often used (Burdorf, 2010; Spielholz et al., 2001), among other reasons.

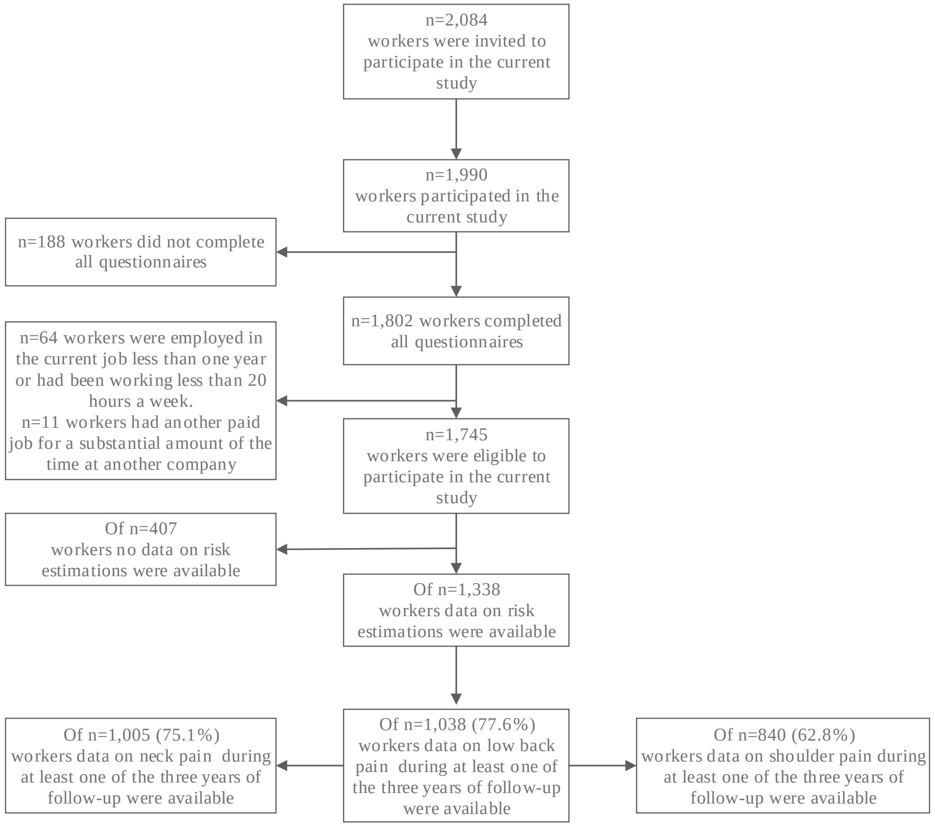

Figure 2.1 | Flow chart of the workers' inclusion process.

Although there are appropriate methods to analyze these ordinal scales (Svensson, 2001), categorization is highly dependent upon a number of factors (e.g., the number of categories used, boundaries of these categories) affecting the accuracy of the measurement that can lead to an underestimation of risk associations (Kociolek & Keir, 2010; Lowe, 2004). Despite reported inaccuracies, we found that the subjectively estimated risk for neck and shoulder pain did predict the occurrence of pain in our study. This might be due to large number of subjects who were observed during a substantial period of their work time. This hypothesis is underlined by reviews presenting comparable (Fejer et al., 2006; Palmer & Smedley, 2007) or even stronger exposure-response associations (Barrero et al., 2009a) in studies using subjective risk estimates compared to more objective measurement tools.

Our finding that risk estimates of LBP are not significantly associated with LBP prevalence corresponds with earlier studies questioning the accuracy of subjective risk estimates (e.g.; Balogh et al., 2004; Hansson et al., 2001). The fact that observers were able to make risk estimates of shoulder and neck pain, but not of LBP, may not directly be attributable to a more complicated causal mechanism. The etiology of MSP has only partly been revealed; it is highly likely that physical load as well as personal and psychosocial factors are involved. This holds not just for LBP, but also for neck and shoulder pain (Eatough et al., 2012; Hartvigsen et al., 2004; Holmstrom et al., 1992; van den Heuvel et al., 2005). More likely, physical loading of the lower back may be harder to assess through visual observation than the physical load on the neck and shoulder. Low-back load depends on a larger number of task variables (i.e. trunk posture, arm posture, load magnitude and load distance) than neck and shoulder load, which mainly depend on neck and shoulder flexion. The accuracy of assessing low-back load seems to be relevant since the risk of LBP was found to be associated with high low-back loads (Coenen et al., 2013b; Marras et al., 2010; Marras et al., 1995; Norman et al., 1998).

Methodological considerations

The results of the current study are based on a prospective cohort study of a large group of workers suggesting high methodological strength (Rothman & Greenland, 2005). A limitation of the present study is that the workers were only observed for a single day, which could alter the reliability of the MSP risk estimates, since it has been shown that variation in work exposure between days may occur (Mathiassen et al., 2003b; Svendsen et al., 2005). To obtain reliable exposure estimates, several sampling strategies can be chosen to reduce the measurement time without losing too much accuracy (e.g. sampling over multiple moments within or across days; Mathiassen et al., 2003b). The choice for a sampling strategy depends on the tasks to be distinguished, variation in exposure within and between days and the reliability of the measurement method chosen (Mathiassen et al., 2003b). For example, Liv and colleagues (2010) showed that when exposure data are correlated within days, efficiency can be improved by distributing the sample widely across the day or across days. We used four randomly selected observation moments for each worker over the course of a workday as it has been shown that a total of four observations are sufficient for group-based assessment of work exposures (Hoozemans et al., 2001). Not taking variability in exposure over days or weeks into account might result in an underestimation of the variability within persons. Nevertheless, as we assigned group-based risk estimates to each individual in a group, this was at least partially compensated by taking variability between subjects into account.

We observed a selection of workers for all task groups, while we assigned median task group values of MSP risk estimates to each individual within a specific task group. This group-based measurement approach is efficient and might lead to more reliable estimates of exposure, since random measurement errors may decrease compared to individual estimates of exposure (Hoozemans et al., 2001; Jansen & Burdorf, 2003). The choice for a group estimate was made based on a pilot study showing that for postural observations, the largest variation derives from within worker variation rather than between-worker variation (van der Beek & Frings-Dresen, 1998). This proposition was confirmed, after collecting the data, by showing small within-group variability and large between-group variability (Ariëns et al., 2001). Therefore, the choice of a group sampling approach in our study seems justified.

Table 2.1 | Descriptive statistics (number of workers (n), gender, age (in years), number of working hours per week and years of employment in the current job) are reported in the upper rows. Self-reported pain is described in the lower rows.

		Study population	Neck pain		LBP		Shoulder pain	
			Study population	Drop-outs	Study population	Drop-outs	Study population	Drop-outs
Descriptive statistics								
n		1338	1005	333	1038	300	840	498
Gender		m=992(74%) f=340	m=743(74%) f=259	m=249(75%) f=81	m=773(74%) f=262	m=219(74%) f=78	m=620(74%) f=217	m=372(75%) f=123
Age (Years)		35.6(8.8)	35.9(8.8)	34.5(8.4)	35.8(8.7)	34.7(8.9)	35.8(8.8)	35.1(8.7)
Hours per week		38.2(4.6)	38.2(4.7)	38.1(4.5)	38.3(4.7)	37.9(7.8)	38.3(4.6)	38.0(7.8)
Years in job		9.6(7.9)	9.9(7.9)	8.6(7.5)	9.8(7.9)	8.8(4.5)	9.8(7.9)	9.2(4.7)
Baseline pain								
Baseline	n		995	323	1028	290	828	488
	Pain		249(25%)	64(19%)	396 (38%)	71(24%)	91(11%)	43(9%)
Follow-up pain								
FU-1	n		996		1015		832	
	Pain		221(22%)		239(32%)		98(12%)	
FU-2	n		973		985		812	
	Pain		205(21%)		317(32%)		77(10%)	
FU-3	n		970		981		824	
	Pain		160(16%)		273(28%)		84(10%)	
At least one year of FU	n		1005		1038		840	
	Pain		334(32%)		528(51%)		187(22%)	

Number of workers who filled in the pain questionnaires and number of workers who reported pain at baseline in the three years of follow-up combined and during each of the three years of follow-up are reported. Percentages provided are the percentages of workers reporting pain (expressed as a percentage of the number of workers who filled in that specific questionnaire). Data of the group of workers who were eligible to participate in the current study (left column), of the groups of workers for whom data were included in the statistical analyses (for neck pain, LBP and shoulder pain, respectively) and those workers who dropped out the analyses (for neck pain, LBP and shoulder pain, respectively) have been described.

Table 2.2 | Associations (odds ratios) for the risk estimates (low, moderate, high) of MSP and the prevalence of MSP during the three years of follow-up in the neck, lower back and shoulders.

Risk Factor	Pain	No Pain	%pain	OR (95% CI)[1]	OR (95% CI)[2]
Neck (n=1046)					1.20 (1.00-1.43)*
Low Risk	58	144	29%	Reference	
Moderate Risk	137	286	32%	1.19 (0.82-1.72)	
High Risk	139	241	36%	1.45 (1.01-2.08)*	
Low-back (n=1120)					1.14 (0.96-1.35)
Low Risk	108	114	48%	Reference	
Moderate Risk	233	243	49%	1.01 (0.74-1.39)	
High Risk	186	154	55%	1.27 (0.91-1.79)	
Shoulders (n=872)					1.28 (1.03-1.59)*
Low Risk	43	194	18%	Reference	
Moderate Risk	83	291	22%	1.29 (0.85-1.94)	
High Risk	61	168	26%	1.64 (1.05-2.55)*	

Both associations taking the lowest risk category as a reference category and associations across all three risk categories are reported

*denotes a significant association of the estimated risk and the reported pain

%pain = percentage of subjects with MSP within the groups of estimated risk of MSP

OR = Odds Ratio

CI = confidence interval

[1] = Associations of three levels of risk using the lowest group as reference

[2] = Associations across the three levels of estimated risk

In our study, associations have been assessed using ORs. It is generally known (e.g.; Twisk, 2003) that ORs can lead to overestimations of relative risks when the prevalence of the dependent variable is high. However, the use of ORs in epidemiological studies is widely accepted. Furthermore, in the present dataset, calculation of risk associations instead of ORs resulted in comparable conclusions (non-reported data).

In this study, consistent with earlier work (Ariëns et al., 2001; Hoogendoorn et al., 2000a), MSP was defined when workers reported regular or prolonged pain in the last 12 months in at least one of the three annual follow-up questionnaires. The prevalence of pain according to this definition is relatively high (32, 51 and 22% for neck pain, LBP and shoulder pain, respectively; Table 2.1). Because of this high prevalence, it is expected that the group of workers reporting prolonged pain in the last 12 months is a heterogeneous group that might attenuate associations with the estimated risk of MSP. It could be that a more strict definition, for example, taking pain severity into account, would have led to stronger associations. Workers with MSP at baseline were included in the current analysis, in contrast with earlier studies on this study population (Ariëns et al., 2001; Hoogendoorn et al., 2000a). Since it is known that recurrence is a typical characteristic of MSP (Hestbaek et al., 2006; van Oostrom et al., 2011), excluding workers with pain at baseline seems rather arbitrary, since it cannot be excluded that workers without complaints at baseline had pain in previous years. Moreover, risk estimates cannot be affected by previous MSP, as observers were not aware of these estimates. Excluding workers with MSP in the past might therefore enhance the healthy worker effect while reducing the external validity of the results. Including these workers, therefore, seems reasonable.

Data on MSP risk estimates and on the reported prevalence of MSP during at least one of the three years of follow-up were available for 1338 workers who reported neck pain (75%), LBP (78%) and shoulder pain (63%; Table 2.1). This rather substantial loss to follow-up could possibly have led to selection or attribution bias. However, descriptive statistics show that the group of workers who dropped out of the cohort during the three years of follow-up did not differ considerably in terms of gender, age and working hours a week (Table 2.1), which renders such bias unlikely. At baseline in the group of workers with follow-up data, pain was slightly higher compared to the group of dropouts, suggesting the opposite healthy worker effect.

We did not correct for confounders, such as age and gender in the analysis. It is plausible that observers incorporated the effect of these confounders in their MSP risk estimates. For example, it is possible that observers, in general, rate the risk of a task differently when it is performed by an old lady compared to a young man. As this already results in an implicit correction for these confounders, extra correction for these confounders seems redundant. Furthermore, group estimates were assigned to all members of each task group, which diminishes the effect of these confounders.

Furthermore, the MSP risk estimation was conducted by observers who were trained to make systematic observations of work postures. It has been shown that postural observations are sufficiently reliable in work-site situations (Bao et al., 2009; van der Beek & Frings-Dresen, 1998). However, since regrettably no inter- and intra-observer reliability tests were performed for the risk estimates, differences in estimation between observers might have occurred. Finally, observers had substantial knowledge of ergonomics and human kinesiology; however, they were not specifically trained to make risk estimations. Ergonomic practitioners may be better trained to make such risk estimations. Therefore, the present results refer to judgments made by observers trained for postural observations and these estimates may not necessarily be the same as judgments by ergonomics experts.

CONCLUSION

From the present study, it can be concluded that trained observers are able to estimate the risk of neck and shoulder pain, however, observers have difficulty predicting an increased risk of LBP. Risk estimation of trained observers, therefore, provides a method that is crude but useful for neck and shoulder pain risk assessment in ergonomics practice and in epidemiological studies.

Chapter 3

Cumulative low-back load at work as a risk factor of low-back pain: a prospective cohort study.

P. Coenen
I. Kingma
C.R. Boot
J.W. Twisk
P.M. Bongers
J.H. van Dieën

Journal of Occupational Rehabilitation, 2013,
23(1): 11-8

ABSTRACT

Much research has been performed on physical exposures during work (e.g. lifting, trunk flexion or body vibrations) as risk factors for low-back pain (LBP), however results are inconsistent. Information on the effect of doses (e.g. spinal force or low-back moments) on LBP may be more reliable but is lacking yet. The aim of the present study was to investigate the prospective relationship of cumulative low-back loads (CLBL) with LBP and to compare the association of this mechanical load measure to exposure measures used previously.

The current study was part of the Study on Musculoskeletal disorders, Absenteeism and Health (SMASH) study in which 1,745 workers completed questionnaires. Physical load at the workplace was assessed by video-observations and force measurements. These measures were used to calculate CLBL. Furthermore, a 3-year follow-up was conducted to assess the occurrence of LBP. Logistic regressions were performed to assess associations of CLBL and physical risk factors established earlier (i.e. lifting and working in a flexed posture) with LBP. Furthermore, CLBL and the risk factors combined were assessed as predictors in logistic regression analyses to assess the association with LBP.

Results showed that CLBL is a significant risk factor for LBP (OR: 2.06 (1.32–3.20)). Furthermore, CLBL had a more consistent association with LBP than two of the three risk factors reported earlier.

From these results it can be concluded that CLBL is a risk factor for the occurrence of LBP, having a more consistent association with LBP compared to most risk factors reported earlier.

INTRODUCTION

In the past decades, epidemiological studies have contributed to our understanding of the etiology of low-back pain (LBP). Risk factors for the occurrence of LBP, can roughly be divided into: personal factors (e.g. age, smoking habits, physical capacity and body weight; Hamberg-van Reenen et al., 2007; Leboeuf-Yde, 2004; Manek & MacGregor, 2005; Wai et al., 2008), psychosocial factors (e.g. stress, social support and job satisfaction; Hartvigsen et al., 2004; Hoogendoorn et al., 2000b; Linton, 2001; Macfarlane et al., 2009), and physical factors (Bakker et al., 2009; Griffith et al., 2012; Hoogendoorn et al., 1999; Kuiper et al., 1999). Of these physical factors, twisting, bending, lifting and whole body vibrations are the most frequently reported ones associated with LBP (Hoogendoorn et al., 2000a; Tiemessen et al., 2008; van Nieuwenhuyse et al., 2006). Nevertheless, some recent reviews suggest that the evidence for a relationship between physical risk factors and LBP is not convincing (Bakker et al., 2009; Wai et al., 2008), and generally, data on exposure-response relationships are scarce and incomplete. It can be argued that the relationships of these physical exposures with LBP might be less reliable than the relationship of low-back load dose (i.e. the effect that physical exposure has in the human body) with LBP, since different exposures (e.g. lifting and bending) affect the same dose (Burdorf, 2010). While parameters of low-back load, like low-back moments or spine compression forces, could be used as such dose measures, information on the dose–response relationship of LBP is limited. Marras et al. investigated the predictive value of a variety of parameters of low-back loading with the risk of LBP (Marras et al., 2010; Marras et al., 1995). Moreover, some other studies suggest that cumulative loads acting on the spine may contribute to LBP (Kumar, 1990; Neumann et al., 2001a; Norman et al., 1998), however, these results are based on retrospective studies. Dose–response relationships obtained from prospective cohort studies have never been reported. The aim of the present study therefore was to investigate the association of cumulative low-back load (CLBL) with LBP, in a large prospective cohort study. Furthermore, the association with LBP of this dose estimate will be compared to associations for exposures reported earlier to be related to LBP. We hypothesized that CLBL, quantified in terms of low-back moments, is associated with LBP and that the association of this dose measure with LBP is more consistent than that of exposure measures that were previously established as risk factors for LBP.

STUDY POPULATION AND METHODS

Population

Data used in this study are part of the Study on Musculoskeletal disorders, Absenteeism and Health (SMASH), a prospective cohort study among Dutch workers on risk factors of musculoskeletal disorders. The study was approved by the medical ethical committee of the Netherlands organization for applied scientific research (TNO). The SMASH study, in which workers from 34 companies with both blue-collar and white-collar jobs from different parts of the Netherlands participated, has been described in more detail previously (Ariëns et al., 2001; Hoogendoorn et al., 2000a).

At baseline 1990 of the 2,048 workers who were invited for the study participated. 1,802 (91 %) of these workers completed all questionnaires at baseline. Forty-six workers were excluded because they had been employed in their current job <1 year or had been working <20 h a week. Eleven workers were excluded because they had another paid job for a substantial amount of time at another company than at which they were recruited. As a result, 1,745 workers were eligible to participate in the current study. Descriptive statistics of these workers are provided in Table 1.

Table 3.1 | Descriptive statistics (number of workers, gender, age, working hours per week and years of employment in the current job) of group of the workers who were eligible to participate in the current study (left column), workers of whom data were included in the statistical analysis (middle column) and workers of whom data were excluded from the statistical analysis (right column).

	Baseline workers	Workers in analysis	Workers not in analysis
N	1745	1086	659
Gender	m=1222 (71%) / f=510	m=759 (70%) / f=327	m=463 (72%) / f=183
Age (years)	35.9 (8.4)	35.6 (8.7)	35.4 (8.9)
Hours per week	38.3 (4.5)	38.2 (4.7)	38.2 (4.7)
Years in current job	9.9 (7.7)	9.6 (7.6)	9.5 (8.0)

Data Collection

At baseline, a number of potential risk factors were measured; questionnaire data were collected and assessment of physical load at the workplace was performed. Furthermore, a 3-year follow-up was conducted in which the prevalence of LBP was assessed annually.

Physical work load was assessed by video-observations and force measurements at the workplace. External force exertion at the hands was measured using force transducers or a weighting scale. Furthermore, workers were video recorded at their workplace during four randomly selected moments of a workday. Each video-recording lasted 5–14 min, depending of the variability in working tasks. Thirty-five observers were recruited from a group of university students of the Faculty of Human Movement Sciences from the VU University Amsterdam. These observers had considerable knowledge on human kinematics and were trained using a standardized protocol to perform structured postural observations. These well trained observers allocated all workers in task groups based on similar tasks and loads according to the International Standard Classification of Occupations. A continuous systematic observation of the video-recordings was used to assess trunk sagittal flexion, arm sagittal elevation, trunk rotation (in the transverse plane) and the presence of an external force in one-fourth of the workers of each task group. Furthermore, the time spend in a sitting position was observed. All data were extrapolated to an 8 h work day. A detailed description of these procedures was given by Hoogendoorn et al. (2000a).

Personal factors such as age and gender were assessed using self-administered questionnaires. A Dutch version of the Karasek's Job Content Questionnaire for psychosocial work characteristics was used to assess job demands, decision authority, co-worker support and supervisor support (Karasek, 1985). The psychometric properties and the construction of these scales have been described by de Jonge et al. (2000). Exercise behavior during leisure time was assessed with the Leisure Time Exercise Questionnaire (Godin et al., 1986). Furthermore, driving a vehicle during work and during leisure time, flexion and rotation of the trunk and moving heavy loads during leisure time were assessed with the Loquest questionnaire (Hildebrandt & Douwes, 1991). A detailed description of all questionnaires has been given earlier (Ariëns et al., 2001; Hoogendoorn et al., 2000a).

At baseline and at each year of the follow-up, the occurrence of LBP was assessed using a self-administered, adapted version of the Nordic Questionnaire (Kuorinka et al., 1987). LBP at baseline was defined when subjects reported regular or prolonged LBP in the previous 12 months before the start of the study. LBP during follow-up was defined as regular or prolonged LBP in the previous 12 months in at least one of the three annually follow-up questionnaires. The baseline population consisted of workers with and without LBP.

Assessment of Low-back Load

For the assessment of CLBL during work, a manikin consisting of a trunk/head, upper arm and a lower arm/hand segment was constructed based on segment orientations obtained from the continuous video-observations (Table 3.2) and segment anthropometrics. As observed postures were supposed to be representative for the task group, average body weight and length within each task group were used for the estimation of segment anthropometrics (segment mass, length and centre of mass; de Leva, 1996; Dumas et al., 2007) and an estimation of the L5-S1 position (de Looze et al., 1992) using regression equations.

For the complete observed period, a top–down calculation of net moments at the L5-S1 joint was performed using a general equation of motion (Hof, 1992). In this calculation, segment gravitational forces of the constructed manikin combined with the measured external forces were taken into account. The calculated moments in the lower back were squared to accommodate for the fact that the moment levels have larger effect on injury risk than the number of repetitions (Brinckmann et al., 1988). Subsequently, CLBL was assessed by calculating the area under the moment curve. Mean task group values of the CLBL during the observed period were assigned to all workers in the same task group and were extrapolated to an entire work week based on the number of working hours of each individual in that task group during a week. All calculations were performed using custom developed Matlab software (version 7.7.0).

Table 3.2 | Observational categories. The table shows a description and corresponding values for the observed variables. The last column shows body orientation values that were used for the calculation of CLBL.

Variable	Observation		CLBL Calculation
	Description	Category	Values
Trunk Flexion	Neutral	<30 degrees	0 degrees
(sagittal plane)	Mild Flexion	30-60 degrees	45 degrees
	Extreme Flexion	60-90 degrees	75 degrees
	Very Extreme Flexion	>90 degrees	90 degrees
Trunk Rotation	Neutral	<30 degrees	0 degrees
(transverse plane)	Twisting	>30 degrees	30 degrees
Arm Elevation	Neutral	<30 degrees	15 degrees
(sagittal plane)	Mild Elevation	30-60 degrees	45 degrees
	Extreme Elevation	60-90 degrees	75 degrees
	Very Extreme Elevation	>90 degrees	90 degrees

CLBL = Cumulative low-back load

Statistical Analyses

The crude effect of CLBL (categorized into five categories, based on 20th percentiles -quintiles-) on LBP was assessed using a logistic regression with LBP during the follow-up (independent of LBP at baseline) as dependent variable, calculating ORs and corresponding 95% CI. The choice for the number of categories is a balance between the power requirements (a sufficient number of workers in each category should remain) and optimizing contrast between the categories. The relationship of CLBL and LBP was checked on linearity by comparing regression coefficients between quintiles. In case of a linear relationship, logistic regression analyses were performed using CLBL as a continuous variable rather than categorized into five categories. In line with earlier reports on the present population (Burdorf, 2010), the variables age, gender, exercise behavior during leisure time, quantitative job demands, decision authority, skill discretion, supervisor support, co-worker support, driving a vehicle during work and leisure time, flexion/rotation of the trunk during leisure time and moving heavy loads during leisure time were considered confounders. A second logistic regression analysis was performed to calculate ORs and corresponding 95% CI for CLBL (independent variable) on LBP during the follow-up (dependent variable), adjusted for these confounders.

To compare the association of the dose measure CLBL with LBP during the follow-up to exposure measures reported earlier, six additional logistic regression analyses were performed. The earlier found risk factors percentage of the working time in a flexed position, number of lifts in an 8 h working day, and number of lifts ≥25 kg in an 8 h working day were used for comparison since they were reported to be significant risk factors for LBP in the same study population earlier (Hoogendoorn et al., 2000a). In the first three analyses, the three exposures reported earlier were separately used as

Table 3.3 | Association of CLBL with LBP based on logistic.

	Regression Model			
Risk Factor	LBP	No LBP	B	OR (95% CI), n=1086†
CLBL				
1st quintile	109	107		Reference
2nd quintile	106	122	-0.15	0.86 (0.59-1.25)
3th quintile	93	129	-0.34	0.71 (0.49-1.04)
4th quintile	93	107	-0.15	0.86 (0.59-1.26)
5th quintile	136	84	0.47	1.60 (1.10-2.35)*
CLBL				
1st quintile				#Reference
2nd quintile			0.05	1.05 (0.70-1.58)
3th quintile			-0.13	0.87 (0.57-1.33)
4th quintile			0.03	1.03 (0.68-1.57)
5th quintile			0.72	2.06 (1.32-3.20)*

B = regression coefficient, OR = Odds Ratio, CI = confidence interval

†Of 1086 workers data on the occurrence of LBP during follow-up, physical exposure at work and all confounders were available.

#Logistic regression adjusted for the confounders: age, gender, exercise behaviour during leisure time, quantitative job demands, decision authority, skill discretion, supervisor support, co-worker support, driving a vehicle during work and leisure time, flexion/rotation of the trunk during leisure time and moving heavy loads during leisure time

CLBL = Cumulative low-back load

LBP = Low-back pain

Table 3.4 | Category values of the five different categories (based on quintiles).

	Category values				
	n	Minimum	Maximum	Mean	Std.
1th quintiles	216	0.09	0.49	0.29	0.11
2nd quintiles	228	0.52	0.71	0.62	0.05
3th quintiles	222	0.74	1.13	1.03	0.13
4th quintiles	200	1.14	1.96	1.52	0.29
5th quintiles	220	1.99	10.83	3.65	2.38
Total	1086	0.09	10.83	1.43	1.16

Number of subjects (n), minimum and maximum, mean and standard deviation of CLBL (all in MNm) in all five quintiles are listed

CLBL = Cumulative low-back load

LBP = Low-back pain

SD = Standard deviation

independent variables consecutively, without and with correction for CLBL. In the other three analyses CLBL was used as independent variable corrected for one of the three above mentioned physical risk factors, consecutively. Associations of all risk factors with LBP separately and corrected as indicated above were compared to assess the risk factor with the most consistent association with LBP. All statistical analyses were performed using SPSS (version 17.0.1).

RESULTS

Population

Of the 1,745 workers eligible for participation in the current study, data on the physical load at workplace were available for 1,463 workers, while data on the occurrence of LBP in at least one follow-up measurement were available for 1,196 workers. For 1,192 workers, data on both physical load at workplace and on the occurrence of LBP were available. Of 1,086 workers, data on physical load at work, the occurrence of LBP and all confounders were available. 416 of these workers (38%) reported LBP at baseline and 537 workers (49%) reported LBP during at least one of the 3 years of follow-up. Data of these workers were used for further analysis (Table 3.1). In contrast to earlier work on the same population (Hoogendoorn et al., 2000a), workers with LBP at baseline were included in the present study.

LBP Risk Model

The regression coefficients of the five CLBL categories, obtained from the logistic regression analyses, revealed a non-linear relationship of CLBL and LBP (Table 3.3). Therefore, categorized CLBL into quintiles (Table 3.4) was used as independent variable in the logistic regression models. A significant crude relation of CLBL and LBP in the group with the highest CLBL compared to the group with the lowest CLBL was shown (OR of 1.60, 95% CI: 1.10–2.35). Also, CLBL adjusted for confounders yielded a significant relationship with the occurrence of LBP in the group with the highest CLBL compared to the group with the lowest CLBL (OR: 2.06, 95% CI: 1.32–3.20; Table 3.3).

To assess the predictive value of CLBL for LBP in comparison to exposures reported earlier, additional logistic regression analyses were performed in which these three risk factors were used as independent variables. Logistic regression analyses adjusted for confounders showed that all three risk factors significantly predicted LBP with ORs of 2.35 (1.46–3.79), 2.22 (1.33–3.36) and 2.38 (1.48–3.82) respectively in the most exposed groups (Table 3.5). However, when corrected for confounders and CLBL, only lifting >15 times ≥25 kg in an 8 h working day compared to no lifts of ≥25 kg was a significant risk factor for LBP (OR: 2.03 (1.23–3.36)), while percentage of the working time in a flexed position and number of lifts in a 8 h working day did not significantly predict LBP. Moreover, when separately corrected for each of these three risk factors, the CLBL remained a significant predictor for LBP in the group with the highest CLBL compared to the group with the lowest CLBL, showing ORs of 1.89 (1.04–3.45), 1.96 (1.15–3.36) and 1.85 (1.17–2.92) respectively (Table 3.5).

Table 3.5 | Associations of the three earlier found risk factors (percentage of the working time in a flexed position, number of lifts in a 8 h working day, number of lifts ≥25 kg in a 8 h working day) with LBP based on logistic regression, adjusted for confounders (left columns) and adjusted for confounders and CLBL (right columns). Besides, association of CLBL with LBP adjusted for all earlier found risk factors separately are shown

Risk Factor	LBP	No LBP	OR (95% CI), n=1086†	OR (95% CI), n=1086†
Time in trunk flexion				
≤5% time ≥30°	256	287	[#]Reference	[¤]Reference
5-10% time ≥30°	96	110	1.01 (0.73-1.47)	1.15 (0.74-1.78)
>10% time ≥30° & ≤5% time ≥60°	120	120	1.15 (0.83-1.58)	0.91 (0.57-1.46)
>5% time ≥60°	65	32	2.35 (1.46-3.79)*	1.45 (0.77-2.73)
Number of lifts				
Never	151	161	[#]Reference	[¤]Reference
Never ≥10 kg/working day	81	94	0.74 (0.50-1.09)	0.69 (0.45-1.06)
Never ≥25 kg/working day	146	156	0.96 (0.68-1.36)	0.77 (0.51-1.17)
1-15 times ≥25 kg/working day	96	107	0.86 (0.59-1.27)	0.73 (0.44-1.19)
>15 times ≥25 kg/working day	63	31	2.22 (1.33-3.72)*	1.60 (0.88-2.92)
Number of lifts ≥25 kg				
Never	378	411	[#]Reference	[¤]Reference
1-15 time/working day	96	107	0.93 (0.67-1.29)	0.92 (0.63-1.34)
>15 times/working day	63	31	2.38 (1.48-3.82)*	2.03 (1.23-3.36)*
CLBL				
1st quintile				[1]Reference
2nd quintile				1.06 (0.70-1.59)
3th quintile				0.83 (0.51-1.33)
4th quintile				1.03 (0.60-1.78)
5th quintile				1.89 (1.04-3.45)*
CLBL				
1st quintile				[2]Reference
2nd quintile				0.97 (0.62-1.51)
3th quintile				0.88 (0.55-1.41)
4th quintile				1.05 (0.62-1.76)
5th quintile				1.96 (1.15-3.36)*

Continuation of table 3.5

CLBL	
1st quintile	[3]Reference
2nd quintile	1.06 (0.71-1.60)
3th quintile	0.85 (0.56-1.31)
4th quintile	0.99 (0.62-1.57)
5th quintile	1.85 (1.17-2.92)*

B = regression coefficient, OR = Odds Ratio, CI = confidence interval †Of 1086 workers data on the occurrence of LBP during follow-up, physical exposure at work and all confounders were available.

[#]Adjusted for the confounders: age, gender, exercise behaviour during leisure time, quantitative job demands, decision authority, skill discretion, supervisor support, co-worker support, driving a vehicle during work and leisure time, flexion/rotation of the trunk during leisure time and moving heavy loads during leisure time.

[º]Adjusted for both abovementioned confounders and CLBL.

[1]Adjusted for both abovementioned confounders and 'Percentage of the working time in a flexed position'

[2]Adjusted for both abovementioned confounders and 'Number of lifts in an 8 hour working day'

[3]Adjusted for both abovementioned confounders and 'Number of lifts ≥25kg in an 8 hour working day'

DISCUSSION

The first aim of the present study was to investigate whether a low-back load dose, in this study expressed in CLBL is a predictor for LBP among workers. In the results, CLBL showed a significant association with the occurrence of LBP in the group with the largest CLBL. From these findings we can conclude that CLBL is a significant predictor of LBP. However, a significantly higher risk of LBP is only shown in the group with the highest levels of CLBL, which are levels of 2.00 MNm and more. As an example, for a moderate lifting task that would lead to a low-back load of 200 Nm, this level of CLBL will be reached when 2.000.000/2002 = 50 of these lifts are performed during a work week. Ergonomic interventions should therefore be targeted mainly to workers who encounter these levels of CLBL which can emerge from combinations of awkward postures and/or high exposure tasks at work.

The second aim, to compare the association with LBP of CLBL to risk factors reported earlier, was attained using additional logistic regression analyses. These results show that CLBL remains a significant risk factor of LBP when corrected for the earlier found risk factors. Moreover, while the risk factors reported earlier are significant risk factors for LBP when corrected for confounders, only one risk factor remains significant when corrected for both confounders and CLBL. From these results we can conclude that CLBL has a more consistent association with LBP than the risk factors time in a flexed position and number of lifts in a working day. This finding supports our hypothesis that a low-back load dose measure provides a stronger relationship with LBP than exposure measures of low-back load since several exposures (e.g. lifting and bending) are incorporated in the dose. The fact that the risk factor number of lifts ≥25 kg in an 8 h working day had a comparable association with LBP may indicate that this exposure metric reflects incidental peak loads which may constitute an independent risk for LBP. Again, this underscores the importance of focusing on peak loads.

Methodological Considerations

The strength of the present study is that the results are based on a large prospective cohort study. This design, in which the prevalence of LBP was measured during a 3-year follow-up allows insight into potential causes of LBP (Rothman & Greenland, 2005). Of the 1,745 workers who were eligible to participate in this study, data on physical load at the workplace, on the occurrence of LBP and on confounders were available for 1,086 workers. Selection or attribution bias may be possible due to this substantial loss to follow-up. However, the group of workers analyzed and the group of workers who were excluded from the statistical analysis due to incomplete data show comparable descriptive characteristics with respect to age, gender, working hours per weeks and years of employment (Table 3.1), thereby reducing the likeliness of these kinds of biases.

In contrast to earlier studies on this study population (Hoogendoorn et al., 2000a), workers suffering from LBP at baseline were included in our analyses. It has been shown that a history of LBP is a good predictor of future LBP since LBP often comes in several episodes (Smedley et al., 1997; van Tulder et al., 2002). Excluding workers with pain at baseline thus seems unreasonable since it cannot be excluded that workers without complaints at baseline, have not had any complaints 2 or 3 years before the baseline measurements. Therefore, we can assume that when excluding these workers, the healthy worker effect will be reinforced. Besides, including workers with a history of LBP makes the present results applicable to a larger part of the working population since excluding these workers would reduce the external validity of the current results. Including workers with pain at baseline seems therefore reasonable. Furthermore, an extra analysis in which only the workers without baseline complaints were analyzed (i.e. the workers who did not report LBP at baseline) showed changes in ORs<0.1 in the associations of CLBL with LBP. These findings, showing that associations of CLBL and LBP do not change considerably, support the consistency of the current results.

A limitation of the present study is the subjective assessment of LBP. It has been shown that diagnosing LBP is complicated. However, subjectively assessed LBP has been shown to have a strong relation with clinically examined LBP (Holmstrom & Moritz, 1991) and sickness absence due to LBP (Roelen et al., 2010). Furthermore, the CLBL assessment method contains some limitations. First, observations based on videos may suffer from errors and potential bias (van der Beek & Frings-Dresen, 1998). Furthermore, movements which are not in the sagittal plane are difficult to assess (Paul & Douwes, 1993) and the outcome of the measurement is dependent on the selected time at the measurement-day, the number of subjects per task group and the number of measurements per subject (Hoozemans et al., 2001). The latter problems were addressed by measuring workers at four random chosen moments of the day and measuring several workers in each task group, to obtain more precise estimates of the exposure within groups (Kromhout et al., 1996). Structured postural observations have been performed by multiple observers. Although, it has been shown that postural video observations are reliable among observers in work-site situations (Bao et al., 2009; van der Beek et al., 1992), inter-observer reliability was not evaluated in the group of observers we recruited. Therefore, because several trained observers classified the body postures, inter- and intra-observer variation cannot be ruled out.

Another source of error in our study might have emerged from the fact that workers were observed at four randomly chosen occasions of the work day for a finite amount of time rather than a complete observation of the whole work day. This choice was made based on a pilot study, in which it has been shown that the largest amount of variation in physical work exposure, is variation in exposure within workers rather than variation in exposure between workers (van der Beek et al., 1994). The appropriateness of our measurement strategy was furthermore supported by showing small within group variability and large between group variability in data on the same cohort (Ariëns et al., 2001). Measuring on multiple occasions on a single work day is therefore considered a feasible and justifiable approach to reduce the amount of observation time. Furthermore, it has been shown that measuring work load at four occasions during a day is sufficient to obtain a reliable estimate of the work exposure (Hoozemans et al., 2001).

A final source of error of the CLBL assessment results from the biomechanical calculation, which contains assumptions concerning the workers' anthropometrics and segment orientations. Furthermore, segment dynamics were not taken into account in this calculation, which may have led to an underestimation of the calculated low-back load. The above mentioned sources of errors in the calculation of CLBL suggest that associations of dose measures with LBP might become even higher when more reliable dose estimates are available. Besides, as an indicator of back load, low-back moments were used, although it may be argued that injury risk and thus potentially LBP is more accurately predicted by spinal forces, either in compression (van Dieën et al., 1999) or shear direction (Marras et al., 2010; Norman et al., 1998). However, a strong correlation of low-back moments with shear forces and compression forces has been reported (van Dieën & Kingma, 2005) reducing the risk of large errors due to the use of moments instead of spine forces. Comparison with Previous Findings The relationship between awkward body postures during work (e.g. trunk flexion, trunk rotation and lifting) and LBP has been reported in several prospective studies in the last decades (Hoogendoorn et al., 2000a; van Nieuwenhuyse et al., 2006). However, several reviews (Bakker et al., 2009; Griffith et al., 2012; Kuiper et al., 1999; Wai et al., 2010) showed that results are inconsistent. The association of low-back load dose measures and the risk of LBP can give more insight in the etiology of LBP. An association of cumulative and peak low-back load with LBP has been has been described before (Kumar, 1990; Neumann et al., 2001a; Norman et al., 1998). However, these associations are based on retrospective studies. The present results are comparable to the earlier findings and thus confirm these findings in a prospective study, thereby providing strong support for a causal relationship between CLBL and LBP.

CONCLUSIONS

From the current study it can be concluded that CLBL is a significant risk factor for LBP with more consistent associations with LBP than risk factors reported earlier. Moreover, CLBL appeared to reflect both the effects of working in a trunk flexed position and number of lifts during work on LBP risk. The risk factor number of lifts ≥25 kg had additional value in predicting the risk of LBP besides CLBL. The results of the present study may have implications for prevention programs for LBP. Interventions aimed at changes in posture and lifting forces, but also reduction of duration of exposure to adverse postures should, according to these findings be considered.

Chapter 4

The contribution of load magnitude and number of load cycles to cumulative low-back load estimations: a study based on in-vitro compression data

P. Coenen
I. Kingma
C.R. Boot
P.M. Bongers
J.H. van Dieën

Clinical Biomechanics, , 2013,
27(10):1083-1086

ABSTRACT

Cumulative low-back load is suggested to be associated with low-back pain, possibly due to (micro-)fractures of spinal segments. Based on available in vitro data it can be assumed that, in order to predict spine segment failure from cumulative compressive loading, load magnitude should be weighted with an exponent higher than one, whereas the number of cycles should be weighted with an exponent lower than 1. The aim of the present study was to assess both exponents based on available in-vitro data.

Data on loading to fatigue fracture of spinal segments under cyclic compression in-vitro were used and converted to survival probability for 5 load levels and 5 levels of number of cycles. Three optimization procedures were used to estimate the exponent of load magnitude and load cycles separately, and load magnitude and load cycles combined. Goodness of fit was assessed by comparing the Akaike's Information Criterion (AIC) between models.

The best fit, based on AIC and average error per data point was obtained with weighting of load magnitude and number of load cycles with exponents of approximately 2.0 and 0.2, respectively.

The results show that a combination of load magnitude and number of load cycles weighted with exponents of approximately 2 and 0.2 respectively provides a suitable measure of cumulative spinal compression loading. This finding may be of relevance for assessing cumulative low-back loads in studies on the etiology of low-back pain.

INTRODUCTION

High mechanical loads on the lower back during manual material handling have been associated with low-back pain (LBP; da Costa & Vieira, 2010; Lötters et al., 2003), possibly due to spinal segment (micro-)fractures (Marras et al., 1993; van Dieën et al., 1999). In addition to peak low-back loading, cumulative low-back load (CLBL) has been suggested to be associated with LBP (Kerr et al., 2001; Norman et al., 1998)

The most common way to calculate CLBL is a linear approach of integrating back load time series (F(t)) during a given period (Callaghan et al., 2001; de Looze et al., 1996; Marras et al., 2010; Norman et al., 1998):

$$\mathrm{Load_{cum}} = \int_{0}^{T} F(t)\ d \qquad \textit{Equation 4.1}$$

which can be simplified to (Kumar, 1990):

$$\mathrm{Load_{cum}} = \sum_{i=1}^{n} N_{\mathrm{cycles}}\,(i) \cdot F(i) \text{ for } i=1,2,\ldots n \qquad \textit{Equation 4.2}$$

in which the (peak) low-back load magnitude of a given work task (F) is multiplied by the number of load cycles (N_{cycles}) of that work task, while these multiplications of all tasks during a work shift (n) are summed. However, it has been argued that high force has more impact on the increase in failure risk than in a high number of cycles (Brinckmann et al., 1988). For example, 15 cycles of 2000 N load would cause a higher risk than in 20 cycles of 1500 N. Thus, alternative calculations of CLBL have been suggested. For example, a linear approach after application of a low-pass filter to spinal loading time series has been suggested by Krajcarski and Wells (2008). Furthermore, non-linear calculations have been suggested as well, for example second order (Seidler et al., 2009; Seidler et al., 2001; Seidler et al., 2003) or fourth order weighting of load magnitude (Jäger et al., 2000), and polynomial calculated CLBL (Parkinson & Callaghan, 2007). Based on this diversity in CLBL calculations, it can be concluded that it is unclear yet how the magnitude of the low-back load contributes to CLBL. Moreover, to our knowledge, number of load cycles is to date always implemented linearly in measures of CLBL. However, visual inspection of in-vitro data (Brinckmann et al., 1988; Hansson et al., 1987; Rapillard et al., 2006) suggests that the contribution of number of load cycles is highly non-linear as well. The aim of this study was therefore to determine the contribution of low-back load magnitude and number of load cycles in CLBL calculations, based on risk of tissue failure. To this end, results of in vitro fatigue failure spine compression experiments of Brinckmann et al. (1988) were used.

METHODS

Analyses of the present study are based on data collected by Brinckmann et al. (1988) who conducted a compression fatigue loading protocol on seventy lumbar motion segments.

First, failure load was established by applying compression in one randomly selected motion segment from each spine until fracture occurred. The mean ultimate strength of all specimen was estimated to be 5.24 (2.07) kN, ranging from 1.80 to 10.40 kN. The remaining motion segments of each spine were tested cyclically in a fatigue testing protocol until fracture or to a maximum of 5000 cycles. For all cyclically loaded motion segments, we derived load level and number of cycles to failure from the original publication. Load range was expressed as a percentage of the predicted ultimate strength. All methodological procedures have been described in detail previously (Brinckmann et al., 1988).

Motion segments were classified into 5 groups based on the load range applied (20–30%, 30–40%, 40–50%, 50–60% and 60–70%). For each group we calculated the probability of survival (no fracture) after 5, 100, 500, 1000 and 5000 load cycles (Table 4.1). These data were transformed into data points by assigning the average survival probability after 5, 100, 500, 1000 and 5000 load cycles to all specimens that had been loaded in a specific load range (Table 4.2). To assess the exponents for load magnitude and number of load cycles in the calculation of CLBL, cumulative loading was defined as:

$$\mathrm{Load}_{\mathrm{cum}} = {\mathrm{N}_{\mathrm{cycles}}}^{\mathrm{N}_{\mathrm{exp}}} \cdot \mathrm{Load}^{\mathrm{F}_{\mathrm{exp}}} \qquad \textit{Equation 4.3}$$

in which load magnitude is weighted with an unknown exponent ($\mathrm{F}_{\mathrm{exp}}$), and multiplied by the number of load cycles which is also weighted with an unknown exponent ($\mathrm{N}_{\mathrm{exp}}$). Since this load is hypothesized to be associated with the probability of survival, a linear relation between cumulative load and survival probability was assumed, so that survival probability can be expressed as:

$$\text{Survival probability} = \text{intercept} - \text{slope} \cdot ({\mathrm{N}_{\mathrm{cycles}}}^{\mathrm{N}_{\mathrm{exp}}} \cdot \mathrm{Load}^{\mathrm{F}_{\mathrm{exp}}}) \qquad \textit{Equation 4.4}$$

Table 4.1 | Probability of a motion segment to survive without compression fracture depending on the relative load and the number of load cycles applied. The table is adjusted from Figure 16 of the original paper (Brinckmann et al., 1988). Note that this original figure shows probability of fatigue fractures whereas here we report survival probability.

Relative loads	Load Cycles				
	10	100	500	1000	5000
60-70% (n=11)	91	37	9	9	0
50-60% (n=13)	100	61	38	15	8
40-50% (n=21)	100	64	45	45	32
30-40% (n=11)	100	100	82	82	73
20-30% (n=12)	100	100	100	100	92

Table 4.2 | Data points obtained from the original data. The average survival probability after 5, 100, 500, 1000 and 5000 load cycles was assigned to all specimens that had been loaded in a specific load range. For example, for the rightmost two lowest cells of Table 4.1, 12 data points were created in which a mean load range of 25 (20–30%) resulted in 92% survival after 5000 load cycles and 11 data points were created in which a mean load range of 35 resulted in a 73% survival probability after 5000 load cycles. This conversion led to a total of 340 data points.

Average Load	Load Cycles	Survival Probability	Number of data points (n=340)
25	10	100	12
35	10	100	11
45	10	100	21
55	10	100	13
65	10	91	11
25	100	100	12
35	100	100	11
45	100	64	21
55	100	61	13
65	100	37	11
25	500	100	12
35	500	82	11
45	500	45	21
55	500	38	13
65	500	9	11
25	1000	100	12
35	1000	82	11
45	1000	45	21
55	1000	15	13
65	1000	9	11
25	5000	92	12
35	5000	73	11
45	5000	32	21
55	5000	8	13
65	5000	0	11

Three optimization procedures were performed using simulated annealing (Goffe et al., 1994) in Matlab (The Mathworks, Natick MA, USA), to calculate intercept, slope and exponent(s) that resulted in the best fit through the data points by minimizing the average absolute error of all data. With regard to the exponents, in the first optimization, F_{exp} was assessed while assuming that N_{exp} is 1. In the second optimization, N_{exp} was assessed while assuming that F_{exp} is 1. In the last optimization, both F_{exp} and N_{exp} were assessed. For the three procedures, the abovementioned exponents as well as the intercept and slope of the best fit were calculated. Average absolute errors were calculated, while the goodness of fit of all fits was assessed using Akaike's Information Criterion (AIC; Akaike, 1974). We used this criterion since it takes into account the higher number of degrees of freedom in the third fit compared to the first two fits. The fit with the smallest AIC is considered the fit with the lowest loss of information. To test for the robustness of the current results, a leave-one-out cross-validation (LOOCV) was performed. This was done by leaving one cluster of data points out of the original sample. Subsequently, exponents were calculated by the abovementioned optimization procedures, based on the remaining sample. These exponents were validated using the 'left out cluster' by calculating the difference in predicted survival probability and actual survival probability. This was repeated such that each cluster of data-points was left out once, while differences between actual and predicted survival probability were averaged over all repetitions.

RESULTS

The probability of survival of the 5 groups of specimen exposed to different load ranges (Table 4.1) was transferred into 340 data points (Table 4.2; Figure 4.1). The first optimization resulted in a F_{exp} of 1.7 (AIC = 1048.64, averaged error = 22.33, LOOCV = 25.00):

$$\text{Survival probability} = 85.5 - 1.4\cdot10^{-5}\cdot(N_{cycles}\cdot \text{Load}^{1.7}) \qquad \textit{Equation 4.5}$$

The second optimization resulted in N_{exp} of 0.2 (AIC = 981.32, averaged error = 15.01, LOOCV = 18.28):

$$\text{Survival probability} = 100.0 - 2.6\cdot10^{-1}\cdot(N_{cycles}{}^{0.2}\cdot \text{Load}) \qquad \textit{Equation 4.6}$$

The third optimization resulted in F_{exp} and N_{exp} of 2.0 and 0.2 (AIC = 948.02, averaged error = 11.53, LOOCV = 14.06):

$$\text{Survival probability} = 100.0 - 5.1\cdot10^{-3}\cdot(N_{cycles}{}^{0.2}\cdot \text{Load}^{2.0}) \qquad \textit{Equation 4.7}$$

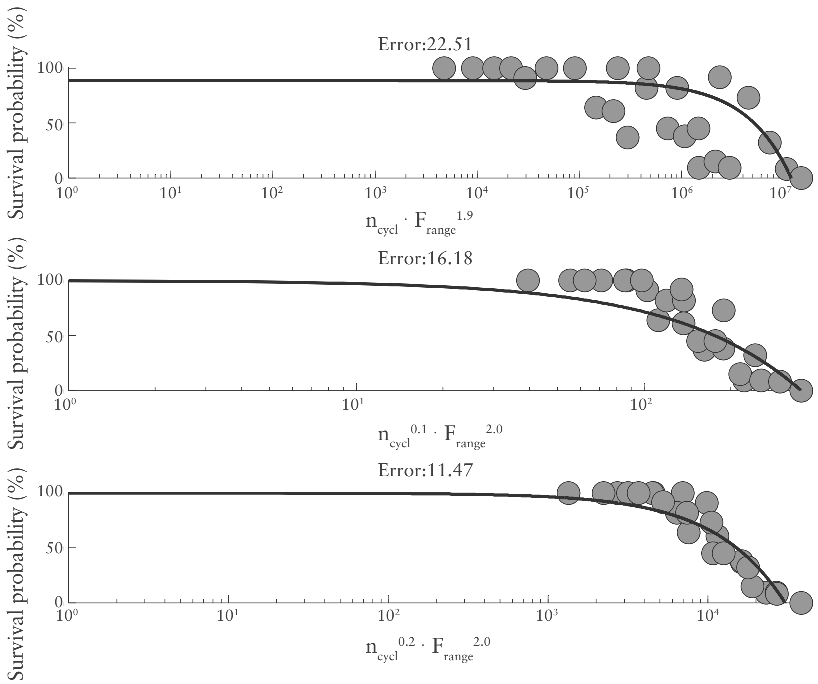

Figure 4.1 | Survival probability plotted against cumulative low-back load. Both the data points (dots) and the optimal fit of the function trough these data points (solid line) are shown. Furthermore, root-mean-square errors in comparison to the data points, averaged over data points are shown. An optimal fit through all data points assessing the relative weighting of load magnitude (upper panel), an optimal fit assessing the relative weighting of number of load cycles (middle panel) and an optimal fit assessing the relative weighting of both load magnitude and number of load cycles (lower panel) are shown. Note that each dot represents at least 11 and at most 22 data points. Dots are scaled to the number of data points they represent; the smallest dot represents 11 data points whereas the largest dot represents 21 data points.

DISCUSSION

The aim of the present study was to determine appropriate exponents for weighting of low-back load magnitude and the number of load cycles in CLBL calculations, based on in vitro compression data. Results show that weighting of load magnitude and number of load cycles with exponents of approximately 2 and 0.2 respectively can be suitable for CLBL estimates:

$$\text{Load}_{cum} = N_{cycles}{}^{0.2}\cdot \text{Load}^{2} \qquad \textit{Equation 4.8}$$

This can be rewritten to:

$$\text{Load}_{\text{cum}} = (\text{N}_{\text{cycles}} \cdot \text{Load}^{2/0.2})^{0.2} \qquad \textit{Equation 4.9}$$

which allows, due to the fact that N_{cycles} is now linear within brackets, summation of multiple (n) load levels, thereby making the equation applicable to work situations with multiple tasks of different load magnitudes:

$$\text{Load}_{\text{cum}} = \Big(\sum_{i=1}^{n} \text{N}_{\text{cycles}}(i) \cdot F(i)^{10}\Big)^{0.2} \quad \text{for } i=1,2,\ldots n \qquad \textit{Equation 4.10}$$

And in fact, this equation can be simplified to:

$$\text{Load}_{\text{cum}} = \Big(\sum_{j=1}^{k} F(j)^{10}\Big)^{0.2} \quad \text{for } j=1,2,\ldots k \qquad \textit{Equation 4.11}$$

where k is the total number of load cycles, that can be summed irrespective of the question whether or not some of them have equal load levels.

Both errors and AIC show a substantial reduction of the information loss in the third fit compared to the first two fits. These results suggest a substantial improvement of the estimation of CLBL when, in addition to exponentially weighting of load magnitude, the number of load cycles is exponentially weighted as well. It should also be noted that a weighting of load magnitude alone resulted in an intercept that deviated from the expected 100% survival at zero cumulative loading. Furthermore, as the LOOCV provides values that are only slightly higher than the calculated averaged absolute errors, it can be concluded that the present findings are robust.

These findings might have important implications for the calculation of CLBL. Concerning the earlier example about the risk of 15 times a 2000 N load compared to 20 load cycles of 1500 N, CLBL of these protocols will lead to $15^{0.2} \cdot 2000^2 = 6.87 \cdot 10^6$ and $20^{0.2} \cdot 1500^2 = 4.20 \cdot 10^6$ loads, a substantial difference in CLBL between the two protocols. This contrasts with the commonly used linear weighting of load magnitude and number of load cycles, which would result in equal CLBL estimates for these two protocols. Moreover, the method we propose might also be applicable to more realistic work situations. For example, combining the two abovementioned work situations might, according to Equations 4.10 lead to a CLBL of $(15 \cdot 2000^{10} + 20 \cdot 1500^{10})^{0.2} = 6.97 \cdot 10^6$. Not taking weighting of the number of load cycles into account can lead to large overestimations in the calculation of CLBL, as when only using the squared weighting of the load magnitude, this would yield a total CLBL of $15 \cdot 2000^2 + 20 \cdot 1500^2 = 1.05 \cdot 10^8$, a more than fifteen-fold higher estimate of the CLBL compared to our method.

It should be noted here that our analyses were performed, based on compression loads that were normalized to the ultimate strength of a specimen rather than on absolute data (N). Application of the current method to comparisons between (groups of) workers, concerning cumulative low-back loads or estimations of survival probability (based on Equation 4.7 and the average ultimate strength of 5.24 kN this would for abovementioned example yields: $100 - 5.1 \cdot 10^{-3}(15 \cdot (100 \cdot 2000/5240)^{10} + 20 \cdot (100 \cdot 1500/5240)^{10})^{0.2} = 87\%$ survival probability), would thus preferably take the capacity of the workers into account, for instance through prediction of individual ultimate strength (Brinckmann et al., 1988) as can for example be predicted in vivo using ultrasound (Nicholson & Alkalay, 2007).

The squared weighting of load magnitude in our best fitting model is consistent with the values proposed by Seidler et al. (2009; 2001), but not consistent with more conventional, linear weighting (e.g., Kumar, 1990; Marras et al., 2010; Norman et al., 1998) or a fifth order polynomial calculated by Parkinson and Callaghan (2007). In the latter study only material of healthy porcines was used instead of humans. Furthermore, no resulting errors were reported, making the results hard to compare to the present data. Besides, in our study, adding a weighting of number of loads turned out to lead to substantial improvement of the CLBL estimation.

It should also be noted that specimens in this study were exposed to one specific cycle time and load magnitude and that the number of load cycles was limited to a maximum of 5000. Whether the present results hold for other exposures (e.g. long sustained exposure or multiple different cyclic exposures), remains to be investigated. Furthermore, specimens in the current study were exposed to compression loads only, while in real life situations loading patterns are more complex and often occur in non-neutral postures (Kingma et al., 2006; Marras et al., 2010). However, compression loading is widely accepted as an important component of low-back loading (Potvin, 1997; van Dieën et al., 1999; Waters et al., 1993).

The choice to use average absolute errors rather than other possible calculations of errors (e.g., RMS errors) is an arbitrary one. However, when re-running our analysis using RMS instead of absolute average errors, we found a similar pattern of errors over optimizations and exponents that only slightly deviated for optimization 1. A limitation of the present study is that we used a multiplicative exponential model only. While we showed that this multiplicative model leads to robust outcomes, other functions may also result in acceptable fits. Furthermore, analyses were performed on data obtained from in-vitro measurements. Therefore, results might not generalize to in vivo situations. Cadaver material, certainly when not tested in a fluid bath does recover poorly from loads and biological repair is definitely absent. So the present study only applies to short term fatigue fracture loading (van der Veen et al., 2005). Roughly, repair of micro-fractures can be estimated to take several weeks. Results of the present study are therefore valid only within this interval.

CONCLUSIONS

It can be concluded that weighting compression forces and number of load cycles with exponents of approximately 2 and 0.2, respectively, provide a suitable metric of cumulative compression loading of the spine for conditions tested in this study. These findings might be relevant for future studies on LBP etiology.

Chapter 5

Effects of the data sampling strategy on bias and power in epidemiologic studies of low-back pain – a bootstrapping approach

P. Coenen
S.E. Mathiassen
I. Kingma
C.R. Boot
P.M. Bongers
J.H. van Dieën

Submitted

ABSTRACT

Studies of work-related low-back pain (LBP) often classify workers into exposure groups for which e.g., lifting or awkward trunk postures are estimated from measurements on a sub-population. The present study investigated combined influences of the sizes of the total study population and the sub-sample on exposure-outcome associations.

At baseline, lifting, trunk flexion, and trunk rotation was observed for 371 of 1131 workers in 19 task groups. Self-reported LBP (dichotomous) was obtained from all workers during three years of follow-up. All three exposures were associated with LBP ($p<0.01$) according to logistic regression.

All possible combinations of $n=10,20,30$ workers per task group and $k=1,2,3,5,10,15,20$ workers being observed were investigated using bootstrapping. The OR and its p-value was determined for each of 10,000 virtual studies at each combination of n and k, and the average OR and the statistical power ($p<0.05$ and $p<0.01$) across the 10,000 studies were assessed.

For lifts and flexed trunk, studies including $n\geq20$ workers in each task group, and $k\geq5$ observed, led to an almost unbiased OR and a power >0.80 (p-level 0.05). A similar performance required $n\geq30$ workers for rotated trunk. Small numbers, k, of observed workers resulted in biased OR, while power was, in general, more sensitive to the total number, n, of workers than to the number, k, of observed workers.

In a group-based exposure assessment strategy, statistical performance may be sufficient if the overall size of the groups is reasonably large, even if exposure is estimated of few workers per group.

INTRODUCTION

In the past decades, numerous epidemiological studies have been conducted on occupational physical exposure risk factors for low-back pain (LBP). Among other factors, exposures such as heavy lifting, trunk flexion, and trunk rotation have been suggested to be risk factors for LBP (Griffith et al., 2012; Lötters et al., 2003). However, the literature on occupational physical risk factors of LBP is not consistent (Bakker et al., 2009; Kwon et al., 2011), one possible reason being that the strategies for assessing physical exposures differ between studies (David, 2005; Punnett & Wegman, 2004).

Several studies on occupational physical risk factors for LBP have adopted a group-based exposure assessment strategy (Ariëns et al., 2001; Burdorf & Jansen, 2006; Hoogendoorn et al., 2000a). Workers are then classified into groups with an expected contrast in exposure, typically based on their job or tasks. The exposure variable(s) of interest is measured only in a sub-sample of workers within each group, and the resulting average exposure of the measured workers is assigned to all workers in the group. Exposure-outcome relationships are then determined using these exposure estimates together with individual data on health outcomes (i.e., LBP) from all subjects in the study population. This exposure assessment strategy is based on the assumption that workers within the same group have similar exposures, i.e. that the groups are homogeneous with respect to exposure, and that exposure variability between groups is comparatively large, so that the exposure contrast between groups will be substantial (Kromhout & Heederik, 1995; Mathiassen et al., 2005).

The effect of the number and allocation of exposure measurements on the statistical properties of a group mean exposure estimate is relatively well documented (Hoozemans et al., 2001; Liv et al., 2010; Mathiassen et al., 2002; Mathiassen et al., 2003a). However, the influence of measurement strategies on the strength and statistics of exposure-outcome associations in logistic regression has, to the best of our knowledge not been thoroughly investigated. A theoretical framework has been presented on the issue of bias and precision in linear regression of continuous outcomes on (continuous) exposure measured with random uncertainty (Tielemans et al., 1998), and even logistic regression has been discussed in this context (Reeves et al., 1998). However, the case of estimating exposure in group(s) from observations of a sub-population while using personal outcome data has not been addressed in any of these studies. Also empirical data to complement theoretical findings have not been presented. Therefore, the present study aimed to assess the combined effect of the sample size of the total population and that of the sample on which exposure is actually observed on exposure-outcome associations in a study of occupational physical exposures and LBP.

STUDY POPULATION AND METHODS

Population

The present study is based on data from the Study on Musculoskeletal disorders, Absenteeism and Health (SMASH). As described in detail previously (Coenen et al., 2013b; Hoogendoorn et al., 2000a), this prospective cohort study recruited workers from 34 companies in the Netherlands. At baseline, 1989 of 2048 invited workers agreed to participate, and questionnaire data on personal factors and work characteristics were obtained from 1802(91%) of these workers. These 1802 workers were classified by experts into 23 task groups, based on their expected physical work load. Within each task group, work was recorded on video from a random sample of roughly one fourth of the workers. After excluding workers dropping out after the baseline measurements, the parent data set for the current study included 1131 workers from those 19 task groups that contained more than 5 observed workers. Video based observation data were available from, in total, 371 workers (Table 5.1).

Exposure and outcome for the parent data set

For each of the 371 workers recorded on video, four recordings were obtained at randomly chosen times during the course of a single work day. Recordings lasted 5-15 minutes each, depending on the variability of the worker's tasks. Recordings were analyzed post-hoc using a structured protocol for assessing three physical exposures, which were shown to be significantly associated with LBP in the same population (Hoogendoorn et al., 2000a); i.e., the number of lifts during an eight hour work week, the percentage of working time with the trunk flexed (defined as >30° trunk flexion), and the percentage of working time with the trunk rotated (defined as >30° trunk rotation). The mean exposure of the observed workers in each of the 19 task groups was assigned to all workers classified into that group. In order to evaluate the task group classification, between-group contrasts for each of the three exposure risk factors were calculated, using:

$$\text{Contrast} = \frac{\text{MSE}_b}{(\text{MSE}_b + s_w)} \qquad \text{Equation 5.1}$$

In which MSE_b is the mean squared error between task groups and s_w is the variability between workers within groups (Kromhout & Heederik, 1995; Mathiassen et al., 2005).

Self-reported LBP was assessed for all 1131 workers once a year for three years after the baseline measurement using a Dutch version of the Nordic Questionnaire (Kuorinka et al., 1987). A case of LBP was registered when a worker reported regular or prolonged LBP during at least one of the three years of follow-up, regardless of baseline status.

Logistic regression analyses using the three exposure variables as continuous independent variables (in which the number of lifts was divided by 100 and percentages of time in flexed or rotated postures were divided by 10) and LBP as the dichotomous dependent variable were executed. Results showed both the number of lifts (per 100 lifts; OR: 1.06 (95%CI: 1.03-1.09), p<0.01), the time working with the trunk flexed (per 10%; OR: 1.31 (95%CI: 1.12-1.52), p<0.01), and the time working with the trunk rotated (per 10%; OR: 1.43 (95%CI: 1.06-1.93), p<0.01), to be significantly associated with LBP in the parent data set.

Simulated sampling strategies

For all 21 possible combinations of n=10,20,30 workers in total per task group and k=1,2,3,5,10,15,20 workers being observed, exposure-outcome associations were assessed using a non-parametric bootstrap simulation procedure as follows (Efron & Tibshirani, 1986; Hoozemans et al., 2001; Liv et al., 2010; Paquet et al., 2005). Within each task group of the parent data set, workers were identified as "observed" and "non-observed" depending on whether exposure data were available or not. For each combination of n and k, k workers in each task group were drawn with replacement from the group of observed workers, and n workers were drawn with replacement from all workers (observed and non-observed combined) in the same task group. This led to a virtual study including n workers in total and k observed workers from each task group. For each virtual study, the three mean exposures (number of lifts, trunk flexion, and trunk rotation) of the k observed workers within each task group were then assigned to all n workers in that particular task group, while the individual LBP status was used as the outcome for each of the n workers. For each virtual study constructed this way, the ORs (with p-levels) for the three associations between each of the exposure variables and LBP were assessed using logistic regression analysis as explained above for the parent data set. For each of the 21 possible combinations of n and k, 10,000 virtual studies were constructed using this procedure. Four measures for each investigated exposure assessment strategy were obtained on the basis of the 10,000 virtual study results, i.e. 1) a pooled estimate of the standard deviation (SD) of the mean exposure estimate within a task group, obtained by first calculating the mean variance between subjects, VAR_{BS}, across the 10,000 replicates of that variance for each specific task group, and then pooling these 19 variances into the average SD of a mean exposure estimate according to the formula:

$$\text{Pooled SD} = \sqrt{\frac{\text{mean}(VAR_{\text{BS}})}{\text{k}}} \qquad \text{Equation 5.2}$$

2) The SD across the 10,000 studies of the LBP prevalence in the population, 3) the mean OR across the 10,000 studies, and 4) the power in each exposure assessment strategy to detect a significant OR at levels p<0.05 and p<0.01, i.e. the proportions of the 10,000 studies resulting in an OR with the mentioned significances. All calculations were performed using customary scripts in Matlab (MATLAB 7.7.0, The MathWorks Inc., Natick, MA, 2000). Logistic regression analyses were implemented using the Matlab statistical toolbox.

Table 5.1 | Parent data set. In the upper panel, the total number of workers (N) and the number of workers observed (K) are shown for each task group, together with LBP prevalence among all workers, and group mean exposures and standard deviations for the three investigated physical exposures: number of lifts at work per week, percent time spent with the trunk flexed more than 30°, and percent time spent with the trunk rotated more than 30°.

In the lower part of the table, pooled descriptive statistics (gender, length, weight, age, working hours per week, years of employment at the current job and proportion of workers with LBP at baseline) are shown for all workers (N) and for those observed (K).

	Workers	LBP	Observed	Lifts		Flexion		Rotation	
Description task groups	(N)		(K)	Mean	StDev	Mean	StDev	Mean	StDev
Mainly sitting work									
Sitting with varying postures	133	39%	61	23.2	83.2	6.3	9.0	12.3	29.2
Sitting with little varying postures (computer work)	57	35%	16	13.4	51.9	7.4	10.5	1.1	2.2
Sitting with little varying postures, in awkward postures (no computer work)	31	68%	11	1.1	3.5	3.6	8.9	2.9	2.7
Sitting with little varying postures, with repetitive movements	95	42%	31	334.3	933.7	2.4	3.1	2.4	3.1
Mainly standing work									
Standing with varying postures (including walking) without external forces	26	58%	9	8.0	18.8	4.1	3.5	2.2	2.8
Standing with varying postures and small external forces	69	38%	23	658.9	781.8	7.2	5.4	2.4	2.9
Standing with varying postures and moderate external forces	87	44%	28	438.1	521.5	10.0	8.9	5.6	4.4
Standing with varying postures and large external forces	65	40%	20	299.5	283.6	11.5	6.5	6.2	5.5
Standing with varying, awkward postures and moderate external forces	66	50%	22	544.4	620.0	13.7	9.1	6.1	4.2
Awkward postures (mainly static exposure)									
Standing in static awkward posture without external forces	42	48%	15	133.6	177.9	8.4	7.3	4.3	3.9
Standing in static awkward posture with small external forces	70	39%	24	194.7	277.2	10.3	6.7	6.7	5.8
Mainly static back exposures by alternating awkward postures	28	61%	11	814.8	1167.3	37.6	30.7	12.1	6.8
Alternating exposures (standing, walking and/or sitting)									
Alternating standing, walking and/or sitting without external forces	167	40%	29	6.4	32.1	5.7	6.1	2.4	3.4
Alternating standing, walking and/or sitting with small external forces	36	50%	13	82.2	71.1	8.9	5.9	4.2	5.9
Alternating standing, walking and/or sitting with moderate external forces	52	42%	15	312.9	179.8	22.0	12.0	4.7	5.8
Alternating standing, walking and/or sitting with large external forces	21	86%	8	2904.0	1483.9	42.5	15.7	19.8	8.1
Alternating standing and walking in static awkward postures, external forces	27	44%	17	379.2	433.4	19.2	11.8	6.8	5.9
Alternating standing and walking in postures, moderate external forces	36	56%	9	577.2	275.8	12.8	6.7	7.4	3.9
Combined functions (as a result of changes in tasks)									
Combined exposures	23	30%	9	252.2	297.0	8.4	7.5	2.4	2.3
Total	1131	44%	371						
Descriptive variables									
Number of workers	1131		371						
Number of males	699(69%)		219(68%)						
Number of females	307(31%)		104(32%)						
Stature (cm)	175.9(9.6)		175.7(9.4)						
Weight (kg)	75.9(13.6)		74.9(12.3)						
Age (years)	35.5(8.8)		35.7(8.8)						
Working hours per week	37.2(6.9)		36.7(6.9)						
Years of employment	9.7(7.5)		9.4(7.3)						
Baseline LBP (number of workers)	366(38%)		125(40%)						

Continuation of table 5.1

RESULTS

Exposure contrasts between groups were 0.55, 0.48 and 0.23 for the number of lifts, time in flexed trunk posture and time in rotated trunk posture, respectively. While task groups did, indeed, differ in mean exposure (Table 5.1), some were very heterogeneous in terms of the workers differing substantially in exposure.

For all three exposure variables, the pooled SD of the group mean exposure decreased as the number of workers, k, for which exposure was actually observed increased (Figure 5.1). This confirmed that more data lead to more precise exposure estimates. Obviously, this effect did not depend on the total number of workers, n, per task group. The SD of the prevalence of LBP in the study population decreased with an increasing total number of workers, n, included in each task group (Figure 5.2), and obviously this effect did not depend on k. The average OR of the association between exposure and LBP increased with larger k (Figure 5.3), while it was affected only little by the total number of workers, n.

Figure 5.4 shows that power increased with both n and k. The effect of the total number of workers, n, on power was stronger than that of the number of observed workers, k. However, the magnitude of these effects differed between risk factors. For number of lifts and time with flexed trunk, a power of 0.80 to detect a significant ($p<0.05$) OR was obtained when at least n=20 workers were included per task group, and the number of actually observed workers in each task group (k) was at least 5. For time working with the trunk rotated, at least n=30 workers per task group were needed to obtain the same power. At the more strict requirement of $p<0.01$, a power of 0.80 was obtained only when the population included at least n=30 workers per task group for lifts and flexed trunk, while this level of power could not be reached at all for the risk factor time working in a trunk rotated posture.

DISCUSSION

The present study dealt with the common group-based assessment strategy in musculoskeletal epidemiology of measuring exposure to risk factors in a sub-population of workers. Mean exposure estimates are then assigned to all workers having similar tasks or jobs, while information on outcomes is available from each individual worker in the total study population. Our study suggests that the probability of finding significant exposure-outcome associations depends more on the total number of workers included in each task group than the number of workers for whom exposure is actually observed. In our setting comprising 19 task groups intended to represent the general working population, studies including at least 30 workers in each group and basing the task group exposures on at least 5 observed workers were sufficient to secure a reasonable power and an almost unbiased estimate of the odds ratio. However, the exact numbers of subject to establish a certain statistical performance differed between the three investigated exposure risk factors (Figures 5.3 and 5.4). Our results may have important implications for future epidemiological studies, since they suggest that a limited research budget would be more efficiently used by collecting outcome data from “many” subjects than by spending extensive efforts on exposure observations, which are often expensive (Trask et al., 2012). As an illustration, reading from Figure 5.4, a statistical power around 0.80 ($p<0.05$) can be reached either by a study design comprising 20 workers per task group and only one is actually observed and by a study including 10 workers per task group, and 10 need to be observed. Thus, the “large” study requires outcome data to be collected from 380 workers, but exposure only from 19, while the “small” study is based on outcome data from only 190 workers, but exposure data from all 190. While the budgets of these two alternatives depend on the unit cost of obtaining exposure and outcome information, it seems likely that the “large” study is cheaper to realize. Notably, while these two sampling strategies have comparable abilities to detect a significant association between exposure and LBP, the former will, however, result in a more biased OR (Figure 5.3).

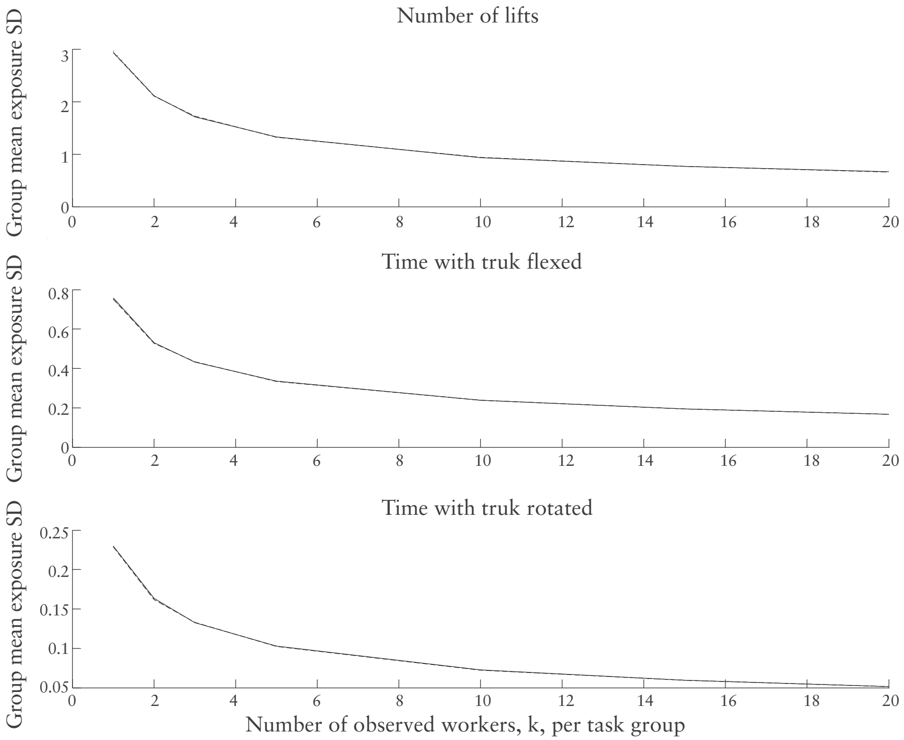

Figure 5.1 | Pooled estimate of the standard deviation (SD) of the group mean exposure in a task group for each of the 21 combinations of n (different lines) and k (x-axis). SD is presented for the exposure variables: number of lifts (upper panel), time with the trunk flexed (middle panel) and time with a rotated trunk (lower panel). Note that the individual curves for different n-values in each panel overlap completely.

As confirmed by our results, more precise (i.e. more certain) group mean exposure estimates will be obtained when data are collected from more workers. Several studies (Allread et al., 2000; Hoozemans et al., 2001; Mathiassen et al., 2005) have shown that the exposure estimate improves still less when still more workers are included in the estimate. Thus, beyond a certain number of observed workers, it may not be warranted to invest more resources in observing even more workers. Similarly, the estimate of the outcome (i.e., the LBP prevalence) will become more precise when more workers are included in a study, and may reach a sufficient precision at a particular number of workers, beyond which further investments may not be justified.

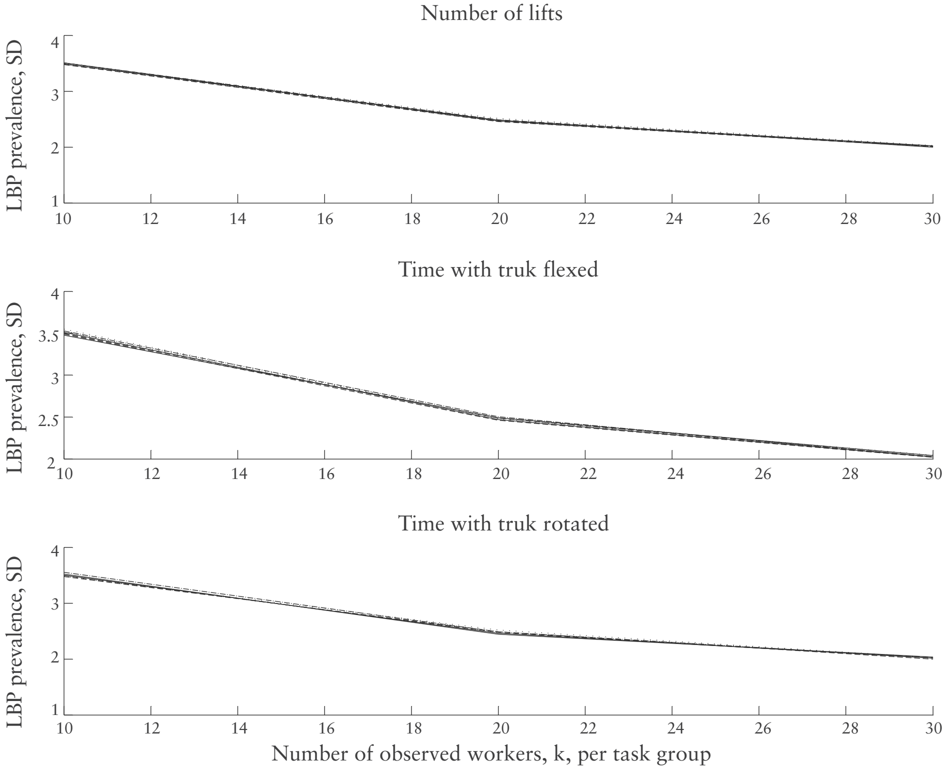

Figure 5.2 | Standard deviation (SD) of the outcome (i.e. LBP prevalence in the entire data set) across the 10,000 replicates for each of the 21 investigated combinations of n (x-axis) and k (different lines). Standard deviations are presented for the exposure variables: number of lifts (upper panel), time with the trunk flexed (middle panel) and time with a rotated trunk (lower panel). Note that the individual curves for different k-values in each panel overlap completely.

The decrease in average ORs with lower numbers of k, i.e. an attenuation of the OR towards 1, is probably a result of increased uncertainty in the estimate of task group exposures, since the OR was, only weakly influenced by the overall number of workers, n, in each task group. Attenuation of exposure-outcome regression coefficients due to uncertainty in the exposure estimates also occurs in simple linear regression of two continuous variables (Tielemans et al., 1998)Non-U.S. , as well as in logistic regression (Reeves et al., 1998), even though a group-based exposure assessment strategy is generally regarded to be an effective measure to avoid biased regression coefficients, in particular in linear regression (25). Our results showed that the bias was, however, not very strong, and only weakly influenced by the overall number of workers in each task group, n.

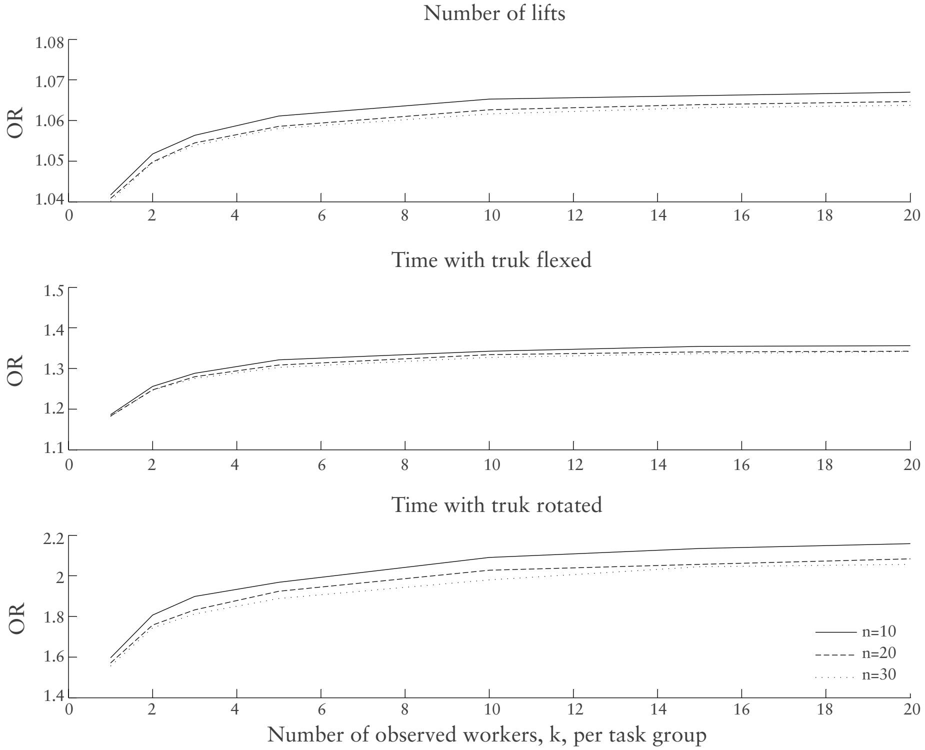

Figure 5.3 | Average odds ratios (OR) for the association between exposure and LBP across the 10,000 replicates for each of the 21 investigated combinations of n (different lines) and k (x-axis). Average ORs are presented for the exposure variables: number of lifts (upper panel), time with the trunk flexed (middle panel) and time with a rotated trunk (lower panel).

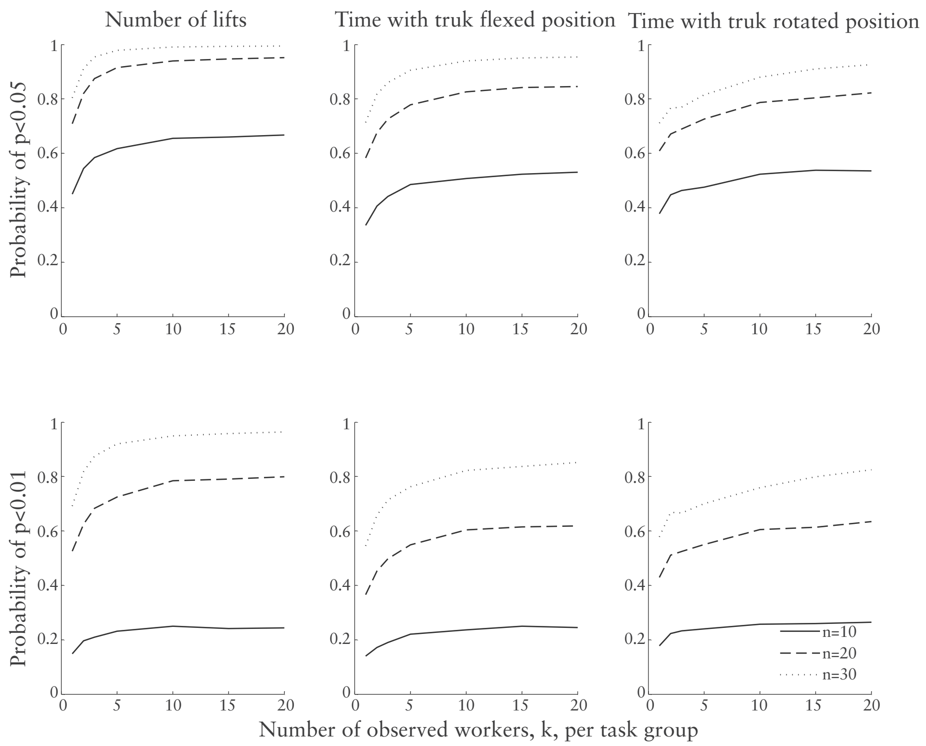

Figure 5.4 | Statistical power, i.e. the probability of obtaining a significant OR for the association between exposure and outcome, for all 21 investigated combinations of n (different lines in each panel) and k (x-axis). Upper and lower panels: significance levels $p<0.05$ and $p<0.01$, respectively. Probabilities of obtaining a significant OR are shown for the exposure variables: number of lifts (left panels), time with the trunk flexed posture (middle panels) and time with a rotated trunk (right panels).

Earlier occupational studies as well as statistical textbooks present equations to calculate the power of a study protocol to obtain statistically significant results, as a function of sample sizes and variability (e.g., of exposures) in the study population (Mathiassen et al., 2002; Mathiassen et al., 2003b; Twisk, 2003). While this literature discusses comparatively simple study designs, the present study confirms the general effect of more data improving power for a more complex design. Our study also adds the observation that the size of k does have an effect on power, but that this effect is weaker than that of changing the total number of workers (Figure 5.4).

In the present study, the video recordings of each particular worker were collected at four randomly chosen occasions during the course of one single day. This may be considered a less efficient choice, since distributing these four occasions over several days would likely have resulted in a more certain exposure estimate for that worker, given that exposure probably varied between days within workers (Hoogendoorn et al., 2000a; Kwon et al., 2011; Liv et al., 2010; Paquet et al., 2005; Twisk, 2003). More certain estimates of the exposures of individual workers in a task group would even lead to a more certain mean exposure estimate for the task group as a whole. Thus, collecting exposure data over multiple days per worker could have led to slightly different conclusions. For example, it might have been necessary to observe less workers to obtain the same exposure-outcome associations as what is now obtained with, for instance, k=5 workers in each task group. However, since the uncertainty of the exposure estimates for individual workers is expected to be the result of random statistical processes, the general conclusions of our study would not change.

An expert classification of tasks (jobs) into groups, based on suspected physical workloads, may result in a grouping scheme that does not effectively capture exposure differences between workers in different tasks. Thus, as it appears even in our material, exposure variability between subjects may be considerable within several of the task groups (Table 5.1), and another categorization of some workers might have resulted in more homogeneous task groups. Task groups were carefully set up by the same trained observers who also collected the video recordings, based on their extensive experience of physical work load assessment in occupational settings. According to the exposure contrast values, classification was reasonably successful for the two variables number of lifts and time in flexed postures. For time working in rotated trunk posture, the contrast was lower, mainly due to task group 1 being very heterogeneous (Table 5.1). The latter is a possible explanation that power was generally less for exposure-outcome relationships based on this risk factor (Figure 5.4). Whether a different grouping scheme, with less or more task groups, possibly defined using other criteria, could have been more effective in disclosing exposure-outcome associations for LBP is an open question. Therefore, studies employing other grouping schemes might reach different results as to the statistical performance of sampling strategies than we did. However, we believe that the trade-off between total study size and number of observed workers would be a consistent finding. Moreover, over results suggest that classifications in future studies of tasks and jobs according to expected exposures could benefit from more comprehensive a priori knowledge. As an example, a pilot study in which observational data of a limited amount of workers is collected and analyzed to identify an optimal classification a priori to the full study could probably lead to a more informed and more effective classification.

The present study addressed only three exposure variables (i.e., lifting, trunk flexion and trunk rotation). In our parent data set, these three exposure variables correlated only weakly, with correlation coefficients of 0.34, 0.09 and 0.09, for lifting *vs.* flexion, lifting *vs.* rotation, and flexion *vs.* rotation, respectively. Therefore, it seems reasonable both to assess the effect of these three exposures on LBP independently of each other and to assume that our general results may apply even to other variables describing trunk exposure, i.e. that the results show a fair external validity.

The present simulations were constructed to include the same number of workers from each task group in a balanced study design. This may have affected exposure-outcome associations, as compared to the more usual situation in epidemiologic studies (and in our parent data set) of groups being of different sizes. As a general rule, the statistical power of a balanced study design will be larger than that of an unbalanced design with the same total number of workers, and so the exposure-outcome associations of our simulated study designs are probably stronger and more precise than those in comparable unbalanced designs of the same total magnitude.

In the current bootstrapping procedure, samples of workers were drawn with replacement from each task group. Therefore, it was possible to “oversample” workers (i.e. obtaining a virtual sample of workers that was larger than the number of unique workers available in the group. Oversampling by more than 100% (i.e., sampling at least twice as many workers as available in the parent data) occurred in 4 out of 19 task groups when selecting k=20 workers for the exposure estimates, while it did not occur for values of k between 1 and 15, and not either in any case of sampling the n workers providing LBP data. We have not been able to identify any discussion in the bootstrapping literature on the acceptability and limits of oversampling, let alone its possible effects on the resulting data distributions (Davison & Hinkley, 1997; Efron & Tibshirani, 1986). However, it is reasonable to assume that effects of oversampling are more prominent if the parent data is small and/or irregularly distributed. We restricted our parent data set to task groups represented by at least 5 observed workers and 21 workers in total (Table 5.1) in order to get a fair representation of workers in the task group, and thus, among other benefits, reduce the possible effect of oversampling. Since results from the sampling strategies containing oversampled exposure data are in line with results from strategies where no oversampling occurred (Figure 5.4), we believe that oversampling did not have serious effects in our study.

In conclusion, the statistical power of an exposure-outcome study design using group-based exposure estimation depended more on the total number of workers included in the study (with personalized outcome data) than on the size of the population on which exposures were actually determined. When, however, exposure was observed on very few workers, the odds ratio of the exposure-outcome relationship was downward biased irrespective of the total population size. Our findings thus suggest that (costly) exposure observations are necessary only on few workers, provided that the overall size of the study population is sufficiently large and everybody is followed up with respect to outcome. These results may contribute to a more informed use of resources in future epidemiological studies.

Chapter 6

Estimation of low-back moments from video analysis a validation study.

P. Coenen
I. Kingma
C.R. Boot
G.S. Faber
X. Xu
P.M. Bongers
J.H. van Dieën

Journal of Biomechanics, 2011,
44(13):2369-2375

ABSTRACT

This study aimed to develop, compare and validate two versions of a video analysis method for assessment of low-back moments during occupational lifting tasks since for epidemiological studies and ergonomic practice relatively cheap and easily applicable methods to assess low-back loads are needed. Ten healthy subjects participated in a protocol comprising 12 lifting conditions. Low-back moments were assessed using two variants of a video analysis method and a lab-based reference method. Repeated measures ANOVAs showed no overall differences in peak moments between the two versions of the video analysis method and the reference method. However, two conditions showed a minor overestimation of one of the video analysis method moments. Standard deviations were considerable suggesting that errors in the video analysis were random. Furthermore, there was a small underestimation of dynamic components and overestimation of the static components of the moments. Intra-class correlations coefficients for peak moments showed high correspondence (>0.85) of the video analyses with the reference method. It is concluded that, when a sufficient number of measurements can be taken, the video analysis method for assessment of low-back loads during lifting tasks provides valid estimates of low-back moments in ergonomic practice and epidemiological studies for lifts up to a moderate level of asymmetry.

INTRODUCTION

As low-back pain (LBP) in society is associated with high social suffering and costs (Lambeek et al., 2011), it is important to consider risk factors involved. Associations between physical risk factors and the occurrence of LBP have been reported extensively with lifting, twisting, bending and whole body vibrations being the most commonly reported ones (Lötters et al., 2003; Wai et al., 2010).

Although posture and force measurements and subsequent biomechanical analyses can provide valid and reliable estimates of back load during occupational handling (Kingma et al., 1996), such measurements are time and money consuming and can hardly be used outside the laboratory setting for epidemiological studies. Accordingly, research has focused on less costly (with respect to time and money) low-back load assessment methods, which can be brought into the work place easily. Direct observation combined with simple measurements (i.e. load distances) was shown to provide reasonable estimates of low-back loads during lifting, although systematic underestimation of loads occurred, possibly due to neglecting segment dynamics (van Dieën et al., 2010). Other efforts focused on video analysis methods (Chang et al., 2003; Hsiang et al., 1998; Sutherland et al., 2008; Xu et al., 2011) by assessing body orientations based on observations of selected key video frames. These methods provided acceptable kinematic accuracy (Chang et al., 2010; Neumann et al., 2001b; Xu et al., 2011). Furthermore, quasi-static biomechanical calculation using these kind of models showed small but significant errors in peak (Chang et al., 2003; Hsiang et al., 1998) and cumulative (Sutherland et al., 2008) lumbar compression forces. Although promising, these methods suffer from some shortcomings. Segment orientations were based on crude categorizations (Hsiang et al., 1998; Sutherland et al., 2008), segment dynamics were not taken into account (Sutherland et al., 2008) or only movements in the sagittal plane could be determined (Chang et al., 2003; Chang et al., 2010). Therefore, better posture matching strategies should be investigated. The aim of the present study was thus to develop, compare and validate (against a reference laboratory-based 3D inverse dynamics method) two versions of a video analysis method for estimation of mechanical back load (expressed in peak and mean moments) during occupational lifting tasks. With this method, we aim to overcome the abovementioned shortcomings by quasi-three-dimensional coding and online posture matching.

METHOD

Participants and procedure

After signing an informed consent, 10 healthy subjects (6 female and 4 male, age 23 (4) years, body mass 67 (7) kg and stature 1.76 (0.12) m) participated in a repeated measures experimental design approved by the ethics committee of the VU University, Amsterdam. Using a height adjustable shelf, subjects lifted a 15 kg box (0.57×0.38×0.37 m) in 12 different conditions: 2 horizontal initial positions of the box (at the front and at 0.57m from the front of the shelf), 3 vertical initial positions of the box (ground, hip and shoulder height) and 2 different types of lifting (symmetric and asymmetric lifting). For the symmetric lifting conditions, the subjects were asked to step towards the box, position the feet symmetrically, grab the box by its handles and lift it to chest height. For the asymmetric lifting conditions, subjects were asked to step towards the box, place the right foot in front of the left foot, grab the box by its handles and lift it with a 180° rotation to chest height. Lifting conditions were unconstrained, so no instructions were given with respect to lifting posture or exact foot placement, therefore, lifting conditions are assumed to resemble occupational tasks.

Reference measurement method

As a reference method, a dynamic three-dimensional linked segment model, described and validated by Kingma and colleagues (1996; 2010) was used. Kinematics of the box, lower arms (and hands), upper arms, trunk (and head) and pelvis were measured using cluster markers strapped to the body segments. Three-dimensional positions of the cluster markers were measured at a sample rate of 50 samples/s using the Optotrak motion capture system (Northern Digital Inc., Waterloo ON, Canada).

Anatomical landmarks were related to cluster markers using a probe with six markers (Cappozzo et al., 1995). Kinematic data were low-pass filtered using a cut-off frequency of 5 Hz. Segment masses, positions of the center of mass and inertia tensors were estimated using regression equations based on individual segment lengths and circumferences (Zatsiorsky, 2002).

Video measurement method

All lifting conditions were recorded with a Canon XM2 camera, while recordings were digitally captured and compressed into AVI format digital videos at a sample rate of 25 Hz. The camera was placed on a tripod which was situated perpendicular to the sagittal plane of the subject's initial lifting posture in the symmetrical lifting conditions. Videos and motion captured data were synchronized using an impulse light which was visible in all videos.

Video analyses were performed by a single observer (PC) using a video coding system with a graphical user interface (Figure 6.1) adjusted from an earlier method (Chang et al., 2003; Xu et al., 2011) using custom-made Matlab software (version 7.7.0). Initially, begin and end frames of the lifting condition were selected by replaying the video. The begin frame is the video frame of the initial lifting posture when the box gets clear from the shelf surface. The end frame is the frame in which the box was closest to the body. Additionally, two equally spaced frames between begin and end frames were selected, to obtain a total of four key frames (Xu et al., 2010b).

For the assessment of body kinematics during lifting, a quasi-three-dimensional manikin consisting of nine segments (right foot, lower leg and upper leg; pelvis, trunk/head, upper arms, forearms/hands) was fitted to the key frame pictures (Figure 6.1). This manikin allows for the following quasi-three-dimensional joint movements: ankle flexion/extension, knee flexion/extension, hip flexion/extension, trunk flexion/extension, trunk rotation, trunk lateral flexion, shoulder flexion/extension, shoulder abduction and elbow flexion/extension. Note that angles of the foot, ankle, knee and hip are required to correctly estimate upper body accelerations. Furthermore, the manikin can be scaled, translated and axially rotated for an optimal fit. Two variants for the composition of the manikin were assessed in the present study. The manikin could be fitted by adjusting the joint angles (video analysis method 1) or an initial guess of joint angles of all segments was calculated based on joint positions that were obtained by clicking on the video frame after scaling, translation and axial rotation of the manikin (video analysis method 2). In this algorithm, the above mentioned segment angles were calculated so that, based on the constrained segments lengths, a minimal difference in joint position compared to the joint position of the ankle, knee, hip, shoulders and hand that was clicked in the video frame was obtained. Subsequently, the observer could adjust joint angles to improve postural matching.

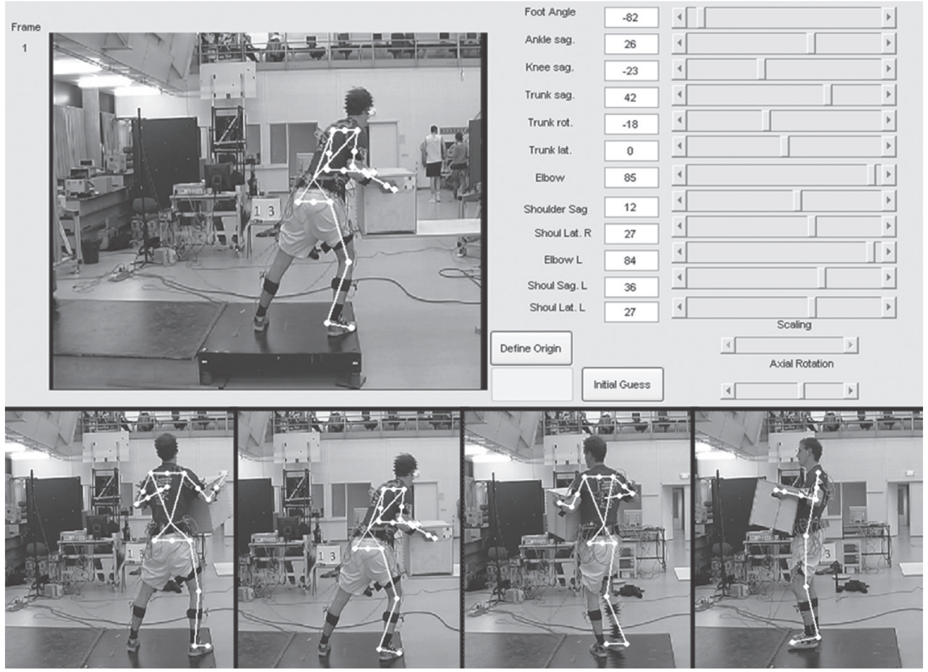
Frame 1
Foot Angle -82
Ankle sag. 26
Knee sag. -23
Trunk sag. 42
Trunk rot. -18
Trunk lat. 0
Elbow 85
Shoulder Sag 12
Shoul Lat. R 27
Elbow L 84
Shoul Sag. L 36
Shoul Lat. L 27

Figure 6.1 | Video analysis method. The upper part of the figure shows the graphical user interface in which a three-dimensional manikin is plotted online to a video key frame by axial rotation, scaling and translation and adjustment of segment angles. The lower part of the figure shows four key frames of an asymmetric lifting condition. These key frames show a representative sample of a video frame of an asymmetric lift as analyzed by the observer.

A cubic spline interpolation of the segment angles over the four key frames was applied to estimate segment angles over the entire lifting trajectory (Xu et al., 2010a). Segment mass, length, position of the center of mass and inertia tensor were estimated based on regression equations using total body mass and stature (Zatsiorsky, 2002). The relative flexion of the pelvis and trunk were estimated from upper body flexion and knee angle using regression equations (Anderson et al., 1985). Furthermore, the position of L5S1 was estimated at 19% of the length of the upper body segment (de Looze et al., 1992) and shoulder width was based on Dumas et al. (2007). The position and acceleration of all segments were constructed by linking all the segments from the right ankle through the hands/box.

Data analysis

To estimate total moments at L5S1 during all lifting conditions in all methods, a top-down calculation of the net moments at L5S1 was performed using external forces (mass and acceleration of the box), segment kinematics and anthropometrics using a global equation of motion (Hof, 1992). Repeated measures ANOVAs were performed with analysis method (reference vs. the two video analysis methods separately) and type of lifting condition (symmetry, horizontal load distance and vertical load distance) as within subject factors; and peak and mean moments as dependent variables. In addition, repeated measures t-tests were used to compare the two video analysis methods with the reference method for each condition separately for peak and mean moments. For all statistical tests, $p<0.05$ was assumed to be significant. To assess the origin of possible errors, static and dynamic components of the total moment at the instant of peak moment were calculated. Furthermore, segment center of mass moment arms with respect to the L5S1 joint were calculated. For the peak moments intra-class correlations coefficients (ICCs) were calculated across subjects and conditions using ICC(3,1) for an individual estimate (Shrout & Fleiss, 1979). ICCs<0.40 were assumed poor, while ICCs 0.40–0.75 are good and ICCs>0.75 are excellent (Fleiss, 1986). Asymmetric components (i.e. trunk rotation, trunk lateral flexion, arm abduction and axial rotation) at the instant of peak moment were calculated from the reference method in all lifting conditions to assess the amount of asymmetry.

RESULTS

An analysis of the resulting asymmetry of the lifts at the instant of peak total moment showed relatively small trunk rotation and trunk lateral flexion (9.1 (4.6)° and 5.4 (3.4)°, respectively), however, a large whole body axial rotation (63.8 (42.5)°) in the asymmetric conditions (Table 6.1; Figure 6.1). Overall peak and mean moments were not significantly different between the reference method and the two video analysis methods, nor was there a significant interaction of analysis method with type of lifting condition (Table 6.2 and 6.3). Averaged peak moment errors were 4.49 (28.27) and 2.41 (27.84) Nm and averaged mean moments errors were 6.21 (13.88) and 1.81 (14.88) Nm, for video analysis methods 1 and 2, respectively. For both mean and peak moments, errors were not larger in asymmetric conditions compared to symmetric conditions (Table 6.2 and 6.3). T-tests on separate conditions showed no significant differences between the reference method and the two video analysis methods concerning peak moments in any of the conditions. However, for mean moments there was an overestimation of the moment in video analysis method 1 in two of the conditions (Table 6.2 and 6.3; Figure 6.2). Typical examples of total moment estimations obtained from video analysis method 2 and the reference method are shown in Figure 6.3. The static component of the moments shows some overestimation in both versions of the video analysis method by 10.28 (24.29) and 7.74 (24.12) Nm, respectively, while the dynamic components of the moment revealed some underestimation in both versions of the video analysis method by –6.82 (15.84) and –6.14 (16.27) Nm, respectively (Table 6.4). Moment arms of all segment centers of mass (Table 6.5) show relatively small errors in moments arms of the trunk and load (≤4 cm), and somewhat larger for the arms (≤12 cm).

Table 6.1 | Asymmetric components of the lifting tasks: trunk rotation, trunk lateral flexion, arm abduction and axial rotation (all expressed in degrees) obtained from the reference method for both the symmetric and asymmetric lifting conditions.

	Symmetric Conditions		Asymmetric Conditions	
Asymmetric components	Mean and Std. (Degrees)		Mean and Std. (Degrees)	
Trunk rotation	2.69	1.18	9.05	4.55
Trunk lateral flexion	1.07	0.62	5.41	3.43
Arm abduction	25.34	13.74	27.76	11.66
Axial Rotation	2.72	3.23	63.78	42.46

Table 6.2 | Outcomes of repeated measures ANOVAs testing for effects in peak moments for both variants of the video analysis method. p-values of within subject effects of the main and two-way interaction effects of the factor 'analysis method' are presented. Furthermore, differences in peak moments between the reference and video analysis methods 1 and 2, respectively, are presented for all lifting conditions separately. Differences averaged over subjects, standard deviations and levels of significance (repeated measures t-test) are presented. Differences averaged over subjects and conditions, all symmetric conditions and all asymmetric conditions are shown as well.

	ANOVA	
Factor	Video Analysis Method 1	Video Analysis Method 2
Analysis	0.47	0.70
Analysis*Vertical	0.87	0.85
Analysis*Horizontal	0.12	0.11
Analysis*Symmetry	0.27	0.43

T-test

Condition				Video Analysis Method 1			Video Analysis Method 2		
Nr.	Symmetry	Vertical	Horizontal	Mean and Std. (Nm)		Sig.	Mean and Std. (Nm)		Sig.
1	Symmetric	Ground	Close	16.00	28.51	0.11	15.56	28.54	0.12
2			Far	13.73	37.53	0.28	13.55	37.74	0.29
3		Shoulder	Close	1.21	20.94	0.86	-3.23	22.40	0.66
4			Far	9.59	23.46	0.23	4.15	19.80	0.52
5		Hip	Close	-3.78	23.69	0.63	-7.87	24.12	0.33
6			Far	9.51	35.25	0.42	5.85	35.21	0.61
All symmetric conditions				7.71	28.52		4.67	28.75	
7	Asymmetric	Ground	Close	6.72	24.58	0.41	6.90	24.32	0.39
8			Far	5.71	23.28	0.46	4.45	22.32	0.54
9		Shoulder	Close	-8.90	19.63	0.19	-8.68	18.60	0.17
10			Far	-1.94	22.03	0.79	0.88	23.03	0.91
11		Hip	Close	-7.99	35.20	0.49	-12.61	34.12	0.27
12			Far	14.06	37.16	0.26	9.98	34.38	0.38
All asymmetric conditions				0.22	27.88		-0.73	26.94	
All conditions				4.49	28.27		2.41	27.84	

Table 6.3 | Outcomes of repeated measures ANOVAs testing for effects in mean moments for both variants of the video analysis method. p-values of within subject effects of the main and two-way interaction effects of the factor 'analysis method' are presented. Furthermore, differences in mean moments between the reference and video analysis methods 1 and 2, respectively, are presented for all lifting conditions separately. Differences averaged over subjects, standard deviations and levels of significance (repeated measures t-test) are presented. Differences averaged over subjects and conditions, all symmetric conditions and all asymmetric conditions are shown as well. Bold numbers indicate significant values ($p<0.05$).

	ANOVA	
Factor	Video Analysis Method 1	Video Analysis Method 2
Analysis	0.08	0.64
Analysis*Vertical	0.88	0.89
Analysis*Horizontal	0.77	0.53
Analysis*Symmetry	0.09	0.12

T-test

Condition				Video Analysis Method 1			Video Analysis Method 2		
Nr.	Symmetry	Vertical	Horizontal	Mean and Std. (Nm)		Sig.	Mean and Std. (Nm)		Sig.
1	Symmetric	Ground	Close	7.28	12.33	0.09	1.61	15.22	0.75
2			Far	3.67	15.98	0.49	0.63	16.48	0.91
3		Shoulder	Close	6.70	9.74	0.06	4.67	9.49	0.15
4			Far	12.56	17.22	**0.04**	7.92	13.79	0.10
5		Hip	Close	5.85	9.50	0.08	-0.70	10.87	0.84
6			Far	11.73	19.40	0.09	4.04	20.71	0.55
All symmetric conditions				7.97	14.26		3.03	14.54	
7	Asymmetric	Ground	Close	1.02	14.50	0.83	-0.53	17.13	0.92
8			Far	-2.26	14.52	0.63	-3.11	19.65	0.63
9		Shoulder	Close	8.96	13.98	0.07	5.34	11.13	0.16
10			Far	4.84	9.74	0.15	1.38	8.61	0.62
11		Hip	Close	5.92	16.45	0.28	-2.93	19.89	0.65
12			Far	8.20	9.29	**0.02**	3.37	13.61	0.45
All asymmetric conditions				4.41	13.37		0.53	15.24	
All conditions				6.21	13.88		1.81	14.88	

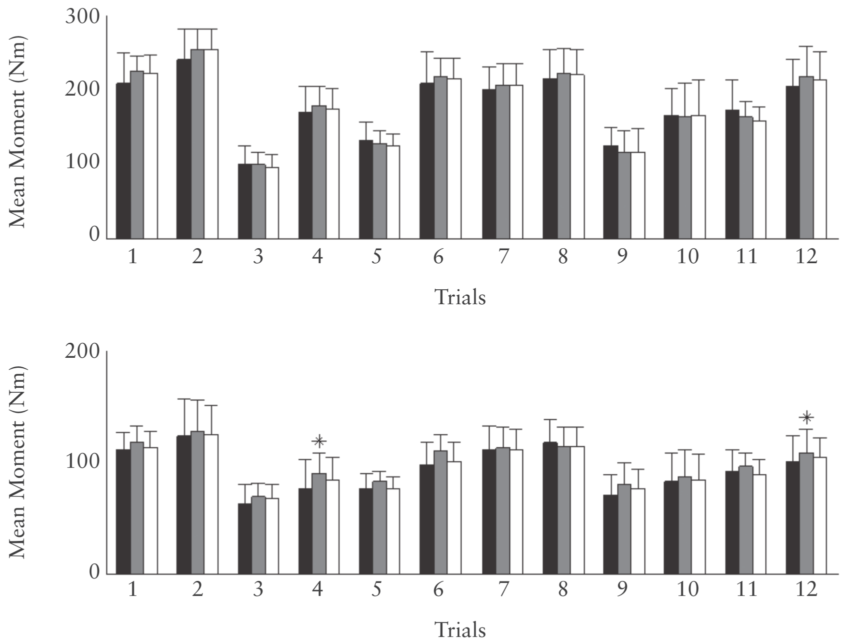

Figure 6.2 | Peak (upper panel) and mean (lower panel) total low-back moments of the 12 lifting conditions. Moments averaged over subjects and standard deviations (error bars) are presented. Moments estimated by the reference method (black bars), video analysis method 1 (gray bars) and analysis method 2 (white bars) are presented. * indicates significant differences ($p<0.05$) of one of the video analysis methods compared to the reference method. Trial numbers correspond to the numbers indicated in Table 6.2 and Table 6.3.

ICCs of peak moments over all pooled individual conditions (12 conditions x 10 subjects) were 0.86 between the reference method and both video analysis methods (Figure 6.4). The ICCs were higher when data were averaged over conditions (0.98 for both versions) and were lower when data were averaged over subjects (0.72 and 0.73 for video analysis methods 1 and 2, respectively; Figure 6.5).

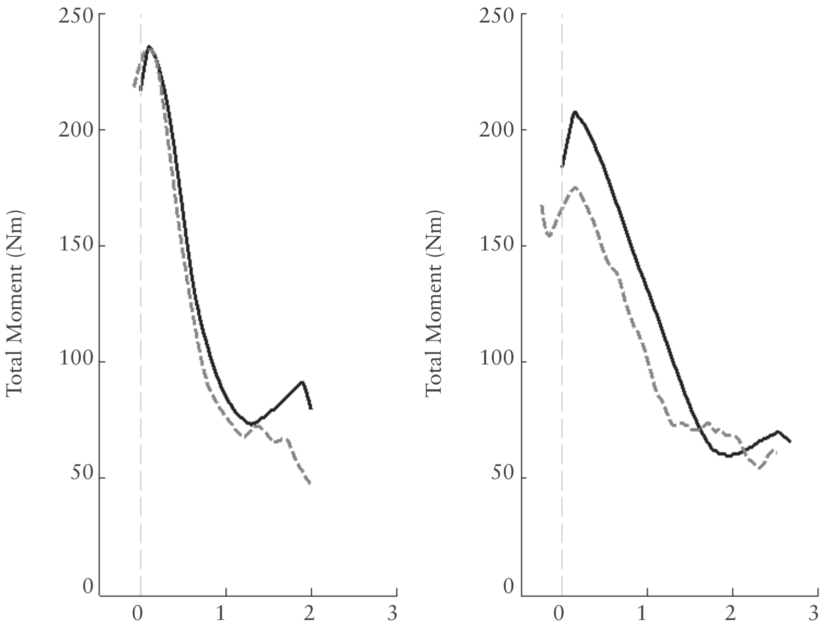

Figure 6.3 | Typical examples of total low-back moments obtained from video analysis method 2 (solid lines) and the reference method (dashed lines) in two lifting conditions. The left panel displays a relatively good fit of the video analysis method to the reference method for a symmetric lifting condition from a floor level initial lifting position. The right panel displays an overestimation of the moment obtained by the video analysis method compared to the reference method in an asymmetric lifting condition from a hip height initial lifting position. The error in the right panel is mainly caused by static errors (i.e. errors in positioning of the manikin). The slightly sharper peak in the video analysis method is a consequence of the spline interpolation based on a limited number of video frames. Examples of video analysis method 1 are comparable.

Table 6.4 | Mean and standard deviations of difference in static and dynamic components of the total moments at instant of peak in both versions of the video analysis method compared to the reference method. The most right columns present the mean and standard deviation of static and dynamic components of the total moment obtained from the reference method.

	Difference in Video Analysis Method 1		Difference in Video Analysis Method 2		Moment from reference method	
	Mean and Std. (Nm)		Mean and Std. (Nm)		Mean and Std. (Nm)	
Static Moments	10.28	24.29	7.74	24.12	162.30	47.06
Dynamic Moments	-6.82	15.84	-6.14	16.27	19.71	14.79

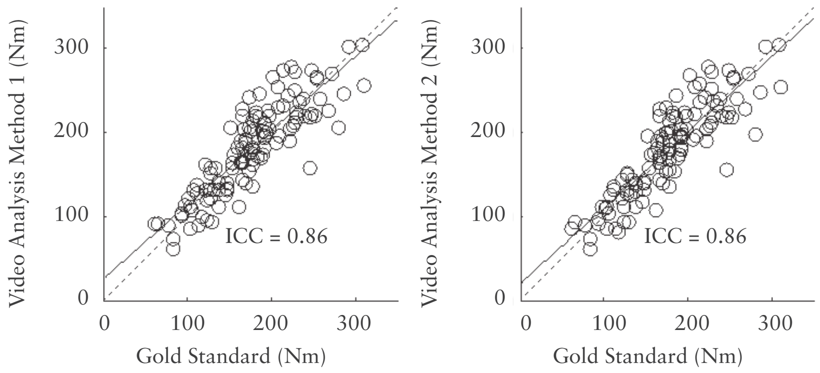

Figure 6.4 | Scatter plots illustrating the relations between peak moment estimated by the reference method and video analysis method 1 (left panel) and video analysis method 2 (right panel). Data of all subjects in all lifting conditions are presented. Furthermore, a linear fit through the data points (solid line) and a x=y reference line (dotted line) are plotted and calculated ICCs are presented.

Table 6.5 | Mean and standard deviations of differences in segment moment arms of the trunk/head, upper arms, lower arms/hand and load segments with respect to the L5S1 joint (expressed in m) for both versions of the video analysis method compared to the reference method. Moment arms are presented for all lifting conditions and for the symmetric and asymmetric lifting conditions separately.

		Video method 1		Video method 2	
Segment		Mean (m)	Std	Mean (m)	Std
Trunk/head	Symmetric conditions	0.02	0.04	0.02	0.04
	Asymmetric conditions	0.01	0.04	0.01	0.04
	All conditions	0.02	0.04	0.02	0.04
Upper Arms	Symmetric conditions	0.09	0.11	0.09	0.11
	Asymmetric conditions	0.04	0.14	0.04	0.14
	All conditions	0.06	0.13	0.06	0.13
Lower Arms	Symmetric conditions	0.10	0.07	0.12	0.08
	Asymmetric conditions	0.05	0.10	0.07	0.10
	All conditions	0.07	0.09	0.09	0.10
Load	Symmetric conditions	-0.04	0.09	-0.02	0.09
	Asymmetric conditions	-0.01	0.10	0.00	0.09
	All conditions	-0.02	0.09	-0.01	0.09

DISCUSSION

In the present study, we aimed to develop, compare and validate two versions of a video analysis method for the assessment of low-back moments during occupational lifting by a comparison with a reference method. ANOVA results revealed no overall differences in peak and mean moments between the reference method and the two video analysis methods. Furthermore, all conditions separately showed no systematic differences for peak moments between the two video analysis methods and the reference method, however, there was an overestimation of the mean moments in two conditions for video analysis method 1. The ICCs revealed a strong correspondence between the video analysis method and the reference method concerning the assessment of peak moments. This correspondence was stronger for data averaged over conditions compared to data averaged over subjects, which can be explained by the higher variance between conditions than between subjects. While we found only 2 small but significant differences between the reference method and one of the video analyses methods, due to the relative small sample size combined with large standard errors, we cannot exclude that with a higher sample size, some more differences might have become significant. However, as can be appreciated from Figure 6.2, the magnitude of the differences was small, so that even if a difference would become significant, it would likely be small. Note however that, while systematic errors in video analysis method 2 were absent, random errors were substantial as shown by the relatively large standard deviations (Tables 6.2 and 6.3). These data indicate that the proposed video analysis method is useful to determine differences in back load between subjects as well as between conditions. However, reliable back load estimation with video analyses does require a substantial number, i.e. about 10, repeated conditions.

The importance of establishing back load during lifting is underlined by in vitro studies showing damage to spinal segments at high peak (Brinckmann et al., 1989; Hansson et al., 1980) and repetitive loads (Brinckmann et al., 1988; Hansson et al., 1987). Furthermore, epidemiological studies have shown that peak (Norman et al., 1998) and cumulative low-back loads (Kumar, 1990; Norman et al., 1998) are biomechanical risk factors for LBP. While back load can be established accurately in the laboratory (Kingma et al., 1996), lifting behavior may differ between laboratory and actual working conditions, which highlights the importance of establishing back load at the workplace (Faber et al., 2011). The results of the present study show that the two versions of the video based method are valid for mean and peak moment determination up to a moderate level of asymmetry, thereby providing a useful tool for epidemiological studies on dose–response relationships and for ergonomic practice.

While errors were not explicitly compared between the two versions of the video analysis method, Tables 6.2 and 6.3 suggest that errors were smaller in video analysis method 2. ICCs were comparable for both video analysis methods. Due to these findings and since video method 2 roughly halves the analysis time compared to video method 1, video analysis method 2 seems to be the best applicable method for future research and ergonomic applications.

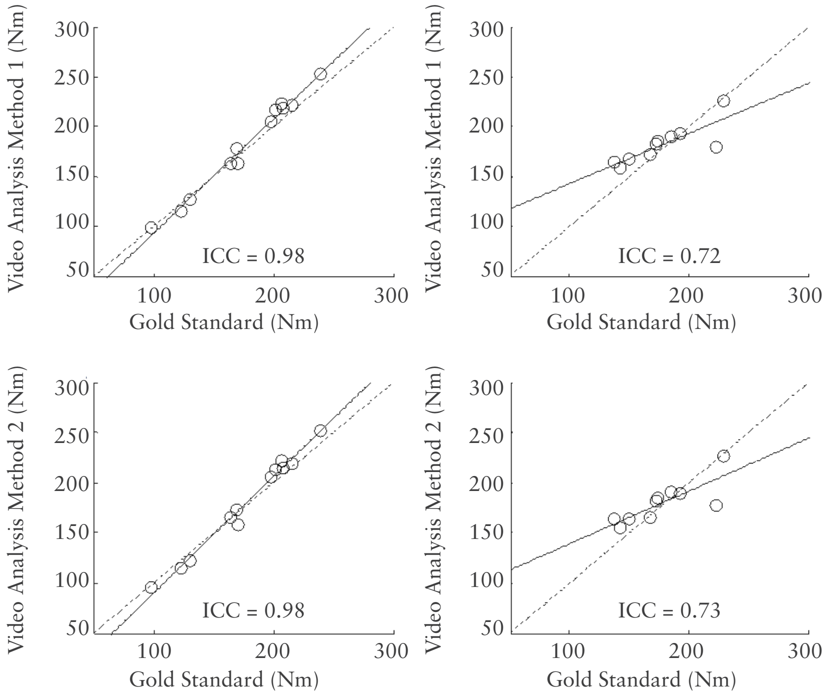

Figure 6.5 | Scatter plot illustrating the relations between peak moment estimated by the reference method and video analysis method 1 (upper row plots) and video analysis method 2 (lower row plots). Data are averaged over conditions (left plots) and over subjects (right plots). Furthermore, linear fits trough the data points (solid line) and x=y reference line (dotted line) are plotted and calculated ICCs are presented.

The video analysis method presented has a number of advantages compared to models presented earlier. Moments were obtained from a dynamical analysis, meaning that not only the gravitational contribution of the moments but also the angular and linear acceleration contributions were taken into account. Since, the dynamic component of the moment accounted for approximately 11 percent of the total moment for the lifting conditions studied and an average error of less than 4 percent of the total moment was made in the dynamic moment component, it can be concluded that by adding dynamic components to the moments, accuracy of the total moment improves. Furthermore, several studies have reported on the problem of assessing movement outside the sagittal plane due to projection biases (Kingma et al., 1998; Paul & Douwes, 1993). With the current model we aimed at decreasing this source of error since we allowed for axial rotation of the manikin and quasi-three-dimensional movements (i.e., trunk rotation, trunk lateral flexion and arm abduction). The validity of this approach was supported by the fact that errors were not larger in asymmetric conditions compared to symmetric conditions. Although errors in symmetric and asymmetric conditions were not explicitly compared, the non-significant interactions of analysis method and symmetry indicate no differences in errors for peak and mean moments between symmetric and asymmetric conditions. Allowing axial rotation of the manikin appeared to be useful as Table 6.1 showed that those rotations were much larger than the out of plane motions of the trunk in the present study, and did not negatively affect the accuracy. A last source of errors that we aimed to overcome with the present method is the error made by crude categorization of segment orientations (de Looze et al., 1994b; van Wyk et al., 2009), since matching of body orientations can be performed on a continuous scale.

Besides the advantages of the presented video analysis method there are some methodological limitations that have to be taken into account. While we could accommodate for body postures deviating from the plane of the video camera, we cannot exclude projection errors. Nevertheless, asymmetric lifting did not result in larger errors than symmetric lifting, suggesting that projection errors did not play an important role. However, in the present study, moderately asymmetric conditions were studied and although these conditions show substantial asymmetric components with respect to the whole body axial rotation, we cannot exclude that larger errors will occur in other lifting conditions, especially in conditions with more asymmetric trunk and arm movements. Furthermore, in the conditions measured in this study, a box with an even distribution of mass was used. It is not known whether this model can also be applied to conditions in which loads with an uneven mass distribution are lifted. In addition, the separate analysis of static and dynamic moment components showed some systematic overestimation of static moments and some underestimation of dynamic components. Most likely, the overestimation of static moments is due to errors in modeling of the trunk. During forward bending, curvature of the trunk occurs, which reduces the distance between hip and shoulder. In the present video methods, the estimated flexion in the hip and L5S1 joints was based on total trunk inclination and the knee angle, as proposed by Andersson et al. (1985). However, this procedure may have caused some errors since modeling the entire trunk in a pelvis and an upper trunk segment might not provide an accurate representation of the trunk curvature (Lariviere & Gagnon, 1999), as shown by the small overestimation of trunk center of mass moment arm. Furthermore, this procedure does not accommodate sideward bending of the pelvis, so that application to asymmetric lifting could introduce errors. However, in the present study, pelvic sideward bending was hardly noticed and asymmetry was adequately covered by allowing for axial rotation of the whole manikin. Furthermore, in asymmetric lifting conditions symmetry in the lower extremities has been assumed, and this might introduce some error in pelvis orientation. The underestimation of the dynamic component of the moment might have been caused by the spline interpolation between the four key data points, which may cause a somewhat smoother movement trajectory compared to what subjects actually do. Improved interpolation or posture prediction algorithms can possibly be used in future studies to improve interpolation accuracy and reduce analysis time (Zhang & Chaffin, 2000). However, benefits from such improvements can be limited as random errors in positioning the manikin will persist (Xu et al., 2010b). Furthermore, all observations have been performed by the same observer. Therefore, no statements can be made about the inter-rater reliability of the present analysis method. However, since the fit of the stick figure is made within the video frame, and can thus be checked visually, the effect of the expertise of the observer can be assumed to be relatively small. Finally, the video analysis method was tested on a group of healthy young subjects. Generalization of these results should be done with caution as it is not obvious that our results will hold for subjects with deviating anthropometry or lifting behavior (e.g. due to LBP; Marras et al., 2004).

CONCLUSION

The present study reports on two variants of a video analysis method, a simple and relatively cheap method for the assessment of low-back loads during occupational lifting. The absence of substantial differences with the reference method supports the validity of the video method of establishing back load in ergonomic practice and epidemiological studies for lifts up to a moderate level of asymmetry. However, the presence of substantial random errors suggests that care should be taken in interpreting results when only few measurements can be taken.

Chapter 7

Inter-rater reliability of a video-analysis method measuring low-back load in a field situation.

P. Coenen
I. Kingma
C.R. Boot
P.M. Bongers
J.H. van Dieën

Applied Ergonomics, 2013,
44(5): 828-834

ABSTRACT

Valid and reliable low-back load assessment tools that can be used in field situations are needed for epidemiologic studies and for ergonomic practice. The aim of this study was to assess the inter-rater reliability of a low-back load video analysis method in a field setting.

Five raters analyzed 50 work site manual material handling tasks of 14 workers. Peak and mean moments at the level of L5S1, and segment angles were obtained using the video analysis method. Intra-class correlation coefficients (ICCs) and median standard deviations across raters were calculated.

ICCs revealed excellent inter-rater reliability (>0.9) for peak and mean moments, ICCs of segment angles were variable. Median standard deviations showed relatively small inter-rater variance for moments (standard deviation <10 Nm) and segment angle variation ranging from 0° to 20°. The proposed video analysis method provides a reliable tool for obtaining low-back loads from occupational field tasks.

INTRODUCTION

High low-back loads that may occur at work (e.g. during lifting, pushing and pulling of objects or working in awkward body positions) are associated with low-back pain (LBP; e.g., Marras et al., 2010; van Dieën et al., 1999). These associations have often been confirmed in epidemiological studies using self-reported exposures or field observations (da Costa & Vieira, 2010; Griffith et al., 2012; Lötters et al., 2003). However, other epidemiological studies did not find support for the association between high low-back loads and LBP, possibly as a result of the lack of appropriate measurement designs (Bakker et al., 2009). Therefore, valid and reliable low-back load assessment methods that can be applied in field settings are needed. Three types of measurement methods can be adopted: self-reports, observational techniques and direct measurement techniques (Burdorf, 2010; David, 2005). Although self-reports are highly efficient, they are assumed to be less reliable than observational techniques and direct measurements (Balogh et al., 2004; Hansson et al., 2001). On the other hand, direct measurement techniques (e.g., measuring muscle activity or body posture recordings using marker tracking or goniometry) are much more accurate but difficult to apply in large scale field studies. In field measurements of low-back load, there thus seems to be a trade-off between efficiency (in terms of time, money and resources) and accuracy. Besides, it can be argued that crude observational low-back exposure measures (e.g., the number of lifts, time spent in a flexed trunk position) provide less detailed information on low-back load than dose metrics (i.e., low-back moments), since different exposures (e.g., lifting and bending) affect the same dose. Therefore, dose-estimates can provide more insight into the etiology of LBP (Wells et al., 2004) and these metrics are more predictive of future LBP than postural exposure measures (Coenen et al., 2013b).

Video-based methods using postural exposure data in biomechanical models to calculate low-back load dose estimates have been shown to be a promising category of observational techniques (e.g., Chang et al., 2010; Coenen et al., 2011; Norman et al., 1998; Potvin, 1997; Sutherland et al., 2008) in the assessment of low-back load metrics such as static (Neumann et al., 2001b), cumulative (Sutherland et al., 2008) or peak low-back moments (Norman et al., 1998). Furthermore, these coding systems allow raters with minimal training and minor use of equipment to collect occupational low-back load data. High inter-rater agreement has been found when using these kinds of models to calculate cumulative low-back moments (Cann et al., 2008; Sullivan et al., 2002). However, testing of these models was only performed in laboratory situations or in mock-ups of field situations, whereas, applicability of these methods for epidemiological studies or in ergonomic practice can best be assessed when applied to actual field situations. The aim of the present study therefore was to test the inter-rater reliability of a low-back load video analysis method in a field setting. The model that will be tested in our study has been validated against a lab-based reference method (Coenen et al., 2011) and inter-rater reliability has been assessed in a laboratory situation (Xu et al., 2011). Although these authors suggest that the method might be valid and reliable in field studies, reliability has not yet been assessed in field settings.

MATERIAL AND METHODS

Data collection

Videos of a wide range of manual materials handling (MMH) tasks were selected from the SMASH cohort that has been described before (Ariëns et al., 2001; Hoogendoorn et al., 2000a). Briefly, in this cohort, risk factors of musculoskeletal disorders were studied in workers from various industrial and service branches, for example, in the metal, chemical, pharmaceutical, food and wood construction industry; waste processing, insurance and distribution companies. The SMASH study consists of a baseline measurement, assessing physical load at the workplace, and baseline and three year follow-up assessment of musculoskeletal symptoms. For the assessment of physical work load, 5–15 min of video recordings at the workplace were taken at four moments during the course of one day. During these recordings, researchers handling the camera were instructed to take a sagittal plane view as much as possible. For all MMH tasks during these 15 min, external forces at the hands were measured using force transducers (during pushing and pulling) or weighing scales measuring mass of the external load (during lifting). Afterward, videos were systematically observed during which MMH tasks, i.e. lifting, pushing and pulling tasks during which external forces are exerted on the hands, were identified. Fifty video fragments were selected representing tasks (38 lifting, 6 pushing and 6 pulling tasks), executed by 14 workers of 10 particular companies. Rather than randomly selecting, we carefully selected these tasks, in order to obtain a wide range of tasks, work postures, task asymmetry, physical workloads and image quality and camera angle relative to the sagittal plane of the subject. Thus, we also included tasks that had not been recorded optimally, e.g. due to occlusion of the view by another worker or with a large angle between the camera plane and the sagittal plane of the subject. The selected workers were 31.9(8.3) years of age and seven workers were female. Six workers reported LBP at baseline. External forces at the hands measured during these tasks were on average 66 (80) N and ranged from almost 0 N to 368 N.

Five raters were recruited among students of the Amsterdam School of Health Professions. Three of them were third year physical therapy students and two of them were fourth year occupational therapy students. The raters were 22.2 (1.8) years of age and had substantial knowledge on kinesiology. After participating in an extensive learning and practice session in which the raters were briefed regarding the purpose of the study and were familiarized with the software, raters analyzed videos of all tasks. Raters analyzed videos independently from each other and were blinded to each other's results.

Video analysis

The video analysis method that was used in this study was described in detail earlier (Coenen et al., 2011). In short, beginning and ending frames of the task were selected from the video fragments by each rater. For lifting tasks, the start of a task was defined as the moment the load is clear from its surface, while the end of the task is the moment in which the end position of the load is reached. For pushing and pulling tasks, the task was defined as the period in which the worker is exposed to external forces at the hands due to resistance of the load. In addition, two intermediate frames, equally spaced in time between the beginning and end frame, were automatically selected to obtain four video frames. In these four video frames, a semi three-dimensional manikin was constructed consisting of nine segments (right foot, lower leg and upper leg; pelvis, trunk/head, two upper arms, two forearms/hands). This manikin allows for semi three-dimensional analysis of movements (ankle flexion/extension, knee flexion/extension, hip flexion/extension, trunk flexion/extension, trunk rotation, trunk lateral flexion, shoulder flexion/extension, shoulder abduction and elbow flexion/extension). Furthermore, the manikin can be scaled, rotated around its longitudinal axis (axial rotation) and translated horizontally and vertically along the video frame (Figure 7.1). Each rater made an optimal fit of the manikin to the four video frames for each of the 50 tasks by adjusting all segment orientations. Subsequently, for each task and rater, a cubic spline interpolation of the segment angles over the four key frames was executed to estimate body kinematics of the worker with a time resolution of 25 Hz. In case a MMH task lasted less than 2 s, only the first and the last frame instead of four video frames were used for cubic spline interpolation to avoid unrealistically high accelerations due to random errors in fitting the manikin. This interpolation method has been validated in a lab-based study before (Xu et al., 2010b). Based on total body mass and stature, individual segment masses and lengths, positions of the center of mass and inertia tensors were estimated using regression equations (Zatsiorsky, 2002). Hand forces were obtained from measured forces (at the time of video recording) in case of pushing and pulling, and from object weight (obtained at the time of video recording) and hand acceleration in case of lifting.

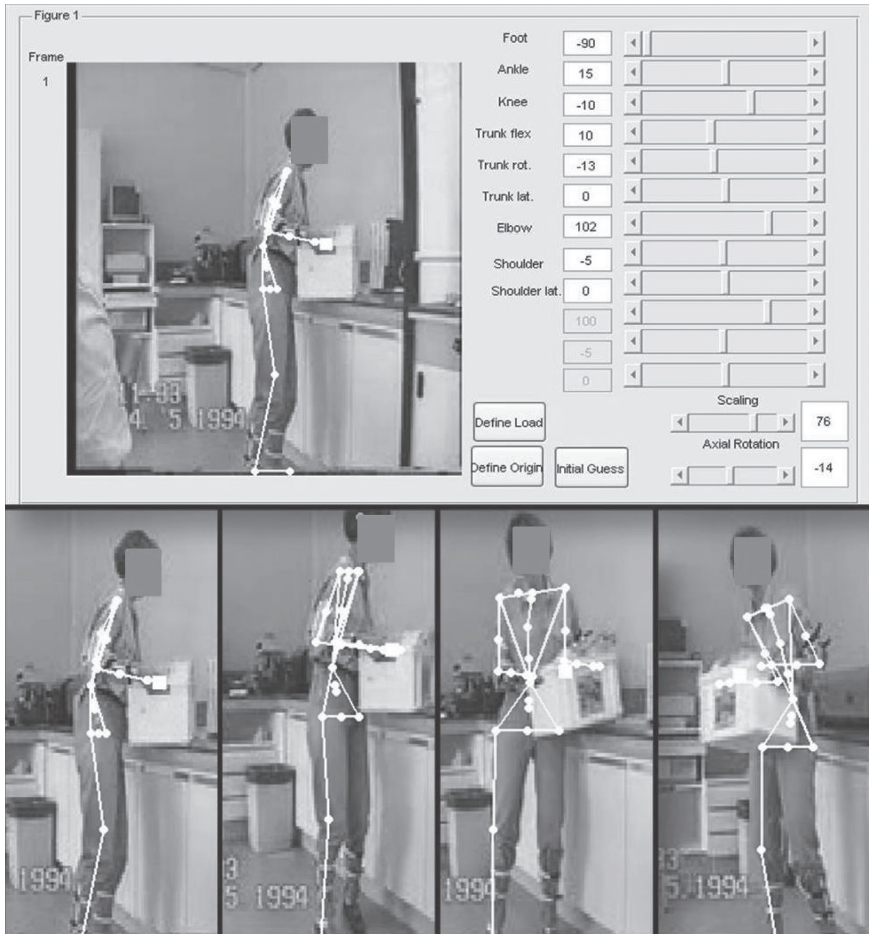
Foot -90
Ankle 15
Knee -10
Trunk flex 10
Trunk rot. -13
Trunk lat. 0
Elbow 102
Shoulder -5
Shoulder lat. 0
100
-5
0
Scaling 76
Axial Rotation -14

Figure 7.1 |Video analysis method. The graphical user interface depicting a three-dimensional manikin plotted onto a video frame is shown (upper part of the figure). In the lower part of the figure, a typical example of four key video frames of a field-based lifting task is shown that was analyzed by one of the observers.

A top-down inverse dynamics calculation using hand forces, segment kinematics (obtained from the interpolated manikin postures) and anthropometrics was performed to calculate dynamic moment components (derived from segment acceleration), static moment components (derived from gravitational forces on upper body segments and external forces at the hands) and total moments (static plus dynamic components) at the level of the L5S1 joint. For further analysis, the resultant moment (i.e., the resultant of the moments around three axes) was considered. Both the moment at the instant of peak total moment and moments averaged over the entire task's time series were obtained. As horizontal load distances of the load with respect to the L5S1 joint is an important input variable for low-back load, horizontal low-back to load distance at the instant of peak moment was assessed. For further analyses, the abovementioned low-back load dose metrics and horizontal load distance and segment orientation angles at the instant of peak moment obtained from the interpolated manikin fit over the workers by each rater, were collected.

Data analysis

Intra-class correlation coefficients (ICCs) were calculated to assess the agreement among the five raters in the estimation of L5S1 peak and averaged moments (total moments; dynamic and static components of the moments), horizontal load distance and the segment angles. ICCs <0.40 were assumed poor, ICCs 0.40–0.75 were assumed good and ICCs >0.75 were assumed excellent (Fleiss, 1986). Furthermore, for the above-mentioned variables, standard deviations over the raters were calculated for each task while the median of these standard deviations over the 50 tasks was calculated to quantify inter-rater variability (Bao et al., 2009; Rothman & Greenland, 2005).

An additional analysis was performed in which inter-rater median standard deviations were assessed for lifting and for pushing/pulling tasks separately for peak and averaged total moments. This analysis was performed to test whether the variability among raters differed in lifting tasks compared to pushing/pulling tasks. Non-parametric Mann–Whitney-U tests were used to test for significant differences between lifting and pushing/pulling tasks assuming p-values <0.05 to be statistically significant.

RESULTS

Peak and mean moments across all tasks were on average 88.17 (15.83) Nm and 68.59 (11.39) Nm respectively. Furthermore, axial rotation across all tasks was on average 29 (31)° at the beginning of the tasks and changed on average 34 (67)° during the tasks.

ICCs were excellent for both peak (ICC = 0.92) and averaged (ICC = 0.91) L5S1 moments (Table 7.1). ICCs were substantially larger, but median inter-rater standard deviations were substantially larger as well for the static (ICC >0.90 and median standard deviation >8.2 Nm) compared to the dynamic (ICC <0.71 and median standard deviation <2.6 Nm) component of L5S1 moments, both with respect to peak (Table 7.1; Figure 7.2) and mean moments (Table 7.1; Figure 7.3). Concerning standard deviation of low-back moments, some occasional outliers for peak (>40 Nm) and mean moments (>30 Nm) were found (Figures 7.2 and 7.3).

ICCs of segment angles ranged from poor (trunk rotation and shoulder abduction), to good (trunk lateral flexion, shoulder flexion and elbow flexion) and excellent (trunk flexion; Table 7.1). Median standard deviations of the segment angles were low (<5°) for the three trunk angles and for shoulder abduction and were higher (>14°) for elbow and shoulder flexion (Table 7.1). Resultant horizontal load distance with respect to the L5S1 joint showed small median standard deviation (0.08 m) and good ICCs. Non-parametric Mann–Whitney-U tests revealed no significant differences for median standard deviations of peak (p = 0.64) and mean moments (p = 0.76) between lifting and pushing/pulling tasks (Figure 7.4).

Table 7.1 | Absolute values (mean and standard deviation over 50 tasks after averaging over 5 observers) and inter-rater reliability estimates (intra-class correlation coefficient (ICC) and median over 50 tasks of the standard deviation over five observers) of low-back moments, and of segment angles and load distance at the instant of peak moment, obtained from the video analysis. Average values, standard deviations and median standard deviations are expressed in Nm for moments, in degrees for segment angles and in meters for load distance.

		Absolute Values		Inter-rater reliability	
Variable		Mean	Std.	ICC	Median Std.
Moments					
Peak moment	Total	88.17	15.83	0.92	8.80
	Static	79.96	12.92	0.93	8.85
	Dynamic	8.20	8.92	0.69	2.54
Mean moment	Total	68.59	11.39	0.91	8.31
	Static	63.65	11.22	0.91	8.63
	Dynamic	4.95	5.20	0.70	1.24
Segment angles					
Trunk flexion		13.87	2.60	0.91	3.58
Trunk rotation		0.14	5.07	0.26	4.89
Trunk lateral flexion		2.08	3.05	0.72	1.88
Elbow flexion right		72.35	10.81	0.63	16.22
Shoulder flexion right		26.33	10.11	0.61	14.49
Shoulder abduction right		4.83	10.36	0.33	4.25
Elbow flexion left		71.76	12.30	0.50	20.71
Shoulder flexion left		24.82	11.05	0.54	15.73
Shoulder abduction left		4.31	10.31	0.26	0.00
Load distance		0.43	0.16	0.63	0.08

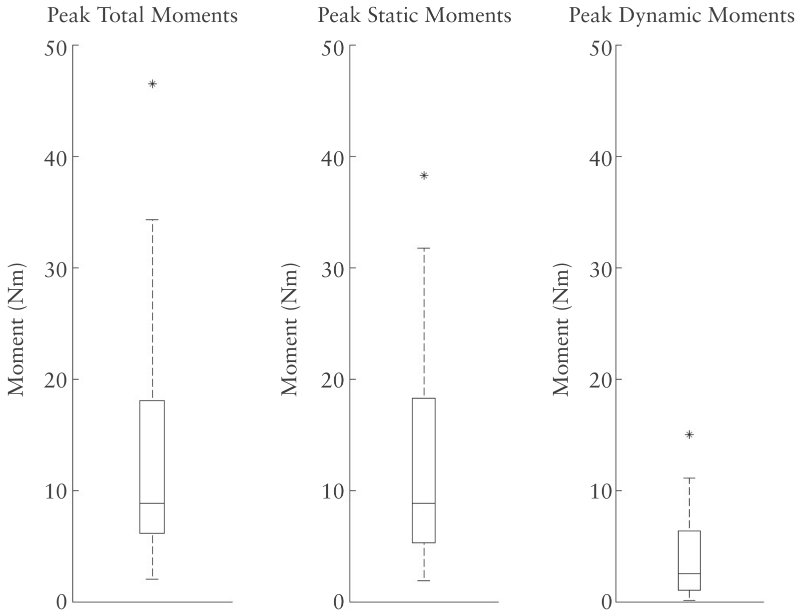

Figure 7.2 | Standard deviations across raters of all rated tasks concerning moments at the instant of peak of the total moment. The middle notch represents the median standard deviation, the box presents the standard deviations of the 25th percentile to the 75th percentile, whiskers represents the 5th to 95th percentile interval and asterisks represent outliers. Total moments (left plot), static component (middle plot) and dynamic component (right plot) of moments are shown. Values were calculated over all 50 tasks.

DISCUSSION

The aim of our study was to evaluate the inter-rater reliability of a video analysis method to estimate low-back load in work field situations. Our main focus was to assess low-back load dose estimates (i.e., low-back peak and mean moments) as these metrics are expected to provide more insight into low-back load than postural exposures (Wells et al., 2004), leading to stronger associations with LBP (Coenen et al., 2013b). Results show excellent ICCs for total low-back moment estimates. Median standard deviations assessing inter-rater variation were relatively low, i.e. about 10% of total moments. Inter-rater reliability was lower for dynamic components of the low-back moments compared to static components. The relatively low inter-rater reliability in dynamic moment components may partly be caused by the fact that inevitable random errors in positioning the manikin are strongly magnified due to double differentiation of position and angle data (Xu et al., 2010b). However, as shown before (Coenen et al., 2011; van Dieën et al., 2010), dynamic components of the moments are only a small percentage of the total moment (i.e., about 10%; Table 7.1). Therefore, errors in dynamic components only contribute for a small part to errors in total moments. However, actual accelerations cannot be obtained from these data. The number of frames is a trade-off between the random errors in individual frames, the effect of which is increasingly magnified by differentiation when time intervals between frames are shorter, and the number of frames required to adequately cover the whole movement. It has been shown that using more than four frames does not improve the results when taking random errors in matching manikins to video frames into account (Xu et al., 2012). In the present study we observed that, as a result of the above-mentioned trade-off, for tasks lasting less than 2 s, using four frames resulted in unrealistically large accelerations. To avoid these unrealistically large accelerations, we decided to use the first and the last frame for interpolation instead of four video frames for tasks lasting less than 2 s. While Xu et al. (2012) showed that (random) errors increase by about 50% when taking 2 instead of 4 samples, we found in tasks with a duration less than 2 s that random errors caused unrealistic accelerations and a subsequent dramatic increase in inter-subject variation (up to over 100%). Regrettably, we could not check the validity of our approach to select 2 s as a threshold. Besides, in the study described by Xu and colleagues, only standardized tasks were studied in a laboratory situation, whereas we studied non-standardized field MMH tasks.

We found no significant differences in inter-rater variation of lifting tasks compared to pushing/pulling tasks for peak and averaged moments, suggesting that the current video analysis method is equally applicable to these three types of MMH tasks. As the tasks selected for our study were only a small proportion of all available tasks in the SMASH cohort, it can be argued that our selection may not be representative for the whole SMASH cohort or for MMH tasks in general. However, the tasks selected for our study were carefully chosen to cover a broad range of tasks from the original SMASH cohort with varying camera angles and occlusion of body segments. Therefore, the selection of workers and tasks used in the current study is considered representative for a broad range of workers, jobs and work settings. As an additional test, ICCs of the low-back loads within all subjects performing more than two tasks were assessed. These ICCs ranged from 0.68 to 0.99 for peak moments and from 0.42 to 0.99 for average moments. These results show that inter-rater agreement varied substantially across workers which is attributable to the variable quality and plane of video images across workers, as well as to the magnitude of the range of low-back loads within workers. While our findings may not be extrapolated to highly asymmetric or highly dynamic tasks, the high ICCS and low standard deviations in our low-back load estimates suggest that the proposed method is applicable for a broad range of tasks, both with and without asymmetry, variation in dynamics and load handled.

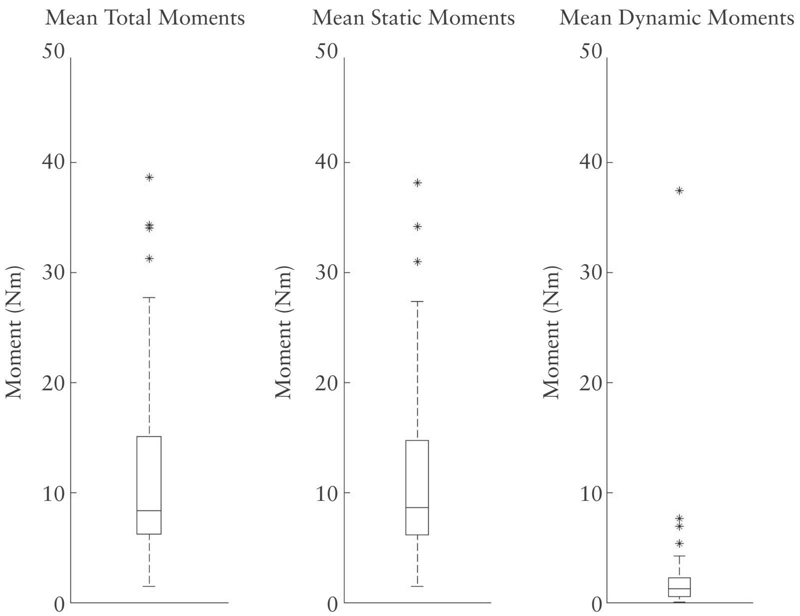

Figure 7.3 | Standard deviations across raters of all rated tasks concerning averaged moments. The middle notch represents the median standard deviation, the box presents the standard deviations of the 25th percentile to the 75th percentile, whiskers represents the 5th to 95th percentile interval and asterisks represent outliers. Total moments (left plot), static component (middle plot) and dynamic component (right plot) of moments are shown. Values were calculated over all 50 tasks.

Excellent inter-rater reliability was shown for trunk flexion angle; raters agreed well for trunk lateral flexion and elbow and shoulder flexion, however, agreement of trunk rotation and shoulder abduction was poor. In part, this may be due to less precise positioning of the manikin in the frontal and transverse plane relative to the sagittal plane. However, also median standard deviations showed varying inter-rater differences for segment angles. Since ICC is the ratio of the between task variance and the total variance (variance between tasks, variance between observers and random variance; Shrout & Fleiss, 1979), the ICC can be poor when the variance in observations is small (Bao et al., 2009). In our study, most raters estimated small movements outside the sagittal plane (e.g. trunk lateral bending, trunk rotation and shoulder abduction), leading to small variations in observations which can explain the poor ICCs for these segment orientations. For example, for shoulder abduction poor agreement was shown (ICCs of 0.33 and 0.26) that can be explained by rather small inter-rater standard deviations (4.25° and 0°; Table 7.1). In addition, trunk rotation and lateral flexion was rather small. However this was not due to little task asymmetry. Substantial asymmetry in the filming of tasks as well as axial rotation of the subjects during the tasks occurred as axial rotation across all tasks was on average 29 (31)° and changed on average 34 (67)°. Notably, however, workers mainly adapted to task asymmetry by whole body rotation rather than by adopting asymmetric postures.

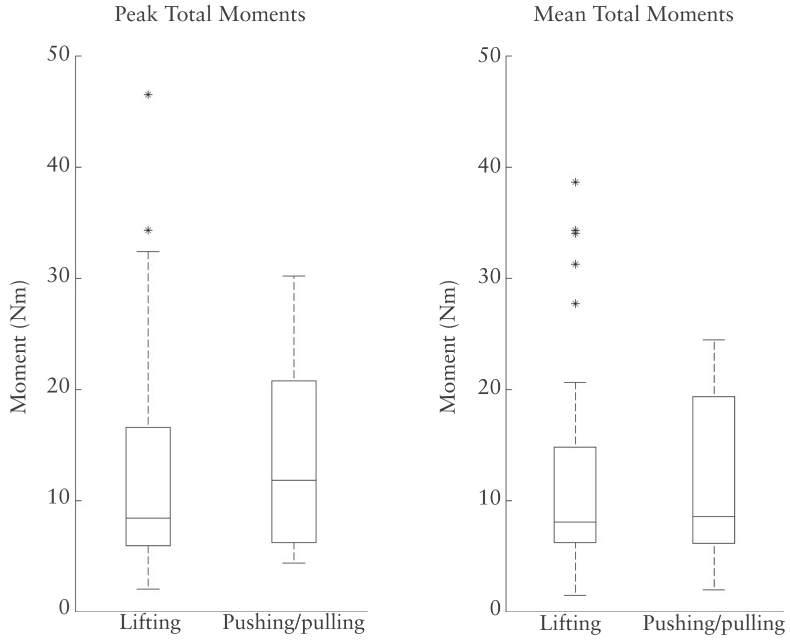

Figure 7.4 | Standard deviations across raters of all rated tasks concerning peak (left plot) and mean (right plot) moments calculated for lifting tasks only and for pushing and pulling tasks. In the figures, the middle notch represents the median standard deviation, the box presents the standard deviations of the 25th percentile to the 75th percentile, whiskers represents the 5th to 95th percentile interval and asterisks represent outliers. Values were calculated over all 50 tasks.

Despite relatively low inter-rater reliability of some postural variables, highly reliable low-back loads were found. A possible explanation is that not all postural variables contribute equally to the low-back load. For example, it is likely that the trunk flexion angle and the horizontal load distance with respect to the L5S1 segment contribute largely to the low-back moments whereas abduction of the shoulder contributes little to the low-back moment. In addition, an error in rating the shoulder angle can be compensated by a concomitant error in rating the elbow angle. This will then lead to a reliable load distance and consequent low-back load. This reasoning is supported by good inter-rater agreement for horizontal load distance of the load with respect to the L5S1 joint and of low-back moments, despite substantial errors in some of the posture variables. Furthermore, other postural variables (e.g., trunk flexion and trunk lateral flexion) do show highly reliable inter-rater reliability.

We did not compare our results to a gold standard as, regrettably, there is no gold standard in assessments of low-back load doses in field situations (Takala et al., 2010). Comparison of measurement tools described in other studies with respect to validity of outcomes is therefore difficult. However, in a lab-based validation study on the same video analysis method (Coenen et al., 2011) we found non-systematic, random errors for peak and mean low-back moments. The present study adds that between-rater differences are rather small (<10%), suggesting that the present video analysis method is a good method for low-back load assessments in field settings.

Although lab-based posture observation studies show comparable inter-observer agreement to the agreement reported here (Bao et al., 2009; Burt & Punnett, 1999), work-site postural observation methods, with and without the use of video recordings, have some drawbacks. They rely on crude categorical estimates, the magnitude of errors increases when joint angles become close to posture boundaries, outcomes heavily rely on the experience of the observer (Kociolek & Keir, 2010; Lowe, 2004; Spielholz et al., 2001), and observers seem to have difficulties to analyze more variables at once (Spielholz et al., 2001). Furthermore, agreement between raters is highly dependent on the number of categories used (Andrews et al., 2008). A postural variable categorized in a low number of categories is more likely to have a high inter-rater agreement, however, may lead to a loss of information (van Wyk et al., 2009). Eventually, large errors may result when using observations of working postures as input in biomechanical models estimating low-back load doses (de Looze et al., 1994b). Due to the reliable estimates of low-back moments and the on-line fitting of body orientations, the proposed video analysis method seems to be more appropriate to assess MMH tasks, especially when estimating low-back loads doses. In studies on comparable video coding systems, Xu et al. (2011) found, except for trunk lateral flexion, high ICCs (>0.75) for segment angles while Sullivan et al. (2002) found ICCs to be high as well for several low-back load metrics. These results are comparable to the ones reported here, however, both studies only reported on lab-tests, whereas we performed a study on field-based tasks.

Despite high inter-rater reliability and small variation among observers, relatively large errors can occur in some occasions. Such errors mainly occur in situations in which a part of the subject's body is occluded from view (e.g. when workers turn their back to the camera or when the view on the worker is, for example, occluded by another worker or by machinery). Although these substantial inter-rater differences occur in only a minor proportion of the tasks, such problems seem to be inevitable in field settings. The possible occurrence of these errors should therefore be noted when obtaining low-back load data from workers in field settings.

We used a relatively small number of raters who had substantial knowledge on kinesiology but no experience on working with low-back load assessment tools. External validity of the current video method can thus be questioned. However, our video analysis method is rather objective as it involves adjusting postures of the manikin to the posture of worker with continuous visual feedback of the manikin stick figure over the video frames. This procedure involves only minor subjective scoring, therefore, no major biases can be expected as a result of the selection of raters.

It has been reported in earlier studies that low-back loading is a risk factor for the occurrence of LBP (Marras et al., 2010; Norman et al., 1998). Both studies found significant differences in several low-back load metrics between workers with and without (risk of) LBP up to about 20%. The errors that we found between raters are substantially smaller than this percentage. Therefore, we expect only minor misclassifications in LBP risk groups due to inter rater variability using the proposed video analysis method.

CONCLUSIONS

The current study shows that the proposed video analysis method is reliable when used by different raters, which makes it applicable in epidemiological studies or ergonomic practice for low-back load dose assessment. Inter-rater reliability for low-back moments is high, while the agreement for rating of the most important segment angles is reasonable. Errors are small enough to limit the likeliness of misclassification in LBP risk groups. Although occasional substantial errors can be made when assessing MMH tasks, this study shows good overall agreement among raters.

Chapter 8

Cumulative mechanical low-back load at work is a determinant of low-back pain.

P. Coenen
I. Kingma
C.R. Boot
P.M. Bongers
J.H. van Dieën

Submitted

ABSTRACT

Reported associations of physical exposures during work (e.g. lifting, trunk flexion or rotation) and low-back pain (LBP) are rather inconsistent. Mechanical back loads (e.g., moments on the low-back) as a result of exposure to abovementioned risk factors has been suggested to be important as such loads provide a more direct relationship with tissue failure and thus LBP. Since information on the effect of such load metrics with LBP is lacking yet, we aimed to assess this effect in a prospective study.

Of 1131 workers, categorized in 19 task groups, LBP was prospectively assessed over three years. Video and hand force recordings of four to five workers per task group (93 in total) were used to estimate mechanical low-back loads (peak load and three cumulative load metrics, i.e., linear weighted load, squared weighted load and load weighted to the 10th power) during manual materials handling (MMH) tasks using a video analysis method. These data were combined with static mechanical load estimates based on structured observation of non-MMH tasks. Associations of mechanical loads and LBP were tested using generalized estimating equations.

Significant effects on LBP were found for cumulative low-back moments (linear and squared weighted; both $p<0.01$ and odds ratios of 3.01 and 3.50 respectively) but not for peak and cumulative moments weighted to the tenth power.

Results of this first prospective study on the effect of mechanical low-back load on LBP support a LBP etiology model of cumulative loads, potentially due to accumulation of micro-damage or fatigue.

INTRODUCTION

Epidemiological studies have contributed to our understanding of the etiology of low-back pain (LBP). According to these studies, LBP is associated with personal risk factors (e.g. age, smoking habits, physical capacity and body weight (Hamberg-van Reenen et al., 2007)), psychosocial risk factors (e.g. stress, social support and job satisfaction (Hartvigsen et al., 2004)), and physical risk factors. Of these physical risk factors, twisting, bending, lifting and whole body vibrations are most frequently reported (Griffith et al., 2012; Lötters et al., 2003). However, it has also been argued that evidence concerning these physical risk factors of LBP is weak, possibly as a result of the use of measurement tools with low accuracy (Bakker et al., 2009). Specifically, measuring physical risk factors for LBP often relies on self-reports or observations that, although proven to be valid and reliable, can have weak associations with LBP (Griffith et al., 2012). Moreover, objective field-based measurements often lack a clear description of all dimensions of the exposure to the risk factors i.e. duration, frequency and magnitude (Takala et al., 2010). It can be argued that mechanical low-back load metrics (e.g., spinal compression forces or moments at the low-back) provide more information than low-back exposure measures (e.g., the number of lifts or time spend in a flexed trunk position). One reason for this is that exposure metrics do not always have consistent relations with load metrics. For example, the mass lifted is a poor predictor of low-back moments (Hoozemans et al., 2001). Furthermore, different exposures affect the same mechanical load. Therefore, load metrics can be expected to be more strongly associated with LBP for which some empirical support has already been provided (Coenen et al., 2013b; Norman et al., 1998).

Several models for the causal chain of LBP etiology have been proposed, all assuming that tissue failure due to mechanical load on the back, as a result of abovementioned variables, is a cause of LBP (Chaffin, 2009; Marras, 2012). In general, two pathways for the occurrence of tissue failure can be considered: LBP may result from instantaneous tissue failure due to peak loads on the low-back, or from cumulative loads. Cumulative loads could cause LBP, for instance through accumulation of micro-damage, or through impaired coordination due to respiratory (Brereton & McGill, 1999; Janssens et al., 2010) or neuromuscular (Sparto et al., 1997; van Dieën et al., 1998) fatigue. The predictive value of a variety of parameters of low-back loading for the risk of LBP has been assessed, showing that both cumulative (Coenen et al., 2013b; Kumar, 1990; Norman et al., 1998) and peak spinal loads (Neumann et al., 2001; Norman et al., 1998) are associated with the LBP prevalence. These findings militate in favor of both of the two abovementioned causal models. However, results are based on retrospective studies or on prospective studies using exposure risk factors rather than low-back load metrics. Such studies can be of paramount importance to gain more insight into the etiology of LBP.

Although there is currently no gold standard for obtaining mechanical low-back load metrics from workers in a field setting (Takala et al., 2010), video based coding methods

(Xu et al., 2011) that assess postural data, which are subsequently used in biomechanical models estimating mechanical low-back load (Coenen et al., 2011; Sutherland et al., 2008), are suitable for this purpose. These methods allow for obtaining accurate mechanical low-back load estimates in field settings without interfering with the worker's tasks. The video-based method that is used in the current study has been validated against a lab-based gold standard (Coenen et al., 2011) and inter-rater reliability of this method has been assessed in the field, showing inaccuracies of approximately 10% of maximum loads (Coenen et al., 2013a). The objective of the present study was to assess the effect of peak and cumulative low-back load metrics on LBP in a prospective cohort study using this video analysis method. To the best of our knowledge, there are currently no data available from prospective studies assessing mechanical low-back load in work site situations. Moreover, it is not yet clear how, in calculating cumulative loads, repetition of loading should be weighted relative to load intensity. As suggested before (Brinckmann et al., 1988; Rapillard et al., 2006), it is likely that the magnitude of peak loads has more impact on the risk of failure than the number of times a load occurs. Therefore, several weightings of these peak loads in the calculation of cumulative loads have been proposed, including raising the loads to a certain power, e.g., squared (Coenen et al., 2012; Seidler et al., 2009), fourth order (Jäger et al., 2000) and even tenth order weighting (Coenen et al., 2012). A higher order weighting reflects a higher importance of load intensity compared to the number of loading cycles. In the current study, the effect of several weightings for cumulative loading will be tested, i.e., linear weighting, squared weighting and tenth order weighting, where the latter two are expected to have a higher predictive value for LBP.

METHODS

Population and data collection

Data used in this study were collected as part of the Study on Musculoskeletal disorders, Absenteeism and Health (SMASH) that aimed to assess risk factors of musculoskeletal disorders among Dutch workers (Ariëns et al., 2001; Hoogendoorn et al., 2000). Briefly, workers from various industrial and service branches, for example, the metal, chemical, pharmaceutical, food and wood construction industry; waste processing, insurance and distribution companies were studied. The study consisted of a baseline measurement, assessing low-back load at the workplace and potential confounders, and a baseline and three year follow-up assessment of musculoskeletal symptoms. Ethical approval for this study was obtained from the *Netherlands* Organization for Applied Scientific Research (TNO) ethics committee. Any identifiable subjects have provided their signed consent to publication and participants gave informed consent before taking part in the study.

Personal factors such as age and gender were assessed using self-administered questionnaires. Furthermore, a Dutch version of the Karasek's Job Content Questionnaire for psychosocial work characteristics was used to assess job demands, decision authority, co-worker support and supervisor support (Karasek, 1985). Exercise behavior during leisure time was assessed with the Leisure Time Exercise Questionnaire (Godin et al., 1986). Driving a vehicle during work and during leisure time, and physical exposure during leisure time were assessed with the Loquest questionnaire (Hildebrandt & Douwes, 1991). The occurrence of LBP was assessed using a Dutch version of the Nordic Questionnaire (Kuorinka et al., 1987). LBP at baseline and during the three consecutive years of follow-up was defined when subjects reported regular or prolonged LBP in the 12 months prior to filling out one of the questionnaires.

At baseline, 1990 of the 2048 workers who were invited for the study participated and 1802 (91%) of these workers completed the baseline questionnaires. Of these workers, LBP data in at least one of the years of follow-up were available for 1131 workers. All these workers filled in the LBP questionnaires at baseline and during the first year of the follow-up, while 1004 and 994 workers filled in the LBP questionnaires during the second and third year of the follow-up respectively. These workers were classified into 19 task groups, based on physical exposure. For 371 workers, approximately 25% of all workers within each task group, 5-15 minutes of video recordings at the workplace were taken at four randomly chosen moments during the course of one day. Furthermore, external forces at the hands during these periods were measured using force transducers (during pushing and pulling) or weighting scales measuring mass of the external load (for lifting tasks). Videos were observed during which manual material handling tasks (MMH tasks; i.e. lifting, pushing and pulling tasks) were identified, yielding a total of 12,924 tasks. In the current study, only task groups with at least four observed workers were included. From each task group, four or if available five workers were analyzed to assess mechanical low-back load. As a result, 4872 MMH tasks of a total of 93 workers were selected for the current study (Table 8.1). On average there were 58±103 MMH tasks per worker, ranging from 0 to 534 tasks. Video recordings of the 4872 MMH tasks were used for the assessment of mechanical low-back load using video analysis as described in the next paragraph.

Table 8.1 | Descriptive statistics of the entire cohort (first column), the group of workers from whom video recordings were available (second column) and the group of workers mechanical loads were calculated from in the current study (third column). Number of subjects, age, gender, LBP during in one of the four questionnaires, number of MMH tasks

	Total	Recorded	Analyzed
Number of workers (n)	1131	371	93
Age (years)	36 (9)	36 (9)	36 (9)
Males (n(%))	800 (71%)	216 (68%)	61 (66%)
LBP (n(%))	600 (53%)	199 (54%)	48 (52%)

Assessment of mechanical low-back load

Ten raters were recruited among students of the Amsterdam School of Health Professions and the Faculty of Human Movement Sciences of the VU University, Amsterdam. After participating in an extensive learning and practice session in which they were familiarized with the software, each rater analyzed videos of a selection of tasks. Raters analyzed videos independently from each other and were asked to rate as many tasks as possible, including tasks that were not recorded optimally (e.g., due to partial occlusion of the view or when the task was not recorded from a sagittal plane view). Furthermore, raters were blinded from the fact whether they rated a worker that either had or had not reported LBP.

Videos of all 4872 MMH tasks were rated, using an earlier developed video-analysis method that has been described extensively before, and was tested on validity and inter-observer reliability (Coenen et al., 2011; Coenen et al., 2013a; Xu et al., 2011). Begin and end frames of the tasks were selected from the video and two intermediate frames were automatically selected to obtain four video frames. On each video frame, a manikin was fitted consisting of nine segments (right foot, lower leg and upper leg; pelvis, trunk/head, two upper arms, two forearms/hands). This manikin allows for semi three-dimensional analysis of movements (ankle flexion/extension, knee flexion/extension, hip flexion/extension, trunk flexion/extension, trunk rotation, trunk lateral flexion, shoulder flexion/extension, shoulder abduction and elbow flexion/extension). Furthermore, the manikin can be scaled, rotated around its longitudinal axis (axial rotation) and translated horizontally and vertically along the video frame, which allows the rater to make an optimal fit of the manikin to the video frame. Subsequently, interpolations of the segment angles over the four key frames were executed to estimate workers' body kinematics (Xu et al., 2010). Based on total body mass and stature, individual segment masses and lengths, positions of the center of mass and inertia tensors were estimated (Zatsiorsky, 2002). A top-down 3D inverse dynamics calculation using hand forces, segment kinematics (obtained from the interpolated manikin postures) and anthropometrics was performed to assess resultant moments at the level of the L5S1 joint. For each MMH task, peak moments were calculated. Workers that did not perform any MMH tasks during the collection of video were assigned a peak load as obtained from an earlier calculation of mechanical low-back load (Coenen et al., 2013b). In this latter study, moments were calculated based on static postures while these postures were based on continuous structured visual observation of all video material of each worker. In these observations, postures were categorized into four categories of trunk flexion, two categories of trunk rotation and four categories of arm elevation.

For cumulative load, a time series for the complete video recordings of the 93 subjects was constructed in which the abovementioned estimation of low-back moments based on observations for non MMH tasks was added to moment time series of all analyzed MMH tasks of the subject. Cumulative moments were then estimated by calculating the area under the moment curve while outcomes were extrapolated to an entire work week (based on the length of the observation and the working hours per week). Peak load was defined as the maximum peak in the complete time series. Three kinds of cumulative moments were calculated: area under the curve, area under the squared curve and area under the curve to the 10th power. Of the four variables (one peak load variable and three cumulative loads), group-based loads (in which average group load estimates are assigned to all members within each task group) were calculated and were used as potential risk factors for LBP in further statistical analyses. To facilitate the interpretation of the ORs presented in the current study, the metrics were divided by 1·102, 1·105, 1·107, 1·1010 for peak moments, non-weighted cumulative moments, squared weighted cumulative moments and moments weighted to the tenth power respectively. Calculations were performed using custom developed Matlab software (version 7.7.0).

Statistical analyses

All analyses were executed for the four load metrics separately. The crude effect of the mechanical low-back loads on LBP were assessed using univariate Generalized Estimating Equations (GEE) with the load (as continuous variables) being the independent variable and LBP (dichotomous outcome of the four measurements –baseline and three years of follow-up-) being the dependent variable. Furthermore, the contribution of a number of potential confounders was explored with multivariate GEE using a forward stepwise selection procedure with the load being the independent variable and LBP being the dependent variable, as described above. Only confounders that led to a change of >10% in the beta depicting the effect of the mechanical load on LBP were included in the model (Twisk, 2006). The following potential confounders were considered, based on previous studies (Coenen et al., 2013b; Hoogendoorn et al., 2000): age, gender, smoking habits, body mass index, physical activity in leisure time, quantitative job demands, decision authority, skill discretion, supervisor support, co-worker support, work security, driving a vehicle during work and leisure time, sitting at work, flexion/rotation of the trunk during leisure time, moving heavy loads during leisure time. In the final four models, the effects of the potential risk factors adjusted for all potential confounders were assessed using multivariate GEE. In all GEE analyses an exchangeable correlation matrix was used. Only for univariate models, quasi likelihood under the independence model criterion (QIC) were calculated depicting the goodness of fit of the models; a lower QIC values was interpreted as a better fit (Pan, 2001). Odds ratios (ORs) and 95% confidence intervals, and corresponding p-levels were estimated for the mechanical low-back loads. P-values <0.05 were considered statistically significant.

To test the robustness of the current selection of 5 workers per task group, we combined our data with 2,339 MMH tasks (74 workers) that had been additionally analyzed (but were not uniformly distributed over the 19 task groups) for other purposes, and we performed 25 random drawings of 5 workers per task group. For each drawing, the effect of the four mechanical load measures on LBP was assessed univariately as described above; p-values of these effects were calculated. All statistical analyses were performed using SPSS (version 20).

RESULTS

Out of all identified MMH tasks, 4,168 (86%) tasks were analyzed. The remaining selected tasks could not be analyzed due to unsatisfactory low quality of the video material (e.g., partial occlusion of the view). On average 52 (90) tasks per subject were analyzed, with an average external force measured at the hands of 72 (60)N. Of these tasks, 3,566 (86%) were lifting tasks, 450 (11%) were pushing tasks and 152 (3%) were pulling tasks.

Linear and squared weighted cumulative load had a significant on LBP, univariately (both p<0.01), as well as when adjusted for confounders (both p<0.01; Table 8.2). Cumulative loads weighted to the tenth power and peak moments did not have a significant effect on LBP, neither when effects were calculated univariately (p=0.70 and p=0.12 respectively), nor when adjusted for confounders (p=0.74 and p=0.73 respectively). Regarding the goodness of fit, a comparable pattern could be found, since linear and squared weighted cumulative loads led to better fits compared to cumulative loads weighted to the tenth power and peak moments. Furthermore, squared cumulative loads led to a slightly better fit than linear weighted cumulative loads. In order to facilitate interpretation of these data, ORs adjusted for confounders for linear and squared weighted cumulative loads (model 3 in Table 8.2) were used to calculate ORs corresponding with a difference in mechanical load of the task groups with the highest mechanical load compared to the group with the lowest mechanical load. This calculation provided ORs of 3.01 and 3.50 for the two metrics respectively.

The robustness analysis of the four mechanical load metrics showed that the estimate of the linear and squared weighting of cumulative loads were robust as comparable p-levels (all <0.01) were shown for all drawings (Figure 8.1). The effect of peak loads was moderately robust leading to univariate significant effects (p<0.05) in six out of the 25 drawings. However, the tenth power weighting of cumulative load was not robust, with a significant effect in one out of the 25 drawings and p-values ranging from <0.01 to 0.87.

Table 8.2 | Effects of mechanical low-back loads on LBP presented the four mechanical low-back load risk factors. For each risk factor, three models are presented: A crude model (Model 1), a model corrected for statistically significant confounders (Model 2) and a model corrected for all potential confounders (Model 3). For univariate models, QIC scores are shown.

		LBP	No LBP	Model1			Model2		Model3	
Moment	Expressed	Mean (St. Dev.)	Mean (St. Dev.)	OR (95% CI)	p	QIC	OR (95% CI)	p	OR (95% CI)	P
Peak	$/10^2$	1.45(0.61)	1.38(0.57)	1.14(0.97-1.35)	0.12	5219.15	1.05(0.81-1.37)	0.69	1.05(0.81-1.36)	0.73
Cumulative	$/10^5$	23.3(12.5)	21.4(6.7)	2.12(1.41-3.16)	<0.01	5197.78	2.01(1.22-3.31)	0.01	2.15 (1.29-3.58)	<0.01
Cumulative^2	$/10^7$	12.2(12.7)	10.1(6.2)	2.16(1.47-3.17)	<0.01	5194.34	2.12(1.39-3.22)	<0.01	2.19(1.35-3.56)	<0.01
Cumulative^{10}	$/10^{10}$	7.61(28.61)	6.60(27.64)	1.21(0.90-1.17)	0.70	5221.55	1.02(0.88-1.18)	0.77	1.03(0.89-1.18)	0.74

Expressed = Defines the units in which the ORs are expressed.

OR (95% CI) = Odds ratio and 95% confidence interval depicting the effect of the mechanical load on LBP

p = level of significance depicting the effect of the mechanical load on LBP

QIC = Quasi likelihood under independence model criterion, depicting the goodness of fit of the univariate mechanical loads.

Model1: Univariate model

Model2: Model adjusted for the following statistically significant confounders:

Peak: Gender, quantitative job demands, work security, physical activity in leisure time, smoking habits, driving a vehicle during work, rotation of the trunk during leisure time, co-worker support, decision authority, supervisor support, moving heavy loads during leisure time, driving a vehicle during leisure time, sitting at work

Cumulative: Gender, co-worker support, driving a vehicle during leisure time, sitting at work

Cumulative2: Gender, smoking habits

Cumulative10: Gender, smoking habits, physical activity in leisure time, quantitative job demands, skill discretion, supervisor support, co-worker support,driving a vehicle during leisure time, sitting at work

Model3: Model adjusted for all potential confounders

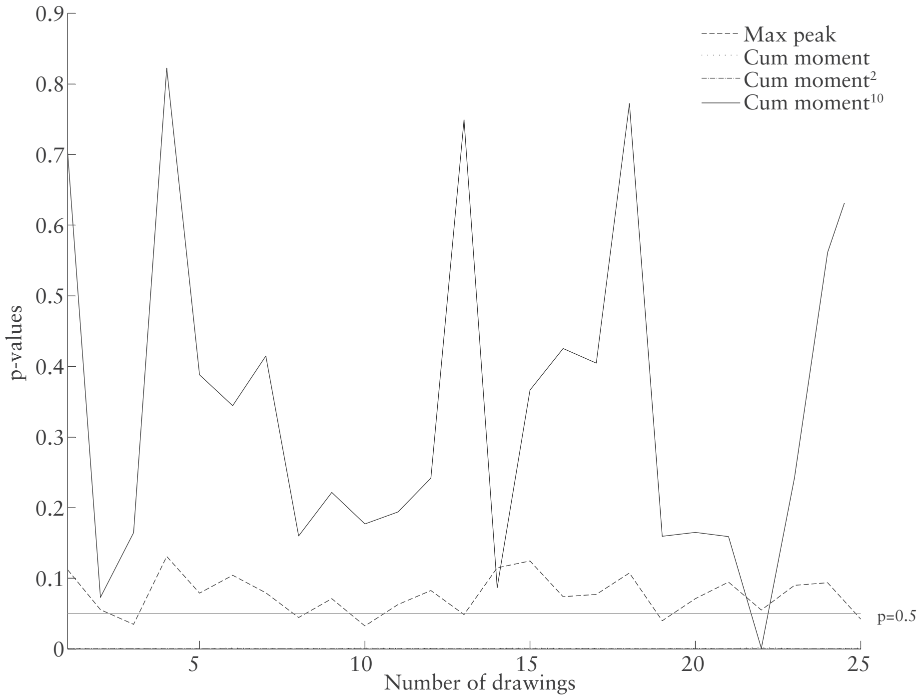

Figure 8.1 | Univariate effect of the four mechanical low-back loads on LBP during 25 random drawings of 5 subjects per task group. Note that all drawings for linear and squared weighted loads have rather small values (all<0.01). A level of significance $p=0.05$ is presented with the grey line.

DISCUSSION

The aim of the present study was to assess the effect of mechanical low-back load metrics on LBP in a prospective cohort study. It can be concluded that cumulative loads are strong predictors of LBP. These findings are in line with the model of LBP etiology due to accumulation of micro-damage and with previous studies showing associations of cumulative mechanical back loads with LBP (Coenen et al., 2013b; Kumar, 1990; Neumann et al., 2001; Norman et al., 1998). Despite the fact that we showed previously that in-vitro failure of spine segments during repeated loading at a constant load levels is best predicted when using a tenth power of load level (Coenen et al., 2012), this metric did not have a significant effect on LBP in our data. The higher the order of the weighting, the larger the contribution of load magnitude to the risk estimate compared to frequency of loading. The latter study was based on a mechanical load protocol applied on in-vitro material. On the one hand, in-vitro material lacks the potential to repair micro-damage, which would cause an overestimation of the importance of the loading frequency. On the other hand, in-vitro testing does not take into account that the risk of low-back injury may increase when respiratory or neuromuscular fatigue causes impaired coordination (Brereton & McGill, 1999; Janssens et al., 2010; Sparto et al., 1997; van Dieën et al., 1998). This leads to an underestimation of the importance of the temporal characteristics of loading. As we show here that squared weighting load has, but load weighting to the tenth power does not have an effect on LBP, the latter characteristic of in-vivo conditions may play an important role here. However, this reasoning may be premature, since the lack of predictive value of the tenth power weighting might also be a result of the fact that the metric is highly affected by inaccuracies in the measurements or actual variation in the work pattern. This can also be deduced from the non-robust nature of the effect of this metric on LBP (Figure 8.1).

As has been suggested before, it is likely that the magnitude of peak loads has more impact on the risk of failure than the number of times a load occurs (Brinckmann et al., 1988; Rapillard et al., 2006). This led us to predict that, in the calculation of cumulative loads, weighted peak loads would be more predictive of LBP than non-weighted peaks. Because squared cumulative loads tended to have a slightly better fit than linear weighted cumulative load, the use of such weighting is recommended for future studies. It should be kept in mind that the design of the present study, with group-based averaging of load metrics and a long follow-up period for the assessment of LBP does not allow any inference on the importance of occasional peak loads leading to acute injury and pain.
In the present study, peak moments did not have an effect on LBP. Although, this lacking effect was moderately robust leading to significant effects in some cases of the repeated drawings univariately, effects were highly non-significant when adjusted for confounders. Therefore, our findings provide stronger support for the cumulative load induced tissue failure model than for the peak load induced tissue failure model. A difference in mechanical load corresponding with a difference of the task groups with the highest mechanical load compared to the group with the lowest mechanical load can be interpreted with ORs of 3.01 and 3.50 for linear and squared cumulative loading respectively. These values suggest substantial risks of LBP in the group of workers with the highest mechanical loads (mainly road workers with high external forces). Prevention of LBP should therefore be targeted on such tasks. Moreover, these ORs are higher than pooled ORs reported in earlier studies for exposures metrics (Griffith et al., 2012). Therefore, the present results are in line with earlier studies suggesting higher associations for mechanical loads as compared to exposure metrics (Coenen et al., 2013b; Norman et al., 1998).

The strength of the present study is that the results are based on a large prospective cohort study and that, for the MMH tasks, low-back loading was assessed more accurately than in epidemiologic studies performed thus far. Furthermore, the current study is based on an assessment of mechanical load that has been proven to be valid (Coenen et al., 2011) as well as reliable among raters in field settings (Coenen et al., 2013a). However, the video analysis method contains some limitations. Only MMH tasks were assessed with the current method, while moments during the remaining part of the video recording were estimated, based on static postures obtained from postural observation categories (Coenen et al., 2013b). This was performed under the assumption that the highest mechanical loads derive from MMH tasks. However, from the current data, it cannot be ruled out that a source of bias is introduced due to this procedure. Therefore, when future techniques allow for continuous measurement of mechanical loads, improvements in the predictive value of mechanical loads can be expected. Furthermore, the video analysis used may yield occasional large errors, e.g., due inherent inaccuracies in manikin fitting (that are amplified in tasks of very short duration). These inaccuracies can originate from occlusion of the view or in highly non-sagittal plane recordings. However, these errors were shown to have a random character (Coenen et al., 2011; Coenen et al., 2013a). Furthermore, as multiple MMH tasks per subject were assessed and as group-based values were calculated in a pool of workers, these random errors are likely to be diminished. However, as has been indicated above, such errors are amplified when using higher order weighting in cumulative load calculations.

In this study, only a limited number of workers were assessed. Mechanical load data were assessed for four or five subjects per task group, introducing the possibility of selection bias, as the rest of the 371 workers, from whom observational data were available, were not analyzed. Such group-based approaches have been adopted before (Ariëns et al., 2001; Hoogendoorn et al., 2000) and have proven to be successful for the assessment of work load in several occupational tasks. Such group-based estimates of work load have been shown to be more reliable than individual estimates (Hoozemans et al., 2001; Paquet et al., 2005), leading to higher predictive values (Jansen & Burdorf, 2003), as individual random errors are reduced. These studies furthermore illustrate that, with an increase in the number of workers sampled, the work load estimate improves less when continuing to add more subjects, which suggests that measuring too many subjects when calculating group-based work load is inefficient. In a simulation study, it was furthermore shown that a total of five workers per task group should be sufficient to obtain significant risk associations for LBP (unpublished data). Furthermore, from the data presented in Figure 8.1, it can be concluded that, at least for the two significantly predictive cumulative load metrics, ORs and p-values comparable to the ones we have reported, are found when varying the selection of workers for low-back load assessment. The current selection of workers is therefore likely to be representative. Moreover, the selection of workers for whom low-back load was measured was highly comparable to the entire group of workers with respect to age, gender and prevalence of LBP (Table 8.1). Therefore, selection bias is not likely to have had a strong impact in the present study. A final source of bias might have emerged from the fact that workers were video-taped at four randomly chosen occasions of the work day for a finite amount of time rather than during the whole work day. Distributing these four occasions over several days might have resulted in a more precise work load estimates, as work load will most likely vary more between days than within days (Mathiassen et al., 2003; Paquet et al., 2005). This issue was addressed by measuring several workers at different days in each task group, to obtain more precise estimates of the work load within groups (Mathiassen et al., 2003; Paquet et al., 2005). The appropriateness of our measurement strategy was furthermore supported by showing small within group variability of observation-based exposure estimates in a previous study on the same population (Ariëns et al., 2001).

From this first prospective study on the effect of mechanical low-back load on LBP, it can be concluded that cumulative low-back loads are predictive for the occurrence of LBP. However, a significant effect was not found for peak loads. Therefore, these findings provide stronger support for a model of LBP etiology due to cumulative loads than for a model based on single peak loads. Information obtained from this study can teach us on the biomechanical etiology of LBP. Such information can be of vital importance for policymakers and ergonomic practitioners when designing LBP prevention programs. Based on the current results, such programs should focus on reducing cumulative low-back loads.

Chapter 9

Epilogue

THESIS SUMMARY

Physical work load is considered an important risk factor for low-back pain (LBP). However, in **Chapter 1** it is also argued that reported associations between physical workload and LBP are rather inconsistent. This inconsistency is a barrier for the understanding of the etiology of LBP. One reason for this lack of knowledge may be inadequate quantifications of mechanical loads in work situations. It was argued that, in most models for LBP etiology, these mechanical loads (i.e., loads at the level of the lower back, for example, low-back moments as indicators for mechanical load, or compression or shear forces on the lumbar spine) are at the end of the causal chain, thereby providing a rather direct relationship with spinal damage. Different exposures (e.g., lifting, twisting and bending) affect the same mechanical load, so that mechanical load can be considered a 'final common pathway' to spine injury. Therefore, obtaining mechanical load metrics in prospective studies seems to be important when striving to obtain more understanding of the etiology of LBP. However, such studies are lacking, probably because of the absence of occupational assessment tools that are easily applicable in field situations. Furthermore, also other measurement issues that affect the outcome of such risk associations are insufficiently understood. Therefore, four aims were addressed in the current thesis. The main findings regarding these aims will be discussed in the following sub-paragraphs. Subsequently, general conclusions will be drawn based on this thesis, and future directions for research and ergonomic practice will be discussed.

The predictive value for LBP of mechanical loads as compared to (subjective) exposure estimates

As mechanical low-back loads have been assumed to be more predictive for LBP than exposures (i.e., obtained from self-reports or from observations), our initial goal was to test this hypothesis in a prospective cohort study. Data presented in **Chapter 2** show that although trained observers were able to predict neck and shoulder pain, they could not predict LBP well. This can be explained by the fact that compared to neck and shoulder load, low-back load depends on a larger number of task variables (i.e. trunk posture, arm posture, load magnitude and load distance) that seem to be difficult to assess subjectively. The finding that risk estimates of LBP are not significantly associated with LBP prevalence questions the accuracy of these subjective risk estimates and advocates for the use of precise measurements rather than estimates.

From the findings reported in **Chapter 3**, we can conclude that cumulative mechanical low-back load, as obtained from calculations of mechanical back load based on posture observation, is a significant predictor of LBP. Moreover, it was shown that this mechanical load metric has a stronger association with LBP than earlier reported exposure risk factors (i.e., time in a flexed position and number of lifts during a working day). These findings support our hypothesis that a mechanical low-back load measure provides a stronger association with LBP than exposure measures. Based on these results it seems justified to develop more precise methods to assess mechanical loads at the workplace. Furthermore, mechanical load variables should be considered in future epidemiological studies to obtain more information on LBP etiology.

The effects of methodological issues on the predictive value of low-back loads for LBP

As a second step towards a better understanding of the LBP etiology we assessed the impact of some methodological issues that are of importance in epidemiological studies on the matter. In cumulative mechanical loads, the (peak) low-back load magnitude of a given work task is often multiplied by the number of load cycles of that particular task, while these multiplications of all tasks during a work shift are summed. However, it has been argued that high forces have more impact on the increase in failure risk than a high number of cycles. **Chapter 4** confirms this hypothesis by a re-analysis of in-vitro mechanical loading to failure data. This analysis showed that weighting compression forces and number of load cycles with exponents of 2 and 0.2, respectively, provides the best prediction of in vitro lumbar spine failure following cumulative loading. This non-linear load-failure association has implications for future studies assessing the effect of cumulative low-back loading for investigation of LBP etiology.

Another methodological issue that we have assessed is the effect of group size in group-based measurement protocols on the statistical power of eventual risk associations (**Chapter 5**). In group-based measurement protocols, workers are grouped according to common characteristics, such as their work tasks. Group-averaged exposure estimates are assigned to all workers in the group on the basis of data measured in a subgroup only, while outcome data (i.e., LBP) are assessed for all workers. Such protocols are often used in epidemiological studies on physical risk factors of LBP. Our results show that the power in such a group-based study depends more on the total number of workers included in the study (using individual outcome data on LBP) than on the size of the subpopulation from which exposures are obtained. Effectively, in order to reach a power of more than 0.80 at a p-level of <0.05, in general, at least 30 workers have to be included in each task group, with exposure measurements of at least 5 of these workers. When exposure was observed from fewer than 5 workers, the odds ratio (OR) of the exposure-outcome relationship was negatively biased. Therefore, findings suggest that although exposure of sufficient workers (≥ 5) should be assessed in order to avoid bias of the OR, it seems to be more efficient to assess LBP from a larger number of workers (≥ 30 per task group).

The applicability of video based quantification of mechanical low-back load in a field situation

Measuring mechanical low-back loads in field settings is a tempting task, as current measurement methods often interfere with the employer's work or only crude metrics are used. Therefore, a video analysis method for the assessment of mechanical low-back loads in the field was developed, based on earlier work. This analysis method can potentially be used in ergonomics practice and in future epidemiological studies as video material can be collected without interfering with the worker's tasks. **Chapter 6** describes a study in which this video analysis method for the assessment of low-back moments during occupational lifting was validated by performing a comparison with a laboratory reference method. No overall differences in peak and mean moments between the reference method and the video analysis methods were found and intra-class correlation coefficients (ICCs) revealed a strong correspondence of the video analysis method and the reference method. In **Chapter 7**, the inter-rater reliability of the video analysis method was tested on video material that had been recorded in field settings. Results from this chapter show excellent agreement among raters (ICC >0.9), while inter-rater variation was relatively low (<10 Nm), for low-back moment estimates of peak and mean moments. However, occasional substantial errors were shown during the assessment of manual material handling (MMH) tasks. These errors appeared to result from amplification of random posture rating errors in tasks of short duration, especially in MMH tasks that are difficult to rate because they were filmed from a non-sagittal view. Despite these errors, it can be concluded that the current video analysis method is valid as well as reliable. The latter is also the case when assessing occupational field tasks.

The etiology of LBP using mechanical load metrics

In the final study described in this thesis (**Chapter 8**), the video analysis method was applied to a large prospective cohort. Mechanical loads were assessed and their association with LBP was estimated. This study shows that cumulative mechanical low-back loads predict LBP. However, the required exponential weighting of force level appeared to be lower than predicted from the in-vitro data analyzed in Chapter 4. Nevertheless, these findings are in favor of the mechanism for the etiology of LBP described in Chapter 1, where cumulative loads play an important role in the cause of LBP, potentially as a result of accumulation of micro-damage, and/or through impaired coordination due to fatigue. As peak loads are not significantly associated with LBP, instantaneous tissue failure due to peak loads on the spine is a less probable cause of LBP based on the current data. However, the latter mechanism for etiology cannot be ruled out, especially as our data suggest that a weighting of load magnitude with a power larger than 1 in calculations of cumulative loads provided a better fit to our data.

GENERAL DISCUSSION

A number of general conclusions can be drawn from the current thesis. First of all, regarding the predictive value for LBP, a clear advantage was shown for the use of mechanical load metrics over exposures obtained by subjective assessments or structured posture observation. This is in line with data from a cross-sectional study (Norman et al., 1998) and with several models arguing that mechanical loads (i.e., loads at the level of the lower back, such as compression forces on the lumbar spine or low-back moments) are at the end of the causal chain and thus provide a more direct relationship with spinal failure and consequently with LBP (Chaffin, 2009; Marras, 2012). This direct relation stems from the fact that these mechanical loads can provide information on duration, frequency and intensity of multiple exposures. Quantification of exposures (i.e., number of lifts or time working in an awkward posture) is not directly related to the quantification of mechanical load variables. Furthermore, mechanical loads also take other mediating factors into account such as psychosocial factors, personal factors and work-related factors (as discussed in Chapter 1). Because of the arguments above, in the present thesis, mechanical loads were considered in order to obtain more information on LBP etiology. In this section the most important sources of error in quantifying low-back load with the methods used in this thesis, and their implications, will be discussed.

The use of posture observations in biomechanical models

Mechanical loads can be obtained by combining information from measured hand forces and structured posture observations in a biomechanical model, as often used in epidemiological studies. Such mechanical loads are predictive for LBP, as has been described in Chapter 3. However, this chapter describes only a first attempt to quantify low-back mechanical load in a prospective study. It has been shown before that using observational data as input for a biomechanical calculation, can lead to large inaccuracies (de Looze et al., 1994b). These inaccuracies can be illustrated by some simple examples based on data of the study described in Chapter 3. In these examples, a static procedure is used, estimating low-back moments from the moments caused by the gravitational force on the upper body with respect to the low-back and of the moments caused by the external force on the hands with respect to the low-back. Let us consider two causes of errors in back load estimates based on the observation of MMH tasks: inaccuracy due to crude categorization of the trunk flexion angle and misclassification of a MMH task. Consider a MMH task that is rated by an observer as being performed in a trunk flexion category ranging from 30 to 60°. When comparing two lifting tasks in which a 15 kg load is lifted with the arms downward and the trunk in the extremes within this category (30 or 60° flexion), moment arms of the upper body and the external force on the hands can differ considerably between these extremes. With 30° trunk flexion, the moment arms of the upper body and de hands are about 20 cm and 30 cm, respectively. However, during 60°

trunk flexion, these values increase to approximately 35 cm and 50 cm, respectively. When performing a static calculation of the low-back moment in these two situations, moments are estimated to be about 125 Nm and 215 Nm, for the 30° and 60° trunk flexion angle respectively (Table 9.1).

Another type of inaccuracy stems from errors in classifying the type of MMH task. Therefore, as a second example we consider a lifting task in which a 25kg load (equivalent to an external force measured at the hands of approximately 250N) is lifted, with the arms downward and the trunk in 30° of flexion. This force, applied at the hands in combination with the gravitational force of the upper body can contribute to a moment at the low-back of 155Nm. However, when this lifting task is incorrectly classified as being a pushing MMH task, the direction of the force vector representing the external force at the hands rotates over 90°. This can lead to a corresponding moment arm that is rather small and can even be in opposite direction relative to the moment arm corresponding to the upper body gravitational force. The moment at the lower back due to these two tasks can therefore differ considerably between these tasks, being about 155Nm and 55Nm for a lifting and a pushing task, respectively (Table 9.1).

Measuring low-back load using a video analysis method

When combining the results from Chapters 6 and 7 on the validity and reliability of our video analysis method with the considerations in the previous paragraph, it becomes clear that the video analysis method is more accurate than the method of static back load estimation based on observational data as used in Chapter 3. The video model has been shown to be applicable in the field, thereby not interfering with the worker's tasks, while measuring in laboratory settings can lead to measuring unrealistic work situations (Faber et al., 2011). Furthermore, moments were obtained from a dynamical analysis, taking not only the contribution of gravitation to the moments, but also the angular and linear acceleration of segments, into account. Our data show that this led to an improvement of the accuracy (Chapter 6) which is in line with earlier studies showing an underestimation of approximately ten percent when ignoring movement dynamics (van Dieën et al., 2010). Furthermore, several studies have reported on the problem of assessing movement outside the sagittal plane due to projection biases (e.g., Kingma et al., 1998). With the current model we decreased this source of error since we allowed for axial rotation of the manikin and quasi-three-dimensional movements (i.e., trunk rotation, trunk lateral flexion and arm abduction). The validity of this approach was supported by the fact that errors were not larger in asymmetric MMH tasks compared to symmetric MMH tasks (Chapter 6). Finally, with the present method we tried to overcome the error made by crude categorization of segment orientations, which can lead to relatively large errors, as has been shown in literature (de Looze et al., 1994b; van Wyk et al., 2009) and in Table 9.1 of this epilogue. Chapters 6 & 7 show that the video method used is both valid when compared to a laboratory gold standard and reliable among raters when used in a field setting.

Table 9.1 | Numerical example showing the consequence of inaccuracies due to the use of crude observational categories of the trunk flexion angle (upper part) and of misclassification of the type of MMH (lower part). Low-back moments were calculated by summing the moment caused by the gravitational force on the upper body with respect to the low-back and the moments caused by the external force on the hands with respect to the low-back (static procedure). The mass of the upper body is assumed to be 40kg and gravitational acceleration was estimated at 10m/s^{2}.

Inaccuracy in observation of trunk angle				
	Lifting 15 kg (30° trunk flexion)		Lifting 15 kg (60° trunk flexion)	
	Force F=m·a	Moment M=d·F	Force F=m·a	Moment M=d·F
Gravitational force upper body	40·10=400N	400·0.20=80Nm	40·10=400N	400·0.35=140Nm
Force measured at the hands	15·10=150N	150·0.30=45Nm	15·10=150N	150·0.50=75Nm
Total		122Nm		215Nm
Inaccuracy in classification of the type of MMH				
	Lifting 25 kg (30° trunk flexion)		Pushing 25 kg (30° trunk flexion)	
	Force F=m·a	Moment M =d·F	Force F=m·a	Moment M=d·F
Gravitational force upper body	40·10=400N	400·0.20=80Nm	40·10=400N	400·0.20=80Nm
Force measured at the hands	25·10=250N	250·0.30=75Nm	250N	250·-0.10=-25Nm
Total		155 Nm		55Nm

We found rather small (<10%) non-systematic, random errors for peak and mean low-back moments when compared to a gold standard or when compared among raters. However, between-rater differences showed some occasional inaccuracies up to 45Nm (for peak moments) and 40Nm (for mean moments). Although these values are much smaller than the errors as shown in Table 9.1, this method may thus have some inaccuracy. One reason may be that modeling the entire trunk in two segments might not provide an accurate representation of the trunk curvature (Lariviere & Gagnon, 1999). Furthermore, both chapters showed relatively large inaccuracy in the dynamic component of the moment, which may partly be due to the interpolation between the four key data points. In addition, in tasks of short duration, random errors are strongly amplified due to double differentiation of position data. Therefore, interpolation or posture prediction algorithms can be assessed in future studies to improve interpolation accuracy. Using interpolation algorithms might additionally help to reduce analysis time.

Predictive values of mechanical low-back loads

In our data, we could not see a higher predictive value of mechanical loads obtained from the more accurate video analysis method as compared to mechanical loads calculated from observation postures. Inaccuracies as described in Table 9.1 are expected to bias the predictive value of posture observation-based estimation of mechanical load for LBP, as it is known that large inaccuracy leads to biased risk associations (Tielemans et al., 1998) and a reduced statistical power (Mathiassen et al., 2002; Mathiassen & Paquet, 2010). Assessment methods that are more accurate are therefore assumed to have a higher power when assessing associations with LBP. However, in practice this is not always the case in epidemiological literature (Griffith et al., 2012), as studies that measure more accurately often measure insufficient numbers of workers, which reduces the power of the study. Therefore, in an additional analysis we tested the effect of back load estimation accuracy by comparing the calculation of mechanical loads based on observations in Chapter 3 of this thesis to mechanical loads assessed more accurately using video analysis as described in Chapter 8. In the same population (as has been described in detail in Chapter 8), mechanical peak and cumulative loads were calculated as obtained from both methods. For individual largest peak moments, Pearson's correlation coefficients showed a poor correlation between these two methods (r=0.48). Differences between the two methods up to 200Nm can be seen (Figure 9.1). Based on the examples in Table 9.1, such differences are not unexpected and are probably mainly due to errors in the posture observation based method. It can therefore be concluded that, although mechanical loads are preferable over exposure metrics when assessing LBP etiology, mechanical loads can contain substantial inaccuracies, especially when observational data are used as input for a biomechanical model. However, these data also show that when individual peak loads are calculated on a group level, higher agreement (r=0.82) between the two methods is found. The video analysis method that we developed was expected to be more accurate. Therefore, as an additional analysis, we compared the predictive value of these two approaches. For both approaches, group-based mechanical load estimates were assessed (obtained from the group of workers from whom video analysis were performed; n=93) and were assigned to all tasks group members (those workers from whom LBP data in at least one of the three years of follow-up are available; n=1131). In this procedure, LBP was defined when a worker reported regular or prolonged LBP during at least one of the three years of follow-up. Crude risk associations were estimated by calculating ORs, 95% confidence intervals and p-values with the load (as continuous variables) being the independent variable and LBP (either case or control) being the dependent variable using logistic regression. From these results it can be concluded that, whereas relatively large differences exist between the two metrics for peak moment (on an individual level; Figure 9.1), both metrics show comparable predictive values for LBP (when group-based mechanical loads are used; Table 9.2). This may partly be caused by the fact that, although the two moments differ considerably on an individual level, these individual variations are diminished when calculating group-based estimates. This is in line with earlier studies that have shown high within-subject variability as compared to between-subject variability in a task group (Allread et al., 2000; Paquet et al., 2005). Measuring multiple subjects at separate occasions and assigning group-based load variables to all group members (as we have done here) is therefore an efficient way to reduce this variance, providing a stable metric that leads to stronger associations with LBP than individual metrics (Jansen & Burdorf, 2003). This implies that an improvement in the accuracy of assessments of a mechanical load on an individual level does not necessarily lead to a more predictive metric on a group level.

For cumulative loads, it was shown that both methods have a higher agreement on an individual level as shown by the relatively high correlation coefficients (r=0.87; Figure 9.1). Group-based mechanical load values agree even more among the two methods (r=0.96), which led to comparable predictive values for LBP for the two estimates of mechanical load (Table 9.2). A reason for this might be that random errors as a result of high inaccuracy in observations as input in a biomechanical model are diminished when calculating cumulative loads. This is at least partly caused by the fact that cumulative loads are based on roughly an hour of observation whereas peak loads occur just in a fraction of this measurement time. Another cause is the fact that, our video analysis method was used only for MMH tasks and mechanical loads during the periods in which no MMH tasks were performed are based on the same observational data in both estimates. This effect of small differences in predictive value for the two estimates is even more diminished when calculating group-based mechanical load. Considering the above, it can be questioned whether a large investment (in term of money and time) for measurements of physical work load is worth the effort. An answer to this question can be deducted from data presented in Chapter 5 clearly showing that at a certain point, it is more beneficial to include more workers in a study to collect LBP outcomes from than more workers to collect exposure data from. When exposure is measured from a sufficient number of workers (≥5 workers per task group), measuring exposure of more workers does not necessarily lead to higher powered risk associations.

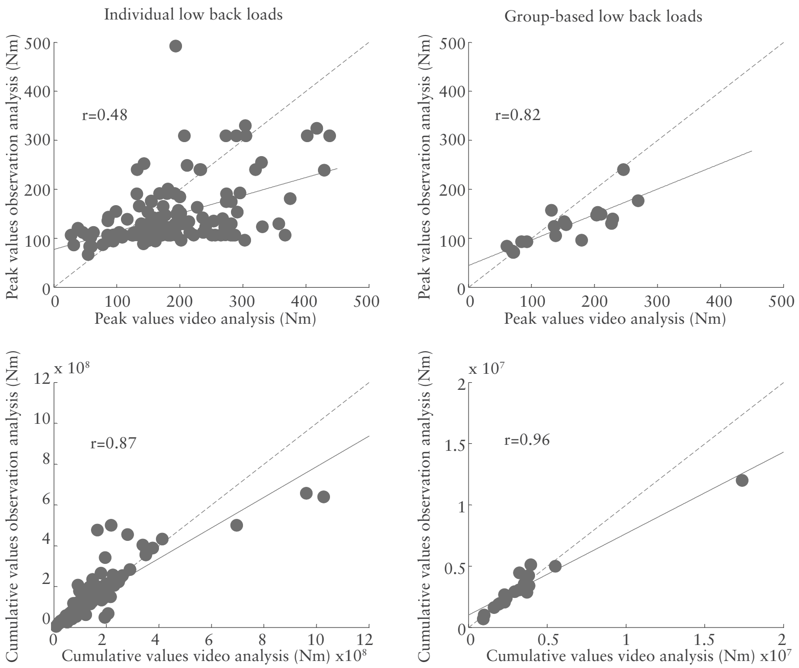

Figure 9.1 | Scatter plot depicting the association of mechanical loads as obtained from structured observational data used in a biomechanical model (Chapter 3 of this thesis; y-axis) and moments obtained from the video analysis method (Chapter 8 of this thesis; x-axis). Peak loads (upper panels) as well as cumulative loads (lower panels) on an individual level (left panels) and on a group level (right panels) are shown. The best fit through the data points (solid line) as well as the x=y reference line (dashed line) are shown. Also, Pearson's correlation coefficients are shown, depicting the correlation between the two methods of calculation of moments.

Non-linear association of low-back load and LBP

Based on several findings presented in this thesis, it can be speculated that there may be a non-linear association between low-back load and LBP. A first finding is that peak loads should be weighted higher when calculating cumulative loads. This is in accordance with data from earlier studies suggesting that that high forces have a more important impact on the increase in risk of failure than the duration of the load (Brinckmann et al., 1988; Hansson et al., 1987; Rapillard et al., 2006). According to this assumption, 15 cycles of 2000N load are presumed to be likely to cause a higher risk of damage than 20 cycles of 1500N. Therefore, multiple non-linear models for the association of mechanical load and LBP have been suggested.

Table 9.2 | Predictive value of peak and cumulative mechanical loads for the occurrence of LBP. Mean moments with standard deviations (Std), ORs, 95% CIs and levels of significance are shown for moments based on a biomechanical model using crude observational variables as input (data from Chapter 3) and moments obtained from a more accurate video analysis method (Chapter 8).

		LBP	No LBP		
Moments	Chapter	Mean (Std)	Mean (Std)	OR (95% CI)	p-value
Peak	Chapter 8	142.35(65.55)	134.82(61.25)	1.002(1.000-1.004)	0.047
Peak	Chapter 3	117.72(40.17)	113.66(36.42)	1.003(1.000-1.006)	0.076
Cumulative	Chapter 8	$1.25\cdot10^{8}(1.24\cdot10^{8})$	$1.05\cdot10^{8}(0.56\cdot10^{8})$	1.003(1.001-1.004)	0.001
Cumulative	Chapter 3	$1.17\cdot10^{8}(8.63\cdot10^{8})$	$1.03\cdot10^{8}(0.47\cdot10^{8})$	1.003(1.001-1.005)	0.001

For example, second order (Seidler et al., 2009; Seidler et al., 2003), fourth order weighting of load magnitude (Jäger et al., 2000), polynomial calculated cumulative load (Parkinson & Callaghan, 2007) or low-pass filtered loading (Krajcarski & Wells, 2008). Our findings show that weighting compression forces and number of load cycles with exponents of approximately 2 and 0.2, respectively, provides a suitable metric for prediction of in vitro failure due to cumulative loading (Chapter 4), which is in line with these suggestions. Therefore, cumulative loads containing weighting were expected to be better predictors for LBP. However, although squared cumulative loads tended to have a better fit than linearly weighted cumulative loads (Chapter 8), differences between these two metrics were marginal. A potential reason might be the lack of discriminating power in the current data. Despite the fact that we showed that in-vitro failure of spine segments during repeated loading at a constant load levels is best predicted when using a tenth power of load level (Chapter 4), this metric was not significantly associated with LBP in our epidemiological data (Chapter 8). Data from Chapter 4 were based on a mechanical load protocol applied on in-vitro material. On the one hand, in-vitro material lacks the potential to repair micro-damage, which would cause an overestimation of the importance of the loading frequency. On the other hand, in-vitro testing does not take into account that the risk of low-back injury may increase when respiratory or neuromuscular fatigue causes impaired coordination (Brereton & McGill, 1999; Janssens et al., 2010; Sparto et al., 1997; van Dieën et al., 1998), leading to an underestimation of the importance of the temporal characteristics of loading. As we show here that squared weighted load is, but load weighted to the tenth power is not associated to LBP, the latter characteristic of in-vivo conditions may play an important role here. However, this reasoning may be premature as the lack of predictive value of the tenth power weighting might also be a result of the fact that the metric becomes highly affected by inaccuracies in the measurements or actual variation in the work pattern. This can also be deduced from Chapter 8 showing that the association of this metric to LBP is very non-robust (Chapter 8). Finally, the suggested non-linearity can also be deducted from Figures 9.1 & 9.2 as it can be hypothesized that associations are likely caused by a relatively small group of workers that experience high mechanical loads and report a high prevalence of LBP, while the majority of workers are in task groups experiencing moderate low-back load and average LBP prevalence. This is in line with the non-linear association of physical work load and LBP that has been suggested before (e.g., McGill, 2009). Non-linear models, as have been discussed already in the past (e.g., Jansen & Burdorf, 2003) can therefore be considered when assessing such risk associations.

An additional simulation procedure was conducted in which the effect of the presence of certain groups in the data-set was assessed. In order to do so (based on the earlier described cohort in Chapter 8, with 19 task groups and with LBP assessed in 1131 subjects and mechanical load assessed in 93 subjects), all 19 task groups were consecutively left out of the cohort using a Jack-knife procedure (Chen et al., 2004; Efron & Gong, 1983). For each virtual study in which one of the groups was left out, logistic regressions were conducted using the peak and cumulative loads consecutively as continuous independent variables and LBP as the dichotomous dependent variable. ORs and p-values were calculated in each virtual study. Results show that, when leaving the group with the highest low-back load out of the cohort although ORs remain above 1, significant associations of both cumulative and peak loads disappear (Figure 9.2). This shows the importance of the presence of high mechanical loads in a cohort.

Etiology of LBP

Despite the limitations discussed above, it was shown in this thesis (Chapters 3 & 8) that cumulative low-back loads are highly predictive for the occurrence of LBP. These findings are in line with earlier studies (Kumar, 1990; Neumann et al., 2001a; Norman et al., 1998). Although peak loads have been shown to be significantly associated with LBP as well in earlier studies (Marras et al., 2010; Neumann et al., 2001a; Norman et al., 1998), this could not be confirmed in this thesis. Therefore, with respect to the etiology of LBP, our findings provide stronger support for a mechanism of LBP etiology due to cumulative loads than for a mechanism based on single peak loads. Such an etiological mechanism based on cumulative load might result from the occurrence of LBP as a consequence of injury or tissue responses due to accumulation of micro-damage or through impaired coordination due to neuromuscular or respiratory fatigue.

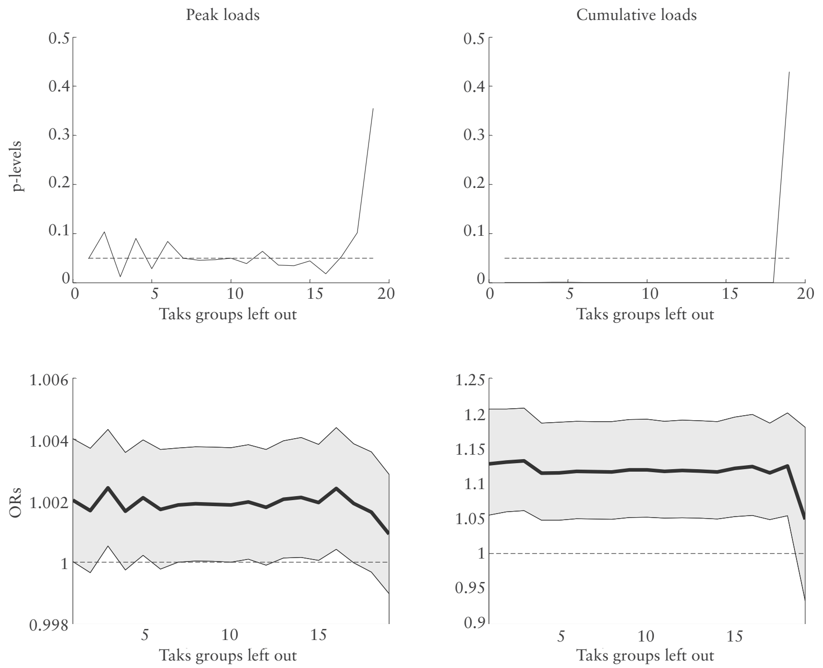

Figure 9.2 | Jack-knife leaving one task group out analysis after which risk associations were calculated. Risk associations were calculated for peak loads (left panels) and cumulative loads (right panels) showing the p-values (upper panels) and ORs and 95% confidence intervals (lower panels) of the associations. Task groups were ranked by magnitude of the mechanical load with group 1 being the group with the lowest mechanical load and group 19 being the group with the highest work load.

Chapter 8 shows a trend of the association of peak loads and LBP. It should be noted that the accuracy of a maximal peak load in an individual is lower than that of a cumulative load, thereby negatively affecting power. The fact that we could not prove the association of peak load to LBP therefore does not prove the absence of this effect. Furthermore, as it was shown in Chapter 8 that including the weighting of peaks in calculations of cumulative loads improves the predictive value for such cumulative loads, peaks should be taken into account. Finally, associations of peak loads and LBP have been shown in earlier studies (Marras et al., 2010; Neumann et al., 2001a; Norman et al., 1998).

DIRECTIONS FOR FUTURE RESEARCH AND ERGONOMIC PRACTICE

Although cumulative load has been shown to be predictive of LBP in this thesis, the exact underlying causal mechanism remains unknown. As an example, more research is needed on the contribution of peak loads on the development of LBP. Although valid and reliable mechanical loads were obtained using the present video analysis method, no substantial improvements in the predictive value were shown relative to observation-based estimation of mechanical loads. With the video assessment tool introduced here, only MMH tasks can be assessed and relatively large occasional errors were shown when using the method. Measurement tools that are able to obtain continuous accurate information on physical work load are therefore required. Several research groups have been working on ambulatory measurement systems that can be used in the field, using goniometers and ultrasonic systems to track body postures (Freitag et al., 2007; Glitsch et al., 2007; Marras et al., 2010). However, these devices are rather bulky and heavy, limiting workers in performing their work which hamper valid work load measurements. Potentially, more easily applicable methods can be found in the direction of ambulatory measurements tools using, for example wireless inertial sensors (Faber et al., 2009a; Faber et al., 2010c) in combination with instrumented force-shoes (Faber et al., 2010b). Also more sophisticated hardware using marker less motion tracking can be used in future studies (e.g., using devices such as Microsoft Kinect; Dutta, 2012). These methods have a low interference with the workers' tasks and allow for collection of large amounts of accurate data. Once these methods have been proven to be applicable in field measurements, they can be used on a larger scale to continuously monitor low-back loading in an epidemiological study.

A second direction for future research is to assess the non-linear association of low-back load and LBP. Although not convincingly demonstrated to be better than linear weighting in Chapter 8, a non-linear association of mechanical low-back load and LBP might be superior to linear weighting of mechanical loads in calculations of cumulative loads. One reason is that, as shown in Chapter 8, squared weighting showed a slightly better fit than linear weighting. Furthermore, it was shown in this epilogue that small task groups experiencing high low-back loads and having a high LBP prevalence play an important role in the calculation of risk associations. Finally, the analyses of in vitro data in Chapter 4 favored a non-linear weighting. Therefore, more research in the direction of modeling risk associations is necessary in order to improve the knowledge on LBP etiology. Furthermore, such non-linear models might be of importance to establish directives for physical work load. Current knowledge lacks the ability to recommend acceptable levels of biomechanical work load (Fallentin et al., 2001).

A final aspect that has shown to be important in this thesis is the variation (within and between subjects) that plays a role in the assessment of physical work load. It was shown that the choice of a measurement strategy (e.g., the size of a sample, allocation of a sample and the use of group vs. individual based load estimates) plays in important role in the assessment of physical work load. Therefore, in future epidemiological studies, assessment strategies should be analyzed a priori, using estimates of relevant variance components (Mathiassen, 2006). Furthermore, also monetary information (such as unit costs, and cost function shapes) could play a role here and should play a role in decisions on measurement strategies (Mathiassen & Bolin, 2011).

Concerning the implications of the present findings for ergonomic practice, it has been shown that cumulative loads are associated with the occurrence of back pain. Therefore, it might be of importance to target prevention on the reduction of cumulative loads. These loads are for a share caused by handling of heavy loads, working in awkward body postures (i.e., working in a trunk flexed posture combined with trunk rotation and large load distances with respect to the low-back) and working in unsafe environments.

As peaks play an important role in the weighting of cumulative loads, such loads should not be overlooked. As an example, these peaks loads can be caused by high low-back loads as a result of handling of heavy loads in awkward postures (i.e., working in a trunk flexed posture combined with trunk rotation and large load distances with respect to the low-back) in a high pace (i.e., causing high body accelerations). Prevention should thus be targeted based on these work situations and peak mechanical loads should be avoided by reducing low-back loads during MMH tasks.

Nederlandse Samenvatting

Dutch Summary

INTRODUCTIE

Lage rugpijn in de samenleving

Lage rugpijn (LRP) is een van de meest voorkomende gezondheidsklachten in onze samenleving en leidt tot verschillende problemen, zowel medisch als financieel (Maetzel & Li, 2002). De puntprevalentie in Nederland is 26% wat inhoudt dat op elk moment, een kwart van de Nederlandse populatie LRP heeft. Ook is er bekend dat, als mensen LRP hebben gekregen, deze pijn vaak terug komt (Andersson, 1999; Picavet & Schouten, 2003) en uiteindelijk chronisch wordt (Kovacs et al., 2005). Bij werknemers blijkt bovendien dat LRP kan leiden tot arbeidsongeschiktheid (Matsudaira et al., 2012), ziekteverzuim (Geuskens et al., 2008) en op jongere leeftijd stoppen met werken (Picavet & Schouten, 2003). De totale kosten voor de Nederlandse samenleving bedragen €4.3 miljard per jaar (0.6% van het bruto nationaal product; Lambeek et al., 2011). Het is daarom evident dat LRP een groot probleem is, wat onderzoek naar deze aandoening rechtvaardigt.

Naast persoonlijke risicofactoren (zoals leeftijd, geslacht, fysieke capaciteit en lichaamsgewicht; Hamberg-van Reenen et al., 2007; Leboeuf-Yde, 2004) en (werkgerelateerde) psychosociale risicofactoren (zoals stress, sociale steun op het werk en de mate waarin een werknemer zijn werk zelf kan indelen; Eatough et al., 2012; Hartvigsen et al., 2004) wordt LRP vaak geassocieerd met fysieke risicofactoren. Voorbeelden van dergelijke fysieke risicofactoren zijn tillen, dragen, duwen, trekken en belastende houdingen zoals een gebogen of een gedraaide romp (da Costa & Vieira, 2010; Griffith et al., 2012; Lis et al., 2007). Het effect van deze risicofactoren op LRP kennen we echter nog onvoldoende. Het verder bestuderen van fysieke risicofactoren van LRP is daarom het uitgangspunt van dit proefschrift. Deze samenvatting bevat alle onderwerpen die aan bod komen in dit proefschrift. Tevens worden de resultaten van de studies die in dit proefschrift zijn beschreven samengevat en bediscussieerd.

Van belasting tot schade

Ondanks het gebrek aan kennis over de etiologie (ontstaansmechanismen) van LRP zijn er binnen de literatuur verschillende modellen geopperd die het causale pad van het ontstaan van LRP beschrijven (Chaffin, 2009; van der Beek & Frings-Dresen, 1998; van Dieën et al., 1999; Wells et al., 2004). Deze modellen veronderstellen dat mechanische belasting op de lage rug als gevolg van blootstelling op het werk (zoals door de bovengenoemde variabelen tillen of buiging in de romp) een belangrijke factor is bij het ontstaan van LRP. Mechanische belastingsmaten (zoals momenten op de lage rug of krachten op de wervelkolom) omvatten meerdere traditionele blootstellingsmaten en houden bovendien rekening met andere factoren zoals geslacht en leeftijd. We veronderstellen daarom dat mechanische belasting een direct verband heeft met schade aan de rug wat uiteindelijk kan leiden tot rugklachten. In eerdere studies is gebleken dat dergelijke mechanische belastingsmaten een sterkere relatie met LRP hebben dan blootstellingsmaten zoals tillen of belastende houdingen. Inzicht in dergelijke maten is daarom belangrijk om de etiologie van LRP beter te begrijpen (Wells et al., 2004).

Twee soorten mechanische belasting kunnen leiden tot LRP: Kortdurende hoge belasting die kan leiden tot acute schade aan de rug of langdurige cumulatieve belasting. Deze langdurige belasting kan op den duur tot schade leiden doordat er steeds meer kleinere schades aan de rug komen. Hierdoor vermindert de capaciteit van de rug om belasting op te vangen. Een andere verklaring is dat door vermoeidheid de balans afneemt of de coördinatie vermindert. Dit zorgt voor blootstelling aan hoge rugbelastingen. Voor al deze theorieën is in de wetenschappelijke literatuur wel enig bewijs gevonden. Wat echter de exacte oorzaak voor klachten is, is nog onvoldoende bekend.

Meten van fysieke belasting

Een belangrijke reden voor het in de vorige alinea's besproken gebrek aan kennis over de relaties tussen fysieke belasting en LRP is dat deze relaties vaak beïnvloed worden door de gekozen meetmethoden (Burdorf, 2010; David, 2005). Bij het kiezen van een meetmethode moet namelijk een afweging gemaakt worden tussen nauwkeurigheid en toepasbaarheid (bijvoorbeeld vanwege investeringen in tijd en kosten) van de methode. Nauwkeurige methoden zijn vaak duur, waardoor naar minder nauwkeurige methoden wordt gegrepen om toch grote groepen proefpersonen te kunnen meten. Fysieke belasting wordt daarom vaak bepaald aan de hand van zelfrapportages door werknemers of aan de hand van observaties door observatoren (subjectieve beoordelingen of gestructureerde observaties van houdingen). Hoewel zelfrapportages en observaties vaak gebruikt worden binnen onderzoek is het bekend dat deze methoden een beperkte nauwkeurigheid hebben (van Wyk et al., 2009). Een derde groep meetmethoden vormen de directe meetmethoden (zoals metingen van spieractiviteit of houdingsmetingen met behulp van markers).
Deze methoden zijn het meest objectief en onderzoekers beschouwen deze als het meest betrouwbaar. Een nadeel is echter dat deze methoden vaak kostbaar zijn en lastig op de werkvloer zijn toe te passen. Videoanalysemethoden vormen een tussenweg om toch op een objectieve en accurate manier mechanische belasting op de werkvloer te meten. Hierbij hoeven immers enkel video-opnames op de werkvloer te worden gemaakt die op een later moment gedetailleerd geanalyseerd kunnen worden.

Variatie in fysieke belasting

Een ander belangrijk aspect waar rekening mee gehouden moet worden bij het ontwikkelen van een meetstrategie voor fysieke belasting is de variatie van de belasting. Fysieke belasting kan nogal verschillen tussen en binnen werknemers. Het is daarom ook lastig om fysieke belasting op de juiste momenten te meten. Omdat het continu meten van fysieke belasting kostbaar is, is het van belang om een meetstrategie te kiezen waarbij op gezette tijden wordt gemeten op een zodanige manier dat de kans op fouten zo klein mogelijk is. Zo is het in voorgaande studies gebleken dat data verzamelen over verschillende dagen en van verschillende werknemers meer informatie oplevert dan wanneer diezelfde meettijd gebruikt zou worden om dezelfde werknemer voor een langere tijd op één dag te meten

(Liv et al., 2010; Svendsen et al., 2005). Daarnaast is ook bekend dat, als er meerdere werknemers gemeten worden of als er per werknemer meerdere meetmomenten worden gedaan, de nauwkeurigheid van de meting op den duur niet meer toeneemt als er nog meer mensen of meetmomenten worden toegevoegd (Mathiassen et al., 2002; Mathiassen et al., 2003). Daarom kiezen onderzoekers vaak voor methoden waarbij binnen een groep werknemers (bijvoorbeeld een beroepsgroep) fysieke belasting in een gedeelte van de groepsleden wordt gemeten. Deze methode is succesvol gebleken om op een efficiënte manier fysieke belasting te meten. Over het effect van het gebruik van dergelijke methoden op uiteindelijke schattingen van risico's is echter nog onvoldoende bekend.

ONDERWERPEN VAN DIT PROEFSCHRIFT

Uit bovenstaande paragrafen blijkt dat het ontstaansmechanisme van LRP nog onvolledig bekend is. Dit komt omdat we nog te weinig weten over bepaalde aspecten van meetmethoden (zoals het meten van variatie in werkbelasting of het kwantificeren van belasting). Bovendien bestaan er maar weinig meetmethoden om mechanische belasting te meten op de werkvloer die eenvoudig toepasbaar zijn. Deze hiaten staan aan de basis van de volgende vier onderwerpen van dit proefschrift (die gedetailleerd besproken zijn in Hoofdstuk 2 tot en met 8 van dit proefschrift).

Dit proefschrift behandelt:

1. De voorspellende waarde van mechanische belastingen in vergelijking met (subjectieve) blootstellingsmaten voor LRP.
2. Methodologische keuzes in onderzoek naar het effect van rugbelasting op LRP.
3. De toepasbaarheid van een videoanalysemodel voor het meten van mechanische belasting.

Deze drie onderwerpen zullen als leidraad dienen voor het verwezenlijken van het einddoel, meer kennis vergaren over

4. De etiologie van LRP.

De voorspellende waarde van mechanische belastingen in vergelijking met (subjectieve) blootstellingsmaten voor LRP.

In Hoofdstuk 2 hebben we op basis van een longitudinale studie (waarbij over een periode van drie jaar het ontstaan van rug-, nek-, en schouderklachten is gevolgd) laten zien dat getrainde observatoren niet in staat waren om, aan de hand van observatie van video's genomen tijdens het werk, LRP te voorspellen, terwijl ze dat wel waren voor nek en schouderpijn. We kunnen dit verklaren doordat mechanische rugbelasting een samenspel is van verschillende factoren (zoals romphouding, armhouding en de grootte van de externe kracht) terwijl mechanische schouder- of nekbelasting dit niet is. De mate van rugbelasting is daardoor lastig (subjectief) in te schatten. De onnauwkeurigheid van deze subjectieve maten pleit voor het gebruik van meer accurate belastingsmaten.

Uit het onderzoek beschreven in Hoofdstuk 3 blijkt dat een mechanische belastingsmaat, bepaald aan de hand van houdingsobservaties, voorspellend is voor LRP. Bovendien is een dergelijke maat een betere voorspeller voor LRP dan dat conventionele blootstellingsmaten (frequentie van tillen of mate van rompflexie) dat zijn. Deze bevindingen ondersteunen daarmee de hypothese dat mechanische belasting een belangrijke maat is die ons meer kan leren over LRP. Op basis van deze twee studies kunnen we concluderen dat mechanische belastingsmaten van toegevoegde waarde zijn in het onderzoek naar het ontstaan van LRP.

Methodologische keuzes in onderzoek naar het effect van rugbelasting op LRP

Het bepalen van het effect van methodologische keuzes op resultaten van epidemiologische studies is een volgende stap om de etiologie van LRP beter te begrijpen. De eerste methodologische keuze is de keuze van een cumulatieve belastingsmaat. Zo wordt bij het bepalen van een cumulatieve belastingsmaat de grootte van de belasting vaak vermenigvuldigd met het aantal keren dat deze belasting optreedt. Zo leidt 15 keer een belasting van 2000 Newton (N) tot een totale cumulatieve belasting van 15*2000=30.000N. Echter, op basis van de kennis die we nu hebben kunnen we ook veronderstellen dat een incidentele hoge belasting meer effect heeft op het risico op schade of rugpijn dan dat het vaker optreden van kleinere belastingen (zoals 20 keer een belasting van 1500N). Hoofdstuk 4 bevestigt deze hypothese. In een analyse van gegevens verzameld uit kadavermateriaal hebben laten zien dat compressiebelasting met een tweede macht gewogen moet worden om de cumulatieve belasting te berekenen. Het aantal keren dat deze belasting voorkomt hoeft maar met een macht 0.2 gewogen te worden. Uit deze gegevens blijkt dat hogere piekbelastingen inderdaad zwaarder meegewogen moeten worden bij het berekenen van cumulatieve belasting. Dit heeft belangrijke implicaties voor vervolgstudies naar dit onderwerp. Zo levert 15 keer een belastingen van 2000N een groter gevaar op voor het ontstaan van LRP dan 20 keer een belasting van 1500N.

Een tweede methodologische keuze die we hebben onderzocht is het effect van groepsgrootte in een groeps-meetprotocol. In dergelijke protocollen wordt een meetpopulatie verdeeld in groepen op basis van een gemeenschappelijk kenmerk (bijvoorbeeld beroepsgroep). Vervolgens wordt de fysieke belasting maar in een deel van iedere groep gemeten omdat wordt verondersteld dat deze groepsbelasting representatief is voor belasting van de groepsleden. Groepsgemiddelden van de fysieke belasting worden dan gebruikt, en gekoppeld aan individuele gegevens over LRP, om LRP te voorspellen. In Hoofdstuk 5 van dit proefschrift wordt het effect van deze veelvoorkomende groepsstrategie onderzocht. De resultaten laten zien dat de mate waarin een fysieke belastingsmaat voorspellend is voor LRP voornamelijk beïnvloed wordt door het totaal aantal mensen in iedere groep. Het aantal mensen in iedere groep waarvan rugbelasting gemeten is, is minder relevant. Echter, als de fysieke belasting van onvoldoende mensen gemeten wordt, verminderd de nauwkeurigheid van de risicovoorspelling.

De toepasbaarheid van een videoanalysemodel voor het meten van mechanische belasting
Meten van fysieke belasting op de werkvloer is een lastige taak omdat de huidige meetmethoden vaak van invloed zijn op de manier waarop de werknemer zijn werk uit kan voeren. Bovendien zijn geavanceerde meetmethoden vaak duur. Bij goedkopere meetmethoden wordt de fysieke belasting vaak alleen in grove eenheden uitgedrukt waardoor de belastingsmeting minder betrouwbaar wordt. Daarom hebben we een videoanalysemethode ontwikkeld waarbij op de werkvloer video-opnames van werknemers gemaakt worden die later kunnen worden geanalyseerd. Hoofdstuk 6 van dit proefschrift beschrijft de validiteit (hoe goed wordt de beoogde belastingsmaat gemeten) van de videoanalysemethode. Hoofdstuk 7 beschrijft de inter-beoordelaarsbetrouwbaarheid van de videomethode (hoe groot zijn de verschillen tussen verschillende beoordelaars bij het toepassen van de methode). De resultaten laten zien dat er geen significante verschillen in piek en gemiddelde belasting zijn tussen de videomethode en een gouden standaard methode waarbij gedetailleerd wordt gemeten met behulp van markers. Bovendien was de samenhang tussen de gegevens verkregen met de videomethode en met deze gouden standaard groot. Daarnaast was de overeenkomt tussen de verschillende beoordelaars goed. Hoewel er sporadisch enkele grote verschillende tussen de beoordelaars te zien waren, waren deze verschillen gemiddeld klein.

De etiologie van LRP
In de laatste studie van dit proefschrift (Hoofdstuk 8) hebben we de eerder beschreven videomethode toegepast op een groot databestand van werknemers. We hebben de mechanische belasting van werknemers gemeten en deze gerelateerd aan het ontstaan van LRP in de daaropvolgende drie jaar. Het bleek dat cumulatieve belasting wel, maar piekbelasting niet voorspellend is voor LRP. De weging zoals voorgesteld in Hoofdstuk 4 zorgde niet voor een betere voorspelling van LRP dan wanneer deze weging niet werd gebruikt. Hoe dan ook laten deze gegevens zien dat cumulatieve belasting een belangrijke rol speelt in het krijgen van LRP. Deze belasting kan leiden tot schade door een opstapeling van kleine schades of door vermoeidheid. Op basis van onze gegevens lijkt piekbelasting, waardoor er acute schade kan optreden aan de lage rug, een minder voor de hand liggende verklaring voor het krijgen van LRP. Deze laatste verklaring kan echter niet worden uitgesloten op basis van de gegevens die we hebben.

DISCUSSIE

We kunnen op basis van dit proefschrift een aantal conclusies trekken. Als eerste is gebleken dat we met mechanische belastingen LRP beter kunnen voorspellen dan met conventionele (subjectieve) blootstellingsmaten van fysieke belasting. Dit is in lijn met de gedachte dat mechanische belasting informatie bevat van duur, frequentie en intensiteit van verschillende blootstellingsmaten, waardoor mechanische belasting een sterkere associatie met schade en daardoor met klachten heeft dan een individuele blootstellingsmaat. Daarom hebben we verder in dit proefschrift deze mechanische belastingsmaten gebruikt. Mechanische belasting, berekend op basis van observatiegegevens in een biomechanisch model, wordt vaak gebruikt in epidemiologische studies. Hoewel deze maten ook in dit proefschrift (Hoofdstuk 3) voorspellend zijn gebleken voor LRP is eerder aangetoond dat dergelijke methoden fikse onnauwkeurigheden kunnen bevatten (de Looze et al., 1994). Deze onnauwkeurigheden blijken onder andere uit de omvang van fouten bij het gebruik van grove houdingscategorieën tijdens observaties. Zo kunnen we bijvoorbeeld een tiltaak beschouwen die wordt beoordeeld in een romphoek-categorie van 30 tot 60°. Tiltaken waarbij 15 kg wordt getild en de romp zich in de beide extremen van deze categorie bevind (30° of 60° buiging) kunnen leiden tot een verschil in mechanische belasting dat tot een factor twee kan oplopen.

Uit de Hoofdstukken 6 en 7 is gebleken dat de videomethode die we hebben ontwikkeld valide en betrouwbaar is. Wat betreft nauwkeurigheid lijkt deze methode daarom beter te zijn dan de eerder gebruikte methode op basis van observatiegegevens in een biomechanisch model. Bovendien is gebleken dat de videomethode goed toepasbaar is in het ergonomische werkveld. Er werden echter wel sporadisch relatief hoge fouten gerapporteerd. Ondanks dat deze fouten veel kleiner zijn dan de eerder genoemde onnauwkeurigheden uit eerdere methoden, bevat ook deze methode nog wat onnauwkeurigheden.

Het is bekend dat onnauwkeurigheden in meetmethoden kunnen leiden tot onnauwkeurigheden in het schatten van risico’s (Tielemans et al., 1998) en dat LRP minder goed te voorspellen is bij onnauwkeurige metingen (Mathiassen et al., 2002; Mathiassen & Paquet, 2010). Toch is het niet altijd zo dat nauwkeuriger gemeten fysieke belasting leidt tot betere voorspellende waarden (Griffith et al., 2012). Dit zou kunnen komen doordat met nauwkeurige meetmethoden vaak onvoldoende proefpersonen gemeten kunnen worden omdat dit te kostbaar is. In Hoofdstuk 9 van dit proefschrift is daarom de voorspellende waarde van mechanische belasting, bepaald uit het videoanalysemodel (accuraat) vergeleken met de belasting bepaald op basis van observatiegegevens (minder accuraat). Uit deze vergelijking blijkt dat hoewel er substantiële verschillen zijn tussen de hoogte van de piekmomenten op individueel niveau, deze verschillen tamelijk goed wegmiddelen op groepsniveau. Bij het bepalen van cumulatieve belasting is het verschil tussen de twee methoden minimaal. Bovendien blijkt dat groepsgebaseerde piek of cumulatieve maten, maar marginaal beter voorspellend zijn als ze nauwkeuriger zijn gemeten. Dit houdt in dat een verbetering van de nauwkeurigheid van de belastingsmaat op individueel niveau niet noodzakelijk tot een verbetering van de voorspellende waarde van de maat op groepsniveau betekent. Men kan zich daarom afvragen of een grote investering (in geld en tijd) voor het verbeteren van de nauwkeurigheid van deze maat de moeite waard is. Een antwoord op deze vraag komt uit Hoofdstuk 5 waarin blijkt dat bij het meten van fysieke belasting van meer dan vijf mensen, de kans op het vinden van een significante associatie met LRP niet substantieel meer toeneemt.

(Non-)lineariteit van het effect van fysieke belasting op LRP

Op basis van verschillende bevindingen van dit proefschrift kan gespeculeerd worden dat er een niet-lineaire relatie bestaat tussen lage rugbelasting en LRP. Uit eerder onderzoek is gebleken dat piekbelasting zwaarder gewogen moet worden in berekeningen van cumulatieve belasting (Brinckmann et al., 1988; Rapillard et al., 2006). Volgens deze bevindingen leveren 15 belastingen van 2000N een groter gevaar op voor het ontstaan van LRP dan 20 belastingen van 1500N. Echter, deze weging leidde slechts tot marginaal betere voorspellende waarde voor LRP. Daarnaast is het ontbreken van een lineaire relatie ook terug te zien in de data uit Hoofdstuk 9, waarin blijkt dat de voorspellende waarde van mechanische belastingen voor een groot deel bepaald worden door een relatieve kleine groep met hoge mechanische belasting en een hoge prevalentie van LRP. De meerderheid van de werknemers had echter een relatief lage belasting en lage LRP prevalentie. Het al dan niet aanwezig zijn van deze non-lineariteit zal echter moeten blijken in vervolgstudies.

Etiologie van LRP

Zowel in Hoofdstuk 3 als Hoofdstuk 8 van dit proefschrift hebben we laten zien dat cumulatieve belasting voorspellend is voor het ontstaan van LRP. Dit is in lijn met eerdere studies (Kumar, 1990; Neumann et al., 2001). We hebben daarom meer bewijs gevonden voor een model van cumulatieve belasting dan voor een model voor het ontstaan van LRP voortkomend uit een enkele piekbelasting. Dergelijke cumulatieve belastingsmodellen kunnen verklaard worden aan de hand van ophoping van micro-schade of door verminderde coördinatie als gevolg van vermoeidheid. Hoewel piekbelasting in dit proefschrift niet significant voorspellend is voor LRP, was dit wel het geval in eerdere studies (Marras et al., 2010; Neumann et al., 2001). Ondanks dat hier geen direct bewijs voor is gevonden in onze gegevens, kunnen we het model van piekbelastingen daarom niet helemaal uitsluiten. De bepaling van piekbelasting is vatbaar voor onnauwkeurigheden in de meetmethode waardoor we mogelijk geen significante associaties hebben gevonden. Dat we geen effecten hebben aangetoond, betekent daarom ook niet dat deze effecten er niet zijn. Omdat gebleken is dat bij berekeningen van cumulatieve belasting, piekbelasting zwaarder moet meegewogen worden, lijkt het bovendien dat pieken wel degelijk een rol spelen en daarom niet over het hoofd gezien mogen worden.

Implicaties voor toekomstig onderzoek en voor het ergonomische werkveld

Hoewel op basis van onze gegevens cumulatieve belasting LRP beter voorspelt dan piekbelastingen, hebben we het exacte causale mechanisme nog niet helemaal ontrafeld. Meer onderzoek is daarom nodig om het exacte mechanisme waarlangs klachten ontstaan te begrijpen. Zo is er meer kennis nodig over de bijdrage van piekbelastingen aan de cumulatieve belasting. Ook is er verbetering van de meetmethoden van mechanische belasting gewenst. Methoden die continu de belasting kunnen meten op de werkvloer, waardoor veel data verzameld kunnen worden, kunnen hier uitkomst bieden. Daarnaast kan de non-lineariteit van de relatie tussen rugbelasting en LRP beter onderzocht worden. Meer kennis over deze relatie kan van belang zijn voor het ontwikkelen van preventieprogramma's om klachten te kunnen voorkomen. Zo is het denkbaar dat alleen boven een bepaalde belastingsgrens het risico op klachten substantieel toe zal nemen terwijl onder deze grens de kansen vergelijkbaar zijn. Als laatste hebben we meer informatie nodig over variatie in fysieke belasting en hoe dit een rol zou moeten spelen bij het ontwikkelen van nieuwe studies en meetmethoden van belasting op de werkvloer. Ook kennis over de kosten van metingen zou hierbij moeten worden meegenomen.

De belangrijkste implicaties van dit proefschrift voor het werkveld is dat cumulatieve belasting een belangrijke risicofactor is voor het ontstaan van rugklachten. Het is daarom van belang om cumulatieve belasting te verminderen bij werknemers om het risico op klachten te reduceren. Cumulatieve belasting kan bijvoorbeeld ontstaan uit het manueel werken met zware lasten, werken in onveilige omgevingen en werken in belastende houdingen (bijvoorbeeld met veel buiging in de romp, in combinatie met rotatie en grote lastafstanden van te tillen voorwerpen ten opzichten van de lage rug). Echter, omdat we piekbelastingen als risicofactor voor LRP niet uit kunnen sluiten, is reductie van piekbelasting eveneens van belang ter preventie van LRP. Deze belastingen ontstaan voornamelijk door hoge lasten als gevolg van het manueel werken met zware lasten in ongunstige houdingen (bijvoorbeeld met veel buiging in romp combinatie met rotatie en grote lastafstanden ten opzichten van de lage rug) in een hoog tempo (met grote lichaamsversnellingen).

References

REFERENCES

Adams, MA, 2004. Biomechanics of back pain. *Acupuncture in Medicine: Journal of the British Medical Acupuncture Society* 22(4): 178-188.

Adams, MA, McNally, DS, Chinn, H, Dolan, P, 1994. Posture and compressive strength of the lumbar spine. *Clinical Biomechanics* 9(1): 5-14.

Akaike, H, 1974. A new look at the statistical model identification *IEEE Transactions on Automatic Control* 19(6): 716-723.

Alexopoulos, EC, Stathi, IC, Charizani, F, 2004. Prevalence of musculoskeletal disorders in dentists. *BMC Musculoskeletal Disorders* 5(16).

Allread, WG, Marras, WS, Burr, DL, 2000. Measuring trunk motions in industry: variability due to task factors, individual differences, and the amount of data collected. *Ergonomics* 43(6): 691-701.

Anderson, CK, Chaffin, DB, Herrin, GD, Matthews, LS, 1985. A biomechanical model of the lumbosacral joint during lifting activities. *Journal of Biomechanics* 18(8): 571-584.

Andersson, GB, 1999. Epidemiological features of chronic low-back pain. *Lancet* 354(9178): 581-585.

Andrews, DM, Holmes, AM, Weiry, PL, Arnold, TA, Callaghan, JP, 2008. Decision times and errors increase when classifying trunk postures near posture bin boundaries. *Theoretical Issues in Ergonomics Science* 9(5): 425-440.

Ariëns, GA, Bongers, PM, Douwes, M, Miedema, MC, Hoogendoorn, WE, van der Wal, G, Bouter, LM, van Mechelen, W, 2001. Are neck flexion, neck rotation, and sitting at work risk factors for neck pain? Results of a prospective cohort study. *Occupational and Environmental Medicine* 58(3): 200-207.

Bakker, EW, Verhagen, AP, van Trijffel, E, Lucas, C, Koes, BW, 2009. Spinal mechanical load as a risk factor for low back pain: a systematic review of prospective cohort studies. *Spine* 34(8): E281-293.

Balogh, I, Orbaek, P, Ohlsson, K, Nordander, C, Unge, J, Winkel, J, Hansson, GA, 2004. Self-assessed and directly measured occupational physical activities -- influence of musculoskeletal complaints, age and gender. *Applied Ergonomics* 35(1): 49-56.

Balogh, I, Orbaek, P, Winkel, J, Nordander, C, Ohlsson, K, Ektor-Andersen, J, 2001. Questionnaire-based mechanical exposure indices for large population studies -- reliability, internal consistency and predictive validity. *Scandinavian Journal of Work Environment & Health* 27(1): 41-48.

Bao, S, Howard, N, Spielholz, P, Silverstein, B, Polissar, N, 2009. Interrater reliability of posture observations. *Human Factors* 51(3): 292-309.

Barrero, LH, Katz, JN, Dennerlein, JT, 2009a. Validity of self-reported mechanical demands for occupational epidemiologic research of musculoskeletal disorders. *Scandinavian Journal of Work Environment & Health* 35(4): 245-260.

Barrero, LH, Katz, JN, Perry, MJ, Krishnan, R, Ware, JH, Dennerlein, JT, 2009b. Work pattern causes bias in self-reported activity duration: a randomised study of mechanisms and implications for exposure assessment and epidemiology. *Occupational and Environmental Medicine* 66(1): 38-44.

van der Beek, AJ, Braam, IT, Douwes, M, Bongers, PM, Frings-Dresen, MH, Verbeek, JH, Luyts, S, 1994. Validity of a diary estimating exposure to tasks, activities, and postures of the trunk. *International Archives of Occupational and Environmental Health* 66(3): 173-178.

van der Beek, AJ, Frings-Dresen, MH, 1998. Assessment of mechanical exposure in ergonomic epidemiology. *Occupational Environmental Medicine* 55(5): 291-299.

van der Beek, AJ, van Gaalen, LC, Frings-Dresen, MH, 1992. Working postures and activities of lorry drivers: a reliability study of on-site observation and recording on a pocket computer. *Applied Ergonomics* 23(5): 331-336.

Bonato, P, Ebenbichler, GR, Roy, SH, Lehr, S, Posch, M, Kollmitzer, J, Della Croce, U, 2003. Muscle fatigue and fatigue-related biomechanical changes during a cyclic lifting task. *Spine* 28(16): 1810-1820.

Bot, SD, van der Waal, JM, Terwee, CB, van der Windt, DA, Schellevis, FG, Bouter, LM, Dekker, J, 2005. Incidence and prevalence of complaints of the neck and upper extremity in general practice. *Annals of the Rheumatic Diseases* 64(1): 118-123.

Brereton, LC, McGill, SM, 1999. Effects of physical fatigue and cognitive challenges on the potential for low back injury. *Human Movement Science* 18(6): 839-857.

Brinckmann, P, Biggemann, M, Hilweg, D, 1988. Fatigue fracture of human lumbar vertebrae. *Clinical Biomechanics* 3(Suppl 1): 1-23.

Brinckmann, P, Biggemann, M, Hilweg, D, 1989. Prediction of the compressive strength of human lumbar vertebrae. *Spine* 14(6): 606-610.

Burdorf, A, 2010. The role of assessment of biomechanical exposure at the workplace in the prevention of musculoskeletal disorders. *Scandinavian Journal of Work Environment & Health* 36(1): 1-2.

Burdorf, A, Jansen, JP, 2006. Predicting the long term course of low back pain and its consequences for sickness absence and associated work disability. *Occupational and Environmental Medicine* 63(8): 522-529.

Burt, S, Punnett, L, 1999. Evaluation of interrater reliability for posture observations in a field study. *Applied Ergonomics* 30(2): 121-135.

Callaghan, JP, McGill, SM, 2001. Intervertebral disc herniation: studies on a porcine model exposed to highly repetitive flexion/extension motion with compressive force. *Clinical Biomechanics* 16(1): 28-37.

Callaghan, JP, Salewytsch, AJ, Andrews, DM, 2001. An evaluation of predictive methods for estimating cumulative spinal loading. *Ergonomics* 44(9): 825-837.

Cann, AP, Connolly, M, Ruuska, R, MacNeil, M, Birmingham, TB, Vandervoort, AA, Callaghan, JP, 2008. Inter-rater reliability of output measures for a posture matching assessment approach: a pilot study with food service workers. *Ergonomics* 51(4): 556-572.

Cappozzo, A, Catani, F, Croce, UD, Leardini, A, 1995. Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical Biomechanics* 10(4): 171-178.

Chaffin, DB, 2009. The evolving role of biomechanics in prevention of overexertion injuries. *Ergonomics* 52(1): 3-14.

Chang, CC, Hsiang, S, Dempsey, PG, McGorry, RW, 2003. A computerized video coding system for biomechanical analysis of lifting tasks. *International Journal of Industrial Ergonomics* 32(4): 239-250.

Chang, CC, McGorry, RW, Lin, JH, Xu, X, Hsiang, SM, 2010. Prediction accuracy in estimating joint angle trajectories using a video posture coding method for sagittal lifting tasks. *Ergonomics* 53(8): 1039-1047.

Chen, JC, Chang, WR, Shih, TS, Chen, CJ, Chang, WP, Dennerlein, JT, Ryan, LM, Christiani, DC, 2004. Using "Exposure Prediction Rules" for Exposure Assessment - An Example on Whole-Body Vibration in Taxi Drivers. *Epidemiology* 15(3): 293-299.

Chen, SM, Liu, MF, Cook, J, Bass, S, Lo, SK, 2009. Sedentary lifestyle as a risk factor for low back pain: a systematic review. *International Archives of Occupational and Environmental Health* 82(7): 797-806.

Coenen, P, Kingma, I, Boot, CR, Bongers, PM, van Dieën, JH, 2012. The contribution of load magnitude and number of load cycles to cumulative low-back load estimations: a study based on in-vitro compression data. *Cinical Biomechanics* 27(10): 1083-1086.

Coenen, P, Kingma, I, Boot, CR, Faber, GS, Xu, X, Bongers, PM, van Dieën, JH, 2011. Estimation of low back moments from video analysis: a validation study. *Journal of Biomechanics* 44(13): 2369-2375.

Coenen, P, Kingma, I, Boot, CRL, Bongers, PM, van Dieën, JH, 2013a. Inter-rater reliability of a video-analysis method measuring low-back load in a field situation. *Applied Ergonomics* 44(5): 828-834.

Coenen, P, Kingma, I, Boot, CRL, Twisk, JWR, Bongers, PM, van Dieën, JH, 2013b. Cumulative low back load at work as a risk factor of low back pain: a prospective cohort study. *Journal of Occupational Rehabilitation* 23(1): 11-18.

Costa-Black, KM, Loisel, P, Anema, JR, Pransky, G, 2010. Back pain and work. *Best Practice & Research, Clinical Rheumatology* 24(2): 227-240.

Côté, JN, 2012. A critical review on physical factors and functional characteristics that may explain a sex/gender difference in work-related neck/shoulder disorders. *Ergonomics* 55(2): 173-182.

Côté, P, van der Velde, G, Cassidy, JD, Carroll, LJ, Hogg-Johnson, S, Holm, LW, Carragee, EJ, Haldeman, S, Nordin, M, Hurwitz, EL, Guzman, J, Peloso, PM, 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* 33(Suppl 4): 60-74.

da Costa, BR, Vieira, ER, 2010. Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. *American Journal of Industrial Medicine* 53(3): 285-323.

David, GC, 2005. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational Medicine* 55(3): 190-199.

Davis, KG, Marras, WS, 2000. Assessment of the relationship between box weight and trunk kinematics: does a reduction in box weight necessarily correspond to a decrease in spinal loading? *Human Factors* 42(2): 195-208.

Davison, AC, Hinkley, DV, 1997. Bootstrap methods and their applications. Cambridge University Press, Cambridge.

Dempsey, PG, 2007. Effectiveness of ergonomics interventions to prevent musculoskeletal disorders: Beware of what you ask. *International Journal of Industrial Ergonomics* 37(2): 169-173.

van Dieën, JH, Faber, GS, Loos, RC, Kuijer, PP, Kingma, I, van der Molen, HF, Frings-Dresen, MH, 2010. Validity of estimates of spinal compression forces obtained from worksite measurements. *Ergonomics* 53(6): 792-800.

van Dieën, JH, Kingma, I, 2005. Effects of antagonistic co-contraction on differences between electromyography based and optimization based estimates of spinal forces. *Ergonomics* 48(4): 411-426.

van Dieën, JH, van der Burg, P, Raaijmakers, TA, Toussaint, HM, 1998. Effects of repetitive lifting on kinematics: inadequate anticipatory control or adaptive changes? *Journal of Motor Behavior* 30(1): 20-32.

van Dieën, JH, van der Veen, A, van Royen, BJ, Kingma, I, 2006. Fatigue failure in shear loading of porcine lumbar spine segments. *Spine* 31(15): E494-498.

van Dieën, JH, Weinans, H, Toussaint, HM, 1999. Fractures of the lumbar vertebral endplate in the etiology of low back pain: a hypothesis on the causative role of spinal compression in aspecific low back pain. *Medical Hypotheses* 53(3): 246-252.

Dolan, P, Adams, MA, 1998. Repetitive lifting tasks fatigue the back muscles and increase the bending moment acting on the lumbar spine. *Journal of Biomechanics* 31(8): 713-721.

Dumas, R, Cheze, L, Verriest, JP, 2007. Adjustments to McConville et al. and Young et al. body segment inertial parameters. *Journal of Biomechanics* 40(3): 543-553.

Dutta, T, 2012. Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. *Applied Ergonomics* 43(4): 645-649.

Eatough, EM, Way, JD, Chang, CH, 2012. Understanding the link between psychosocial work stressors and work-related musculoskeletal complaints. *Applied Ergonomics* 43(3): 554-563.

Efron, B, Gong, G, 1983. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician* 37(1): 36-48.

Efron, B, Tibshirani, R, 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1(1): 54-77.

Eriksen, W, Bruusgaard, D, Knardahl, S, 2004. Work factors as predictors of intense or disabling low back pain; a prospective study of nurses' aides. *Occupational Environmental Health* 61(5): 398-404.

Faber, A, Giver, H, Stroyer, J, Hannerz, H, 2010a. Are low back pain and low physical capacity risk indicators for dropout among recently qualified eldercare workers? A follow-up study. *Scandinavian Journal of Public Health* 38(8): 810-816.

Faber, GS, Kingma, I, Bruijn, SM, van Dieën, JH, 2009a. Optimal inertial sensor location for ambulatory measurement of trunk inclination. *Journal of Biomechanics* 42(14): 2406-2409.

Faber, GS, Kingma, I, Kuijer, PPFM, van der Molen, HF, Hoozemans, MJM, Frings-Dresen, MHW, van Dieën, JH, 2009b. Working height, block mass and one- vs. two-handed block handling: the contribution to low back and shoulder loading during masonry work. *Ergonomics* 52(9): 1104-1118.

Faber, GS, Kingma, I, Martin Schepers, H, Veltink, PH, van Dieën, JH, 2010b. Determination of joint moments with instrumented force shoes in a variety of tasks. *Journal of Biomechanics* 43(14): 2848-2854.

Faber, GS, Kingma, I, van Dieën, JH, 2007. The effects of ergonomic interventions on low back moments are attenuated by changes in lifting behaviour. *Ergonomics* 50(9): 1377-1391.

Faber, GS, Kingma, I, van Dieën, JH, 2010c. Bottom-up estimation of joint moments during manual lifting using orientation sensors instead of position sensors. *Journal of Biomechanics* 43(7): 1432-1436.

Faber, GS, Kingma, I, van Dieën, JH, 2011. Effect of initial horizontal object position on peak L5/S1 moments in manual lifting is dependent on task type and familiarity with alternative lifting strategies. *Ergonomics* 54(1): 72-81.

Fallentin, N, Viikari-Juntura, E, Waersted, M, Kilbom, Å, 2001. Evaluation of physical workload standards and guidelines from a Nordic perspective. *Scand Journal of Work Environment & Health* 27(Suppl 2): 1-52.

Fejer, R, Kyvik, KO, Hartvigsen, J, 2006. The prevalence of neck pain in the world population: a systematic critical review of the literature. *European Spine Journal* 15(6): 834-848.

Ferguson, SA, Gaudes-MacLaren, LL, Marras, WS, Waters, TR, Davis, KG, 2002. Spinal loading when lifting from industrial storage bins. *Ergonomics* 45(6): 399-414.

Fleiss, JL, 1986. The Design and Analysis of Clinical Experiments. John Wiley and sons, Toronto.

Freitag, S, Ellegast, R, Dulon, M, Nienhaus, A, 2007. Quantitative measurement of stressful trunk postures in nursing professions. *The Annals of Occupational Hygiene* 51(4): 385-395.

Gallagher, S, Marras, WS, Litsky, AS, Burr, D, Landoll, J, Matkovic, V, 2007. A comparison of fatigue failure responses of old versus middle-aged lumbar motion segments in simulated flexed lifting. *Spine* 32(17): 1832-1839.

Geuskens, GA, Hazes, JMW, Barendregt, PJ, Burdorf, A, 2008. Predictors of sick leave and reduced productivity at work among persons with early inflammatory joint conditions. *Scandinavian Journal of Work, Environment & Health* 34(6): 420-429.

Glitsch, U, Ottersbach, HJ, Ellegast, R, Schaub, K, Franz, G, Jager, M, 2007. Physical workload of flight attendants when pushing and pulling trolleys aboard aircraft. *International Journal of Industrial Ergonomics* 37(1-2): 845-854.

Godin, G, Jobin, J, Bouillon, J, 1986. Assessment of leisure time exercise behavior by self-report: a concurrent validity study. *Canadian Journal of Public Health* 77(5): 359-362.

Goetzel, RZ, Hawkins, K, Ozminkowski, RJ, Wang, S, 2003. The health and productivity cost burden of the "top 10" physical and mental health conditions affecting six large U.S. employers in 1999. *Journal of Occupational and Environmental Medicine* 45(1): 5-14.

Goffe, WL, Ferrier, GD, Rogers, J, 1994. Global Optimization of Statistical Functions with Simulated Annealing. *Journal of Econometrics* 60(1-2): 65-100.

Granata, KP, Gottipati, P, 2008. Fatigue influences the dynamic stability of the torso. *Ergonomics* 51(8): 1258-1271.

Griffith, LE, Shanon, HS, Wells, RP, Walter, SD, Cole, DC, Côté, P, Frank, J, Hogg-Johnson, S, Langlois, L, 2012. Individual participant data meta-analysis of mechanical workplace risk factors and low back pain. *American Journal of Public Health* 102(2): 309-318.

Gunning, JL, Callaghan, JP, McGill, SM, 2001. Spinal posture and prior loading history modulate compressive strength and type of failure in the spine: a biomechanical study using a porcine cervical spine model. *Clinical Biomechanics* 16(6): 471-480.

Hägg, GM, Luttmann, A, Jäger, M, 2000. Methodologies for evaluating electromyographic field data in ergonomics. *Journal of Electromyography and Kinesiology* 10(5): 301-312.

Hamberg-van Reenen, HH, Ariens, GA, Blatter, BM, van Mechelen, W, Bongers, PM, 2007. A systematic review of the relation between physical capacity and future low back and neck/shoulder pain. *Pain* 130(1-2): 93-107.

Hansson, GA, Arvidsson, I, Ohlsson, K, Nordander, C, Mathiassen, SE, Skerfving, S, Balogh, I, 2006. Precision of measurements of physical workload during standardised manual handling. Part II: Inclinometry of head, upper back, neck and upper arms. *Journal of Electromyography and Kinesiology* 16(2): 125-136.

Hansson, GA, Balogh, I, Bystrom, JU, Ohlsson, K, Nordander, C, Asterland, P, Sjolander, S, Rylander, L, Winkel, J, Skerfving, S, 2001. Questionnaire versus direct technical measurements in assessing postures and movements of the head, upper back, arms and hands. *Scandinavian Journal of Work Environment & Health* 27(1): 30-40.

Hansson, T, Roos, B, Nachemson, A, 1980. The bone mineral content and ultimate compressive strength of lumbar vertebrae. *Spine* 5(1): 46-55.

Hansson, TH, Keller, TS, Spengler, DM, 1987. Mechanical behaviour of the human lumbar spine. II. Fatigue strength during dynamic compressive loading. *Journal of Orthopaedic Research* 5(4): 479-487.

Hartvigsen, J, Bakketeig, LS, Leboeuf-Yde, C, Engberg, M, Lauritzen, T, 2001. The association between physical workload and low back pain clouded by the "healthy worker" effect: population-based cross-sectional and 5-year prospective questionnaire study. *Spine* 26(16): 1788-1792.

Hartvigsen, J, Lings, S, Leboeuf-Yde, C, Bakketeig, L, 2004. Psychosocial factors at work in relation to low back pain and consequences of low back pain; a systematic, critical review of prospective cohort studies. *Occupational and Environmental Medicine* 61(1): e2.

Hestbaek, L, Leboeuf-Yde, C, Kyvik, KO, Manniche, C, 2006. The course of low back pain from adolescence to adulthood: eight-year follow-up of 9600 twins. *Spine* 31(4): 468-472.

van den Heuvel, SG, Ariëns, GA, Boshuizen, HC, Hoogendoorn, WE, Bongers, PM, 2004. Prognostic factors related to recurrent low-back pain and sickness absence. *Scandinavian Journal of Work, Environment & Health* 30(6): 459-467.

van den Heuvel, SG, van der Beek, AJ, Blatter, BM, Hoogendoorn, WE, Bongers, PM, 2005. Psychosocial work characteristics in relation to neck and upper limb symptoms. *Pain* 114(1-2): 47-53.

Hildebrandt, VH, Douwes, M, 1991. Physical load and work: Questionnaire on musculoskeletal load and health complaints (Lichamelijke belasting en arbeid: vragenlijst bewegingsapparaat). *Voorburg: Ministry of Social Affairs and Employment*.

Hof, AL, 1992. An explicit expression for the moment in multibody systems. *Journal of Biomechanics* 25(10): 1209-1211.

Holmstrom, E, Moritz, U, 1991. Low back pain -- correspondence between questionnaire, interview and clinical examination. *Scandinavian Journal of Rehabilitation Medicine* 23(3): 119-125.

Holmstrom, EB, Lindell, J, Moritz, U, 1992. Low back and neck/shoulder pain in construction workers: occupational workload and psychosocial risk factors. Part 2: Relationship to neck and shoulder pain. *Spine* 17(6): 672-677.

Hooftman, WE, van Poppel, MN, van der Beek, AJ, Bongers, PM, van Mechelen, W, 2004. Gender differences in the relations between work-related physical and psychosocial risk factors and musculoskeletal complaints. *Scandinavian Journal of Work Environment & Health* 30(4): 261-278.

Hoogendoorn, WE, Bongers, PM, de Vet, HC, Douwes, M, Koes, BW, Miedema, MC, Ariens, GA, Bouter, LM, 2000a. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. *Spine* 25(23): 3087-3092.

Hoogendoorn, WE, van Poppel, MN, Bongers, PM, Koes, BW, Bouter, LM, 1999. Physical load during work and leisure time as risk factors for back pain. *Scandinavian Journal of Work Environment & Health* 25(5): 387-403.

Hoogendoorn, WE, van Poppel, MN, Bongers, PM, Koes, BW, Bouter, LM, 2000b. Systematic review of psychosocial factors at work and private life as risk factors for back pain. *Spine* 25(16): 2114-2125.

Hoozemans, MJ, Burdorf, A, van der Beek, AJ, Frings-Dresen, MH, Mathiassen, SE, 2001. Group-based measurement strategies in exposure assessment explored by bootstrapping. *Scandinavian Journal of Work Environment & Health* 27(2): 125-132.

Hoozemans, MJ, Kingma, I, de Vries, WH, van Dieën, JH, 2008. Effect of lifting height and load mass on low back loading. *Ergonomics* 51(7): 1053-1063.

Howarth, SJ, Callaghan, JP, 2012. Compressive force magnitude and intervertebral joint flexion/extension angle influence shear failure force magnitude in the porcine cervical spine. *Journal of Biomechanics* 45(3): 484-490.

Hoy, D, Bain, C, Williams, G, March, L, Brooks, P, Blyth, F, Woolf, A, Vos, T, Buchbinder, R, 2012. A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism* 64(6): 2028-2037.

Hsiang, SM, Brogmus, GE, Martin, SE, Bezverkhny, IB, 1998. Video based lifting technique coding system. *Ergonomics* 41(3): 239-256.

ISCO, 1968. International Standard Classification of Occupations (ISCO). International Labour Office, Geneva.

Jäger, M, Jordan, C, Luttmann, A, Laurig, W, Group, D, 2000. Evaluation and assessment of lumbar load during total shifts for occupational manual materials handling jobs within the Dortmund Lumbar Load Study - DOLLY. *International Journal of Industrial Ergonomics 25*(6): 553-571.

Jansen, JP, Burdorf, A, 2003. Effects of measurement strategy and statistical analysis on dose-response relations between physical workload and low back pain. *Occupational and Environmental Medicine* 60(12): 942-947.

Janssens, L, Brumagne, S, Polspoel, K, Troosters, T, McConnell, A, 2010. The effect of inspiratory muscles fatigue on postural control in people with and without recurrent low back pain. *Spine* 35(10): 1088-1094.

Johanson, E, Brumagne, S, Janssens, L, Pijnenburg, M, Claeys, K, Paasuke, M, 2011. The effect of acute back muscle fatigue on postural control strategy in people with and without recurrent low back pain. *European Spine Journal* 20(12): 2152-2159.

de Jonge, J, Reuvers, MM, Houtman, IL, Bongers, PM, Kompier, MA, 2000. Linear and nonlinear relations between psychosocial job characteristics, subjective outcomes, and sickness absence: baseline results from SMASH. Study on Musculoskeletal Disorders, Absenteeism, Stress, and Health. *Journal of Occupational Health Psychology* 5(2): 256-268.

Karasek, R, 1985. Job content instrument user guides: revision 1.1. *Los Angeles, CA: Department of Industrial and System Engineering, University of Southern California*.

Kerr, MS, Frank, JW, Shannon, HS, Norman, RW, Wells, RP, Neumann, WP, Bombardier, C, 2001. Biomechanical and psychosocial risk factors for low back pain at work. *American Journal of Public Health* 91(7): 1069-1075.

Kingma, I, de Looze, MP, Toussaint, HM, Klijnsma, HG, Bruijnen, TBM, 1996. Validation of a full body 3-D dynamic linked segment model. *Human Movement Science* 15(6): 833-860.

Kingma, I, de Looze, MP, van Dieën, JH, Toussaint, HM, Adams, MA, Baten, CT, 1998. When is a lifting movement too asymmetric to identify low-back loading by 2-D analysis? *Ergonomics* 41(10): 1453-1461.

Kingma, I, Faber, GS, Bakker, AJ, van Dieën, JH, 2006. Can low back loading during lifting be reduced by placing one leg beside the object to be lifted? *Physical Therapy* 86(8): 1091-1105.

Kingma, I, Faber, GS, van Dieën, JH, 2010. How to lift a box that is too large to fit between the knees. *Ergonomics* 53(10): 1228-1238.

Kociolek, AM, Keir, PJ, 2010. Reliability of distal upper extremity posture matching using slow-motion and frame-by-frame video methods. *Human Factors* 52(3): 441-455.

Koes, BW, van Tulder, MW, Thomas, S, 2006. Diagnosis and treatment of low back pain. *BMJ* 332(7555): 1430-1434.

Kovacs, FM, Abraira, V, Zamora, J, Fernandez, C, 2005. The transition from acute to subacute and chronic low back pain: a study based on determinants of quality of life and prediction of chronic disability. *Spine* 30(15): 1786-1792.

Krajcarski, S, Wells, R, 2008. The time variation pattern of mechanical exposure and the reporting of low back pain. *Theoretical Issues in Ergonomics Science* 9(1): 45-71.

Kromhout, H, Heederik, D, 1995. Occupational epidemiology in the rubber industry: implications of exposure variability. *American Journal of Industrial Medicine* 27(2): 171-185.

Kromhout, H, Tielemans, E, Preller, L, Heederik, D, 1996. Estimates of individual dose from current measurements of exposure *Occupational Hygiene* 3(1-3): 23-39.

Kuiper, JI, Burdorf, A, Verbeek, JHAM, Frings-Dresen, MHW, van der Beek, AJ, Viikari-Juntura, ERA, 1999. Epidemiologic evidence on manual material handling as a risk factor for back disorders: a systematic review. *International Journal of Industrial Ergonomics* 24(4): 389-404.

Kumar, S, 1990. Cumulative load as a risk factor for back pain. *Spine* 15(12): 1311-1316.

Kuorinka, I, Jonsson, B, Kilbom, A, Vinterberg, H, Biering-Sorensen, F, Andersson, G, Jorgensen, K, 1987. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Applied Ergonomics* 18(3): 233-237.

Kwon, BK, Roffey, DM, Bishop, PB, Dagenais, S, Wai, EK, 2011. Systematic review: occupational physical activity and low back pain. *Occupational Medicine* 61(8): 541-548.

Lambeek, LC, van Tulder, MW, Swinkels, IC, Koppes, LL, Anema, JR, van Mechelen, W, 2011. The trend in total cost of back pain in The Netherlands in the period 2002-2007. *Spine* 36(13): 1050-1058.

Lariviere, C, Gagnon, D, 1999. The influence of trunk modelling in 3D biomechanical analysis of simple and complex lifting tasks. *Clinical Biomechanics* 14(7): 449-461.

Leboeuf-Yde, C, 2004. Back pain -- individual and genetic factors. *Journal of Electromyography and Kinesiology* 14(1): 129-133.

de Leva, P, 1996. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *Journal of Biomechanics* 29(9): 1223-1230.

Linton, SJ, 2001. Occupational psychological factors increase the risk for back pain: a systematic review. *Journal of Occupational Rehabilitation* 11(1): 53-66.

Lis, AM, Black, KM, Korn, H, Nordin, M, 2007. Association between sitting and occupational LBP. *European Spine Journal* 16(2): 283-298.

Liv, P, Mathiassen, SE, Svendsen, SW, 2010. Theoretical and empirical efficiency of sampling strategies for estimating upper arm elevation. *Annals of Occupational Hygiene* 55(4): 436-449.

Loomis, D, Kromhout, H, 2004. Exposure variability: concepts and applications in occupational epidemiology. *American Journal of Industrial Medicine* 45(1): 113-122.

de Looze, MP, Kingma, I, Bussmann, JBJ, Toussaint, HM, 1992. Validation of a dynamic linked segment model to calculate joint moments in lifting. *Clinical Biomechanics* 7(3): 161-169.

de Looze, MP, Kingma, I, Thunnissen, W, van Wijk, MJ, Toussaint, HM, 1994a. The evaluation of a practical biomechanical model estimating lumbar moments in occupational activities. *Ergonomics* 37(9): 1495-1502.

de Looze, MP, Toussaint, HM, Ensink, J, Mangnus, C, van der Beek, AJ, 1994b. The validity of visual observation to assess posture in a laboratory-simulated, manual material handling task. *Ergonomics* 37(8): 1335-1343.

de Looze, MP, Visser, B, Houting, I, van Rooy, MA, van Dieën, JH, Toussaint, HM, 1996. Weight and frequency effect on spinal loading in a bricklaying task. *Journal of Biomechanics* 29(11): 1425-1433.

Lötters, F, Burdorf, A, Kuiper, J, Miedema, H, 2003. Model for the work-relatedness of low-back pain. *Scandinavian Journal of Work Environment & Health* 29(6): 431-440.

Lowe, BD, 2004. Accuracy and validity of observational estimates of shoulder and elbow posture. *Applied Ergonomics* 35(2): 159-171.

Macfarlane, GJ, Pallewatte, N, Paudyal, P, Blyth, FM, Coggon, D, Crombez, G, Linton, S, Leino-Arjas, P, Silman, AJ, Smeets, RJ, van der Windt, D, 2009. Evaluation of work-related psychosocial factors and regional musculoskeletal pain: results from a EULAR Task Force. *Annals of the Rheumatic Diseases* 68(6): 885-891.

Maetzel, A, Li, L, 2002. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Practice & Research. Clinical Rheumatology* 16(1): 23-30.

Manek, NJ, MacGregor, AJ, 2005. Epidemiology of back disorders: prevalence, risk factors, and prognosis. *Current Opinion in Rheumatology* 17(2): 134-140.

Marras, WS, 2012. The complex spine: the multidimensional system of causal pathways for low-back disorders. *Human Factors* 54(6): 881-889.

Marras, WS, Ferguson, SA, Burr, D, Davis, KG, Gupta, P, 2004. Spine loading in patients with low back pain during asymmetric lifting exertions. *the Spine Journal* 4(1): 67-75.

Marras, WS, Ferguson, SA, Burr, D, Schabo, P, Maronitis, A, 2007. Low back pain recurrence in occupational environments. *Spine* 32(21): 2387-2397.

Marras, WS, Granata, KP, Davis, KG, Allread, WG, Jorgensen, MJ, 1999. Effects of box features on spine loading during warehouse order selecting. *Ergonomics* 42(7): 980-996.

Marras, WS, Lavender, SA, Ferguson, SA, Splittstoesser, RE, Yang, G, 2010. Quantitative dynamic measures of physical exposure predict low back functional impairment. *Spine* 35(8): 914-923.

Marras, WS, Lavender, SA, Leurgans, SE, Fathallah, FA, Ferguson, SA, Allread, WG, Rajulu, SL, 1995. Biomechanical risk factors for occupationally related low back disorders. *Ergonomics* 38(2): 377-410.

Marras, WS, Lavender, SA, Leurgans, SE, Rajulu, SL, Allread, WG, Fathallah, FA, Ferguson, SA, 1993. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. The effects of workplace factors, trunk position, and trunk motion characteristics on risk of injury. *Spine* 18(5): 617-628.

Mathiassen, SE, 2006. Diversity and variation in biomechanical exposure: what is it, and why would we like to know? *Applied Ergonomics* 37(4): 419-427.

Mathiassen, SE, Bolin, K, 2011. Optimizing cost-efficiency in mean exposure assessment - cost functions reconsidered. *BMC Medical Research Methodology* 11: 76.

Mathiassen, SE, Burdorf, A, van der Beek, AJ, 2002. Statistical power and measurement allocation in ergonomic intervention studies assessing upper trapezius EMG amplitude. A case study of assembly work. *Journal of Electromyography and Kinesiology* 12(1): 45-57.

Mathiassen, SE, Burdorf, A, van der Beek, AJ, Hansson, GA, 2003a. Efficient one-day sampling of mechanical job exposure data -- a study based on upper trapezius activity in cleaners and office workers. *American Industrial Hygiene Association Journal* 64(2): 196-211.

Mathiassen, SE, Moller, T, Forsman, M, 2003b. Variability in mechanical exposure within and between individuals performing a highly constrained industrial work task. *Ergonomics* 46(8): 800-824.

Mathiassen, SE, Nordander, C, Svendsen, SW, Wellman, HM, Dempsey, PG, 2005. Task-based estimation of mechanical job exposure in occupational groups. *Scandinavian Journal of Work Environment & Health* 31(2): 138-151.

Mathiassen, SE, Paquet, V, 2010. The ability of limited exposure sampling to detect effects of interventions that reduce the occurrence of pronounced trunk inclination. *Applied Ergonomics* 41(2): 295-304.

Matsudaira, K, Konishi, H, Miyoshi, K, Isomura, T, Takeshita, K, Hara, N, Yamada, K, Machida, H, 2012. Potential risk factors for new onset of back pain disability in Japanese workers: findings from the Japan epidemiological research of occupation-related back pain study. *Spine* 37(15): 1324-1333.

McGill, SM, 2009. Evolving ergonomics? *Ergonomics* 52(1): 80-86.

Moray, N, 2000. Culture, politics and ergonomics. *Ergonomics* 43(7): 858-868.

Neumann, WP, Wells, RP, Norman, RW, Frank, J, Shannon, H, Kerr, MS, 2001a. A posture and load sampling approach to determining low-back pain risk in occupational settings. *International Journal of Industrial Ergonomics* 27(2): 65-77.

Neumann, WP, Wells, RP, Norman, RW, Kerr, MS, Frank, J, Shannon, HS, Group, OW, 2001b. Trunk posture: reliability, accuracy, and risk estimates for low back pain from a video based assessment method. *International Journal of Industrial Ergonomics* 28(6): 355-365.

Nicholson, PH, Alkalay, R, 2007. Quantitative ultrasound predicts bone mineral density and failure load in human lumbar vertebrae. *Clinical Biomechanics* 22(6): 623-629.

van Nieuwenhuyse, A, Somville, PR, Crombez, G, Burdorf, A, Verbeke, G, Johannik, K, van den Bergh, O, Masschelein, R, Mairiaux, P, Moens, GF, 2006. The role of physical workload and pain related fear in the development of low back pain in young workers: evidence from the BelCoBack Study; results after one year of follow up. *Occupational and Environmental Medicine* 63(1): 45-52.

Norman, R, Wells, R, Neumann, P, Frank, J, Shannon, H, Kerr, M, 1998. A comparison of peak vs cumulative physical work exposure risk factors for the reporting of low back pain in the automotive industry. *Clinical Biomechanics* 13(8): 561-573.

van Oostrom, SH, Verschuren, WM, de Vet, HC, Picavet, HS, 2011. Ten year course of low back pain in an adult population-based cohort--the Doetinchem cohort study. *European Journal of Pain* 15(9): 993-998.

Palmer, KT, Smedley, J, 2007. Work relatedness of chronic neck pain with physical findings -- a systematic review. *Scandinavian Journal of Work Environment & Health* 33(3): 165-191.

Pan, W, 2001. Akaike's information criterion in generalized estimating equations. *Biometrics* 57(1): 120-125.

Paquet, V, Punnett, L, Woskie, S, Buchholz, B, 2005. Reliable exposure assessment strategies for physical ergonomics stressors in construction and other non-routinized work. *Ergonomics* 48(9): 1200-1219.

Parkinson, RJ, Callaghan, JP, 2007. The role of load magnitude as a modifier of the cumulative load tolerance of porcine cervical spinal units: progress towards a force weighting approach. *Theoretical Issues in Ergonomics Science* 8(3): 171-184.

Paul, JA, Douwes, M, 1993. Two-dimensional photographic posture recording and description: a validity study. *Applied Ergonomics* 24(2): 83-90.

Picavet, HS, Schouten, JS, 2003. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 102(1-2): 167-178.

Plamondon, A, Gagnon, M, Desjardins, P, 1996. Validation of two 3-D segment models to calculate the net reaction forces and moments at the L(5)/S(1) joint in lifting. *Clinical Biomechanics* 11(2): 101-110.

Potvin, JR, 1997. Use of NIOSH equation input to calculate lumbosacral compression forces. *Ergonomics* 40(7): 691-707.

Punnett, L, 2004. Work related neck pain: how important is it, and how should we understand its causes? *Occupational and Environmental Medicine* 61(12): 954-955.

Punnett, L, Fine, LJ, Keyserling, WM, Herrin, GD, Chaffin, DB, 1991. Back disorders and nonneutral trunk postures of automobile assembly workers. *Scandinavian Journal of Work, Environment & Health* 17(5): 337-346.

Punnett, L, Pruss-Utun, A, Nelson, DI, Fingerhut, MA, Leigh, J, Tak, S, Phillips, S, 2005. Estimating the global burden of low back pain attributable to combined occupational exposures. *American Journal of Industrial Medicine* 48(6): 459-469.

Punnett, L, Wegman, DH, 2004. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology* 14(1): 13-23.

Rapillard, L, Charlebois, M, Zysset, PK, 2006. Compressive fatigue behavior of human vertebral trabecular bone. *Journal of Biomechanics* 39(11): 2133-2139.

Reeves, GK, Cox, DR, Darby, SC, Whitley, E, 1998. Some aspects of measurement error in explanatory variables for continuous and binary regression models. *Statistics in Medicine* 17(19): 2157-2177.

van Rijn, RM, Huisstede, BM, Koes, BW, Burdorf, A, 2010. Associations between work-related factors and specific disorders of the shoulder -- a systematic review of the literature. *Scandinavian Journal of Work Environment & Health* 36(3): 189-201.

Roelen, CA, Koopmans, PC, Groothoff, JW, 2010. Subjective health complaints in relation to sickness absence. *Work* 37(1): 15-21.

Rothman, KJ, Greenland, S, 2005. Causation and causal inference in epidemiology. *American Journal of Public Medicine* 95(Suppl 1): 144-150.

Seidler, A, Bergmann, A, Jager, M, Ellegast, R, Ditchen, D, Elsner, G, Grifka, J, Haerting, J, Hofmann, F, Linhardt, O, Luttmann, A, Michaelis, M, Petereit-Haack, G, Schumann, B, Bolm-Audorff, U, 2009. Cumulative occupational lumbar load and lumbar disc disease -- results of a German multi-center case-control study (EPILIFT). *BMC Musculoskeletal Disorders* 10: 48.

Seidler, A, Bolm-Audorff, U, Heiskel, H, Henkel, N, Roth-Kuver, B, Kaiser, U, Bickeboller, R, Willingstorfer, WJ, Beck, W, Elsner, G, 2001. The role of cumulative physical work load in lumbar spine disease: risk factors for lumbar osteochondrosis and spondylosis associated with chronic complaints. *Occupational and Environmental Medicine* 58(11): 735-746.

Seidler, A, Bolm-Audorff, U, Siol, T, Henkel, N, Fuchs, C, Schug, H, Leheta, F, Marquardt, G, Schmitt, E, Ulrich, PT, Beck, W, Missalla, A, Elsner, G, 2003. Occupational risk factors for symptomatic lumbar disc herniation; a case-control study. *Occupational and Environmental Medicine* 60(11): 821-830.

Shrout, PE, Fleiss, JL, 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* 86(2): 420-428.

Smedley, J, Egger, P, Cooper, C, Coggon, D, 1997. Prospective cohort study of predictors of incident low back pain in nurses. *BMJ* 314(7089): 1225-1228.

Sparto, PJ, Parnianpour, M, Reinsel, TE, Simon, S, 1997. The effect of fatigue on multijoint kinematics, coordination, and postural stability during a repetitive lifting test. *The Journal of Orthopaedic and Sports Physical Therapy* 25(1): 3-12.

Spielholz, P, Silverstein, B, Morgan, M, Checkoway, H, Kaufman, J, 2001. Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. *Ergonomics* 44(6): 588-613.

Stewart, WF, Ricci, JA, Chee, E, Morganstein, D, Lipton, R, 2003. Lost productive time and cost due to common pain conditions in the US workforce. *Journal of the American Medical Association* 290(18): 2443-2454.

Sullivan, D, Bryden, P, Callaghan, JP, 2002. Inter- and intra-observer reliability of calculating cumulative lumbar spine loads. *Ergonomics* 45(11): 788-797.

Sutherland, CA, Albert, WJ, Wrigley, AT, Callaghan, JP, 2008. A validation of a posture matching approach for the determination of 3D cumulative back loads. *Applied Ergonomics* 39(2): 199-208.

Svendsen, SW, Mathiassen, SE, Bonde, JP, 2005. Task based exposure assessment in ergonomic epidemiology: a study of upper arm elevation in the jobs of machinists, car mechanics, and house painters. *Occupational and Environmental Medicine* 62(1): 18-27.

Svensson, E, 2001. Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine* 33(1): 47-48.

Takala, EP, Pehkonen, I, Forsman, M, Hansson, GA, Mathiassen, SE, Neumann, WP, Sjogaard, G, Veiersted, KB, Westgaard, RH, Winkel, J, 2010. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scandinavian Journal of Work Environment & Health* 36(1): 3-24.

Tielemans, E, Kupper, LL, Kromhout, H, Heederik, D, Houba, R, 1998. Individual-based and group-based occupational exposure assessment: some equations to evaluate different strategies. *The Annals of Occupational Hygiene* 42(2): 115-119.

Tiemessen, IJ, Hulshof, CT, Frings-Dresen, MH, 2008. Low back pain in drivers exposed to whole body vibration: analysis of a dose-response pattern. *Occupational Environmental Medicine* 65(10): 667-675.

Trask, C, Mathiassen, SE, Wahlstrom, J, Heiden, M, Rezagholi, M, 2012. Data collection costs in industrial environments for three occupational posture exposure assessment methods. *BMC Medical Research Methodology* 12: 89.

Twisk, JWR, 2003. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge University Press, New York.

Twisk, JWR, 2006. Applied mulitlevel analysis. Cambridge University Press, Cambridge, United Kingdom.

van Tulder, M, Koes, B, Bombardier, C, 2002. Low back pain. *Best Practice & Research, Clinical Rheumatology* 16(5): 761-775.

van der Veen, AJ, Mullender, M, Smit, TH, Kingma, I, van Dieën, JH, 2005. Flow-related mechanics of the intervertebral disc: the validity of an in vitro model. *Spine* 30(18): E534-539.

Verbeek, JH, Martimo, KP, Karppinen, J, Kuijer, PP, Viikari-Juntura, E, Takala, EP, 2011. Manual material handling advice and assistive devices for preventing and treating back pain in workers. *Cochrane Database of Systematic Reviews* 18(3): CD005958.

Wahlstrom, J, Mathiassen, SE, Liv, P, Hedlund, P, Ahlgren, C, Forsman, M, 2010. Upper arm postures and movements in female hairdressers across four full working days. *The Annals of Occupational Hygiene* 54(5): 584-594.

Wai, EK, Rodriguez, S, Dagenais, S, Hall, H, 2008. Evidence-informed management of chronic low back pain with physical activity, smoking cessation, and weight loss. *the Spine Journal* 8(1): 195-202.

Wai, EK, Roffey, DM, Bishop, P, Kwon, BK, Dagenais, S, 2010. Causal assessment of occupational lifting and low back pain: results of a systematic review. *the Spine Journal* 10(6): 554-566.

Wang, Y, Videman, T, Battie, MC, 2012a. ISSLS prize winner: Lumbar vertebral endplate lesions: associations with disc degeneration and back pain history. *Spine* 37(17): 1490-1496.

Wang, Y, Videman, T, Battie, MC, 2012b. Lumbar vertebral endplate lesions: prevalence, classification, and association with age. *Spine* 37(17): 1432-1439.

Waters, TR, Putz-Anderson, V, Garg, A, Fine, LJ, 1993. Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics* 36(7): 749-776.

Welch, L, Haile, E, Boden, LI, Hunting, KL, 2009. Musculoskeletal disorders among construction roofers - physical function and disability. *Scandinavian Journal of Work, Environment & Health* 35(1): 56-63.

Wells, R, van Eerd, D, Hägg, G, 2004. Mechanical exposure concepts using force as the agent. *Scandinavian Journal of Work, Environment & Health* 30(3): 179-190.

Westgaard, RH, Winkel, J, 1997. Ergonomic intervention research for improved musculoskeletal health: A critical review. *International Journal of Industrial Ergonomics* 20(6): 463-500.

Winkel, J, Mathiassen, SE, 1994. Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations. *Ergonomics* 37(6): 979-988.

van Wyk, PM, Weir, PL, Andrews, DM, Fiedler, KM, Callaghan, JP, 2009. Determining the optimal size for posture categories used in video-based posture assessment methods. *Ergonomics* 52(8): 921-930.

Xu, X, Chang, CC, Faber, GS, Kingma, I, Dennerlein, JT, 2010a. Comparing polynomial and cubic spline interpolation of segment angles for estimating L5/S1 net moment during symmetric lifting tasks. *Journal of Biomechanics* 43(10): 583-586.

Xu, X, Chang, CC, Faber, GS, Kingma, I, Dennerlein, JT, 2010b. Interpolation of segment Euler angles can provide a robust estimation of segment angular trajectories during asymmetric lifting tasks. *Journal of Biomechanics* 43(10): 2043-2048.

Xu, X, Chang, CC, Faber, GS, Kingma, I, Dennerlein, JT, 2011. The validity and inter-rater reliability of video-based posture observation during asymmetric lifting tasks. *Human Factors* 53(4): 371-382.

Xu, X, Chang, CC, Faber, GS, Kingma, I, Dennerlein, JT, 2012. Estimation of 3-D peak L5/S1 joint moment during asymmetric lifting tasks with cubic spline interpolation of segment Euler angles. *Applied Ergonomics* 43(1): 115-120.

Zatsiorsky, VM, 2002. Kinetics of Human Motion. Human Kinetics, Champaign, IL.

Zhang, X, Chaffin, D, 2000. A three-dimensional dynamic posture prediction model for simulating in-vehicle seated reaching movements: development and validation. *Ergonomics* 43(9): 1314-1330.

ABOUT THE AUTHOR

Pieter Coenen was born on January 22nd 1985 in Tilburg. After graduating from secondary school at Durendael Oisterwijk in 2004, he started studying at the Faculty of Human Movement Sciences at the VU University in Amsterdam. After obtaining his Masters degree in 2009, Pieter worked as a research assistant at the Faculty of Human Movement Sciences at the VU University in Amsterdam and later at the Duyvensz-Nagel Research Laboratory at the Rehabilitation Center Amsterdam (currently: Reade). In 2009, Pieter started his PhD of which the results are presented in this thesis entitled 'On the origin of back pain'. Furthermore, Pieter was involved in multiple research projects regarding epidemiology, biomechanics and ergonomics. These projects were among others a result of collaborations with TNO and the University of Gävle. Pieter followed the postgraduate epidemiology program at the VU University Medical Center which will soon grant him the 'Epidemiologist B' registration.

LIST OF PUBLICATIONS

Papers in international peer reviewed journals

Coenen P, Kingma I, Boot CR, Faber GS, Xu X, Bongers PM, Dieen JH. 2011. Estimation of low back moments from video analysis a validation study. *Journal of Biomechanics* 44(13): 2369-2375.*

Coenen P, Kingma I, Boot CR, Douwes M, Bongers PM, Dieen JH. 2012. Risk of musculoskeletal pain estimations by trained observers during work site observations - a prospective cohort study. *Ergonomics 55*(11): 1373-1381.*

Coenen P, Kingma I, Boot CR, Bongers PM, Dieen JH. 2012. The contribution of load magnitude and number of load cycles to cumulative low-back load estimations: a study based on in-vitro compression data. *Clinical Biomechanics* 27(10): 1083-1086.*

Coenen P, van Werven G, van Nunen MP, van Dieen JH, Gerrits KHL, Janssen TW. 2012. Robot-assisted walking versus overground walking in stroke patients: an evaluation of muscle activity. *Journal of Rehabilitation Medicine* 44(4): 331-337.

Mazaheri M, **Coenen P**, Parnianpour M, Kiers H, van Dieën JH. 2013. Low back pain and postural sway during quiet standing with and without sensory manipulation: a systematic review. *Gait & Posture* 37(1): 12-22.

Coenen P, Kingma I, Boot CR, Twisk JWR, Bongers PM, Dieen JH. 2013. Cumulative low back load at work as a risk factor of low back pain: a prospective cohort study. *Journal of Occupational Rehabilitation* 32(1): 11-18.*

Coenen P, Kingma I, Boot CR, Bongers PM, Dieen JH. 2013. Inter-rater reliability of a video-analysis method measuring low back load in a field situation. *Applied Ergonomics* 44(5): 828-834.*

Douwes M, Boocock M, **Coenen P**, van den Heuvel S, Bosch T. In press. Predictive validity of the Hand Arm Risk assessment Method (HARM). *International Journal of Industrial Ergonomics*.

Coenen P, Kingma I, Boot CR, Bongers PM, van Dieën JH. Submitted. Cumulative mechanical low-back load at work is a determinant of low-back pain.*

Coenen P, Mathiassen SE, Kingma I, Boot CR, Bongers PM, van Dieën JH. Submitted. Effects of the data sampling strategy on bias and power in epidemiologic studies of low-back pain – a bootstrapping approach.*

Coenen P, Formanoy M, Douwes M, Bosch T, de Kraker H. Submitted. Validity and inter-observer reliability of subjective hand-arm vibration assessments.

*Papers marked with a * are part is this thesis.*

Publications in Dutch journals

Coenen P, van Werven G, van Nunen MP, Stolwijk J, Konijnenbelt M, Janssen TWJ, Gerrits KHL. 2009. Robot gestuurde looptherapie in Nederland: Een overzicht van de resultaten en ervaringen. *Revalidata* 150: 10-12.

van Engelen SJPM, Willigenburg NW, **Coenen P**. 2011. De wervelkolom biomechanisch bekeken: hoe belasting kan leiden tot gebreken. *Physios* 4: 14-23.

Formanoy M, **Coenen P**, Douwes M, Bosch T, de Kraker H. 2013. Het meten van hand-arm trillingen. *Tijdschrift voor Ergonomie* 38(2): 18-23.

Conference papers

van Werven G, **Coenen P**, van Nunen MP, Gerrits KHL, Janssen TW. Muscle activity patterns during robotic walking and overground walking in subjects with stroke. *International state-of-the-art congress Rehabilitation: Mobility Exercise and Sports* 2010, Amsterdam.

Janssen TW, van Werven G, **Coenen P**, van Nunen MP, Gerrits KHL. 2010. Muscle activity during robotic treadmill versus overground walking in individuals with hemiplegia. *Medicine & Science in Sports & Medicine* 42(5):S111.

Coenen P, Kingma I, Boot CR, Bongers PM, Dieen JH. Cumulative low back load as a risk factor for low back pain. *World Physical Therapy Congress* 2011, Amsterdam, the Netherlands.

Coenen P, Kingma I, Boot CR, Faber GS, Xu X, Bongers PM, Dieen JH Occupational low back load assessment using a video analysis method. *ISB 2011*, Brussels, Belgium.

Coenen P, Kingma I, Boot CR, Bongers PM, Dieen JH. Field based low-back load assessment using video analysis, *Premus 2013*, Busan, South Korea.

Coenen P, Mathiassen SE, Kingma I, Boot CR, Bongers PM, Dieen JH. The effect of group-based exposure measurement strategies on the statistical significance of an association between trunk rotations and low-back pain, *Premus 2013*, Busan, South Korea.

Bosch T, Douwes M, Boocock M, **Coenen P**, van den Heuvel S. Predictive validity of the Hand Arm Risk assessment Method (HARM), Premus 2013, Busan, South Korea.

Coenen P, Mathiassen SE, Kingma I, Boot CR, Bongers PM, Dieen JH. Group-based exposure measurement strategies and their effects on trunk rotation and low-back pain exposure-outcome associations. *EPICOH 2013*, Utrecht, the Netherlands.

BEDANKT! THANKS!

Nu genoeg over mijzelf... Iemand moet het zware werk doen. Maar het was natuurlijk nooit zover gekomen zonder: promotieteam, oppositie, co-auteurs, TNO 'orde in de chaos' hulptroepen, FBW collega's, roomies, 'B656 bad boys', stoere mannen, CBF-members uit Gävle en paranimfen. En misschien nog wel belangrijker vrienden, familie, Beau Geste leden en voetbalmaten voor jullie broodnodige afleiding, mentale steun of gewoon zomaar.

In het bijzonder moeten toch even genoemd worden….

Eduardo
Cécile Jack Dennerlein
Peer Nienke Pa en Ma Maaskantje
Michiel de Looze Marco Hoozemans
SvendErik Mathiassen Roel Jasmijn Jaap
Xu Xu
Swenneke van den Heuvel Lex Burdorf
Jos Twisk Marjolein Douwes Niek
Paulien Joop Allard van der Beek
Idsart dr Lee Gert Faber

Allen, heel hartelijk dank! Thanks a lot to all of you!