From UNIT OF INTERVENTION & IMPLEMENTATION RESEARCH FOR WORKER HEALTH, INSTITUTE OF ENVIRONMENTAL MEDICINE
Karolinska Institutet, Stockholm, Sweden

FROM HEALTH OUTCOMES TO VALUE ASSESSMENTS:

Preference-based evaluation of interventions and valuation of productivity costs among working adults

Emmanuel Aboagye

Stockholm 2017

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Eprint AB 2017

ISBN 978-91-7676-740-5

FROM HEALTH OUTCOMES TO VALUE ASSESSMENTS:
Preference-based evaluation of interventions and valuation of productivity costs among working adults

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Emmanuel Aboagye

Principal Supervisor:
Professor Irene Jensen
Karolinska Institutet
Department of Environmental Medicine
Division of Intervention & Implementation Research for Worker Health

Co-supervisor(s):
Assistant Professor Malin Lohela Karlsson
Karolinska Institutet
Department of Environmental Medicine
Division of Intervention & Implementation Research for Worker Health

Associate Professor Iben Axén
Karolinska Institutet
Department of Environmental Medicine
Division of Intervention & Implementation Research for Worker Health

Associate Professor Lydia Kwak
Karolinska Institutet
Department of Environmental Medicine
Division of Intervention & Implementation Research for Worker Health

Opponent:
Associate Professor Trine Kjær
University of Southern Denmark
Department of Business and Economics
Division of Health Economics Research

Examination Board:
Associate Professor Martin Henriksson
Linköping University
Department of Medical and Health Sciences
Division of Health Care Analysis

Associate Professor Eva Rasmussen Barr
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Physiotherapy

Professor Mikael Forsman
Karolinska Institutet
Department of Environmental Medicine
Division of Occupational medicine

To Tanja, Milana and Elias

ABSTRACT

Background: This thesis applies value assessment frameworks to non-specific low back pain (LBP) in a working population to evaluate health outcomes and to estimate productivity costs. LBP is a common diagnosis in primary health care. It is characterized by recurrent pain episodes and is a major factor in the increasing sickness absence among workers. LBP contributes substantially to the burden of disease and to the economic burden in terms of productivity loss to employer and society at large. Interventions are thus needed to prevent recurring LBP and its associated burden. Such interventions need to be properly evaluated from a variety of perspectives, including that of the worker, the employer and society at large, in order to help us better understand how to support worker-health policies and a sustainable working life.

Aims: In the first part of this thesis, the frameworks used aim to investigate the cost-effectiveness of and preferences for secondary prevention interventions for LBP. The second part of the thesis focuses on the production loss measure validity test and on deriving wage multipliers to estimate productivity costs. The thesis investigates productivity costs associated with sickness absence, sickness presenteeism and work environment-related problems from the employer's perspective.

Methods: In Study 1, a cost-effectiveness analysis was conducted among working adults (n =159) to compare the effect and costs of yoga for the prevention of LBP with strength exercises and evidence-based advice. This analysis was conducted from two perspectives, namely that of the employer and that of society at large. The outcome quality adjusted-life years (QALYs) were examined in a 12-month follow-up randomized controlled trial. In Study 2, a discrete choice experiment was conducted among working adults with LBP (n =112). This experiment used the conditional logit model to examine the influence of exercise attributes and individual characteristics on preferences for exercise to prevent LBP. In Study 3, using the Pearson Correlation and the Bland and Altman's test of agreement, the convergent validity of the Swedish health-related and work environment-related production loss measures (HRPL and WRPL respectively) were tested against the Health and Work Performance Questionnaire (HPQ) among working adults (n = 88). The HPQ is an extensively psychometrically tested and widely-used instrument. The Intraclass Correlation Coefficient (ICC) and Bland and Altman's tests of repeatability were used as tests of stability (n = 44). In Study 4, wage multipliers for managers (n = 758) were derived using an ordinal probit model to predict the costs of productivity loss from sickness absence, sickness presenteeism and work environment-related problems.

Results: The cost-effectiveness analysis in Study 1 demonstrates that yoga is less costly and improves quality of life (QALY) compared with strength exercises and evidence-based advice. For an additional QALY worth EUR 11,500 for society, yoga yielded a positive incremental net benefit of EUR 1,542 and EUR 2,860 compared with strength exercise and evidence-based advice respectively. Yoga could also be cost-effective, compared with evidence-based advice, if an employer considers that the improvements in QALY justify the additional cost of the intervention (EUR 150 per worker with LBP). These results only hold for those who adhered to the recommendations of exercising twice a week. The discrete choice experiment in Study 2 demonstrates that the most preferred exercise option was medium to high-intensity cardiovascular training performed in a group with trainer supervision at least once to twice per week. The most preferred types of incentive were exercise during work hours and wellness allowances (Friskvårdbidrag). The individual characteristic that most consistently influenced preferences for exercise was age. The convergent validity test of the production loss measures conducted in Study 3 showed moderately strong correlations (i.e. $r = 0.46$ for the HRPL and $r = 0.31$ for WRPL), as expected. The ICC for HRPL assessments was 0.90 and 0.91 for WRPL between the different testing occasions. This suggests that the test–retest reliability was good. Study 4 demonstrates that sickness absence, sickness presenteeism and work environment problems significantly impact team workers' productivity when job characteristics (i.e. teamwork, ease of substitution of workers and time sensitivity of output) are taken into account. To determine the economic implications for the employer in terms of the cost, it was estimated that the median wage multipliers from the sampled occupations in the study were 1.92 for sickness absence, 1.65 for health-related presenteeism due to acute illness, 1.58 for health-related presenteeism due to chronic illness, and 1.70 for work environment problems.

Conclusions: Studies 1– 4 gave rise to the following conclusions: a) Yoga may be considered a cost-effective early intervention for the prevention of LBP, but further investigations are warranted. b) Where preferences for exercise to prevent LBP are concerned, the most important factors were the age and exercise attributes such as type of exercise, frequency, level of supervision and incentives. This implies that providers and employers could improve participation in exercise for working adults with non-specific LBP by focusing on the exercise characteristics which are most attractive. c) The validity test of the HRPL and

WRPL suggests that the measures have convergent validity and good stability. This finding may have implications for advancing methods of assessing production loss as an outcome, which represents a major cost for employers. d) After job characteristics had been taken into account, the economic impact of sickness absence, sickness presenteeism and work environment-related problems on team productivity exceeded the cost of wages in a number of occupations. This implies that there could be productivity gains for employers if the cost of health and work environment-related productivity losses can be reduced.

Keywords: Low back pain, yoga, exercise, cost-effectiveness analysis, discrete choice experiment, preferences, work performance, productivity loss, production loss, RCT, intervention, secondary prevention, occupational safety and health, validity, reliability, sickness absence, sickness presenteeism, work environment, wage multipliers.

LIST OF SCIENTIFIC PAPERS

I. Aboagye E, Lohela-Karlsson M, Hagberg J, Jensen I. Cost-effectiveness of early interventions for non-specific low back pain: a randomized controlled study investigating medical yoga, exercise therapy and self-care advice. *J Rehabil Med 2015*, 47(2): 167-173.

II. Aboagye E, Hagberg J, Axén I, Kwak L, Lohela-Karlsson M, Skillgate E, Dahlgren G, Jensen I. Individual preferences for physical exercise as secondary prevention for non-specific low back pain: A discrete choice experiment. *Manuscript Submitted*.

III. Aboagye E, Jensen I, Bergström G, Hagberg J, Axén I, Lohela-Karlsson M. Validity and test-retest reliability of an at-work production loss instrument. *Occupational Medicine* 2016, 66(5): 377-82.

IV. Strömberg C, Aboagye E, Hagberg J, Bergström G, Lohela-Karlsson M. Estimating the effect and economic impact of absenteeism, presenteeism and work environment problems on reductions in productivity from a managerial perspective. Accepted for publication in *Value in Health* (2017).

CONTENTS

1 INTRODUCTION 9
2 Background..12
2.1 Prevalence, risk and burden of LBP 12
2.2 Economic evaluation 13
2.2.1 Identification of costs 13
2.2.2 Types of economic evaluation 14
2.2.3 Health-related quality of life (HRQL) 15
2.2.4 Quality-adjusted life year (QALY)..17
2.3 Discrete Choice Experiments (DCEs) 18
2.4 Productivity losses..20
2.5 Psychometric evaluations..20
2.6 Valuing productivity costs 21
3 Literature review of previous research 23
3.1 Effect and cost-effectiveness of yoga exercise for LBP 23
3.2 Preferences for exercise interventions for LBP 24
3.3 Psychometric properties of instruments used to measure productivity loss 25
3.4 The wage multiplier approach...26
4 Aims...28
4.1 Outline of the thesis..28
5 METHODS...29
5.1 Study design and sample size..29
5.2 Measures 31
5.3 Statistical analyses...35
5.4 Ethical considerations 39
6 RESULTS...40
6.1 Health-related quality of life among working adults with LBP 40
6.2 Cost-effectiveness of yoga for LBP 41
6.2.1 Complementary CEA analysis..42
6.3 Preferences for exercise among working adults with LBP 43
6.4 Convergent validity of the HRPL and WRPL measures against the HPQ work performance..45
6.5 Test-retest reliability of HRPL and WRPL measures 47
6.6 The impact on teamwork productivity of health and work environment-related problems 47
6.7 Wage multipliers for sickness absence, sickness presenteeism and work environment problems...51
7 DISCUSSION 55
7.1 Comparing findings with previous research..55
7.1.1 Cost-effectiveness of yoga intervention for LBP 55
7.1.2 Preferences for exercise among working adults with LBP.............56

7.1.3 Validity and test-retest reliability of an at-work production loss measure 57
7.1.4 The cost of productivity loss to the employer 58
7.2 General strengths and limitations 58
7.3 Implications for health policy/ actions 61
7.4 Suggestions for future research 62
8 CONCLUSION 65
9 Acknowledgements 66
10 References 67

LIST OF ABBREVIATIONS

MSDs	Musculoskeletal Disorders
HRQL	Health-Related Quality Of Life
DALYs	Disability-Adjusted Life Years
LBP	Non-specific Low Back Pain
CEA	Cost-Effectiveness Analysis
COI	Cost of Illness
QALYs	Quality-Adjusted Life Years
CUA	Cost-Utility Analysis
VAS	Visual Analogue Scale
SG	Standard Gamble
TTO	Time Trade-Off
TLV	Dental and Pharmaceutical Benefits Agency
ICER or ICUR	Incremental Cost-Effectiveness (Utility) Ratio
DCEs	Discrete Choice Experiments
RP	Revealed Preference
SP	Stated Preference
QOL	Quality of Life
HPQ	Health and Work Performance Questionnaire
VOLP	Valuation of Lost Productivity
WPAI	Work Productivity and Activity Impairment Questionnaire
OMPSQ	Örebro Musculoskeletal Pain Screening Questionnaire
OHS	Occupational Health Services
HRPL	Health-Related Production Loss
WRPL	Work Environment-Related Production Loss
INB	Incremental Net Benefit
EUR	Euros
SEK	Swedish Kronor

1 INTRODUCTION

In this thesis, value assessment frameworks were applied to non-specific low back pain (LBP) in a working population to evaluate health outcomes and to estimate productivity costs. Interventions to prevent LBP need to be properly evaluated from a variety of perspectives, including that of the worker, the employer and society at large, in order to help us better understand how to support worker health policies and a sustainable working life. The interest in sustainable working life in this thesis focuses on the economic perspective of work and health where individual workers are seen as a resource and preventive activities in the workplace can be provided to improve worker health (Ahonen 2015).

Non-specific Low Back Pain (LBP) is a common musculoskeletal disorder among workers. LBP is often recurrent and at times persistent (Axen and Leboeuf-Yde 2013). According to European data sources, up to 25% of workers reported back pain in 2005 (EU-OSHA 2010). LBP can represent a considerable burden to workers in terms of disability which impacts their productivity and can cause sickness absence (EU-OSHA 2010). Research has shown that the burden of LBP among workers is mainly caused by work-related factors (EU-OSHA 2010, Punnett et al. 2005). It is an economic burden for employers and society at large because it reduces productivity (Murray et al. 2012).

The prevalence of LBP among workers and its considerable impact on working life make it important to take action to prevent recurring, disabling back pain. In other words, secondary prevention measures that aims to reduce the more serious consequences of LBP are needed (Balague et al. 2012).

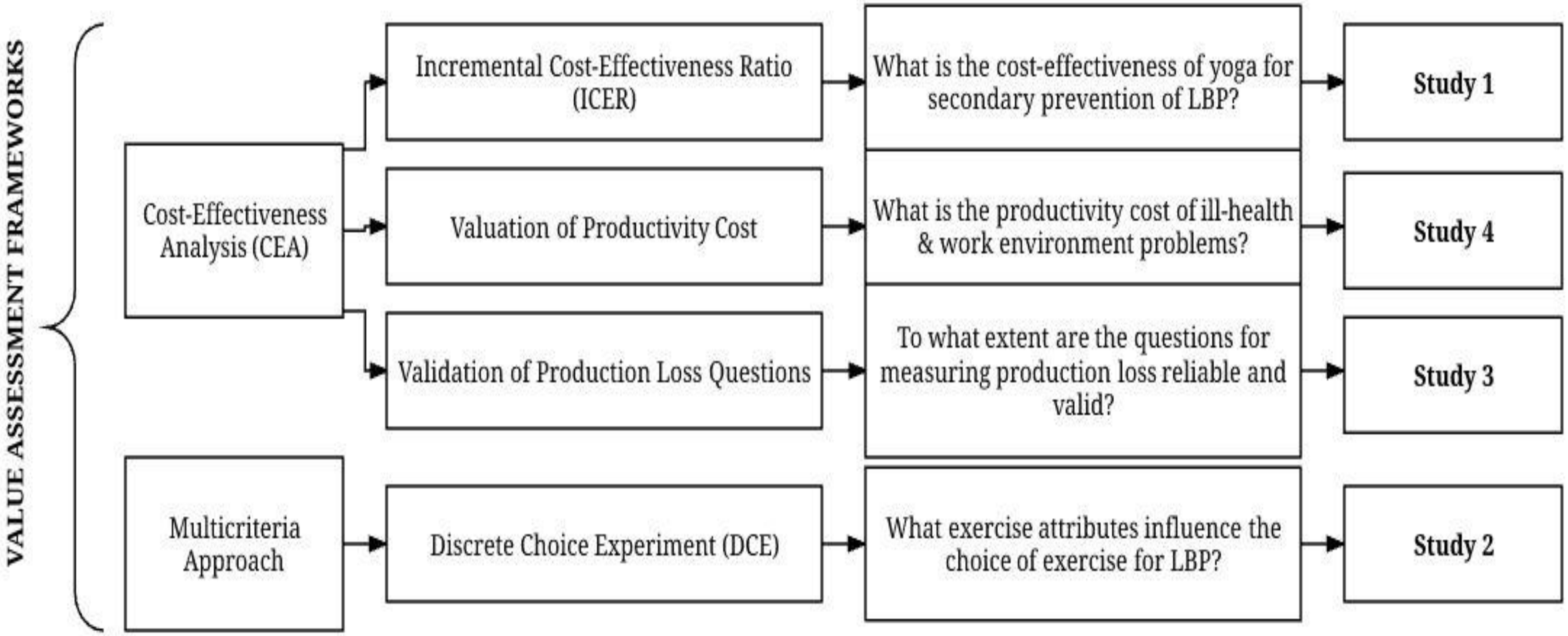
The first part of this thesis evaluates exercise interventions that are designed to prevent LBP among working adults. In order to evaluate the cost of and preferences for exercise, this thesis use the preference-based measure EuroQol-5D (EQ-5D) and self-reported preference measures for exercise interventions to prevent LBP (Chapman et al. 2011).

Global comparative studies have found that the most significant costs of LBP are indirect ones such as productivity losses due to sickness absence and sickness presenteeism (i.e. going to work when ill) (Bevan et al. 2009, Wieser et al. 2011, Lambeek et al. 2011). Lamers et al. (2005) found an association between low quality of life and reduced productivity and sickness absence in persons with LBP. The cost of productivity losses due to sickness absence and reduced productivity while at work can be an incentive for decision-makers (particularly employers) to adopt preventive actions to address their workers' health problems. However, because of the challenge of measuring productivity loss, cost calculations are rare in cost-effectiveness analyses (Lensberg et al. 2013).

The second part of this thesis evaluates the psychometrics of a production loss instrument intended to measure reduced work performance while at work (sickness presenteeism). It further estimates the extra costs associated with productivity loss (due to sickness absence, sickness presenteeism and work environment problems) that employers may have to bear over and above wage costs.

An overview of the research questions and the overall research framework are presented in Figure 1. The CEA framework was used to evaluate costs and effectiveness of interventions (Doshi and Willke 2017). Exercise preferences were investigated by means of a multi-criteria approach (Marsh et al. 2016).

Figure 1. Overview of the value assessment frameworks and the research questions



2 BACKGROUND

Before looking at just why it is essential to properly evaluate interventions for LBP in terms of the costs and preferences, the thesis will start with a more general discussion about LBP as a major health problem. This is followed by a presentation of the general methodologies of the kinds of economic evaluation that can be used to evaluate interventions. The main question of how to estimate productivity costs for economic evaluations of health interventions is also discussed.

2.1 Prevalence, risk and burden of LBP

LBP is a common diagnosis in primary health care; it affects 60% to 80% of the general population at some point in their lifetime (Hoy et al. 2014). LBP is commonly classified as acute if it lasts less than six weeks, subacute up to three months, and chronic when it lasts more than three months (Leboeuf-Yde et al. 2013). However, recent research has shown that it is common to experience a chain of episodes during the lifetime. In other words, LBP is a recurrent and persistent condition (Axen and Leboeuf-Yde 2013). Only in about 10% of cases is it possible to establish a specific cause for LBP, and thus the majority of people have "non-specific" LBP (Savigny P et al. 2009). In this thesis, LBP is the non-specific type unless stated otherwise.

The risk of developing LBP can increase with exposure to certain work-related factors (Sundelin et al. 2015). Work-related physical factors such as bending, twisting and prolonged sitting are associated with LBP (Sterud and Tynes 2013). Psychosocial factors at work such as low social support, a high perceived work load, time pressure, low job control, perceived stress and high psychological job demands also increase the risk of LBP (Sundelin et al. 2015). Individual biological (age and gender) as well as psychosocial factors (eg. coping and social support) have also been shown to be risk factors for LBP (Jones et al. 2005, Hansson and Jensen 2004).

LBP is a known cause of sickness absence, work disability and high healthcare utilization in the Swedish workforce (Hansson and Jensen 2004, Ekman et al. 2005). A study that compared sick leave patterns in different musculoskeletal disorders (MSDs) in Sweden found that LBP was associated with short-term sick leave as well as contributing to a significant share of recurrent sick leave (Hubertsson et al. 2014). Visits to health practitioners, out-of-pocket costs, and self-prescribed medications were shown to be higher for individuals with LBP than for the general population (Joud, Petersson and Englund 2012).

Society is affected by the direct healthcare costs of working adults with LBP and by the indirect cost of their reduced work capacity (Woolf and Pfleger 2003). According to the Swedish Social Insurance Agency (SSIA), LBP costs amounted to just above 7% of the total health insurance costs paid in sickness benefits (Lidwall 2011). The indirect costs of LBP, caused by productivity loss due to sickness absence and reduced work performance, have been shown to be substantially higher than the direct cost of LBP (Persson et al. 2015, Ekman et al. 2005). The higher costs and utilization of health care suggest that effective preventive interventions for LBP could have a significant cost-saving potential.

2.2 Economic evaluation

Health economic evaluation is a comparative analysis of at least two alternative interventions that compete for scarce resources (Drummond et al. 2005). It is not possible to determine whether an intervention should be discontinued or continued without comparing its costs and consequences with those of another intervention. The aim of an economic evaluation of health interventions is therefore to provide information about whether the additional resources required for improving health are justified (Drummond et al. 2005). Thus, an economic evaluation complements the regulation of health interventions by agencies such as the National Board of Health and Welfare and the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) which check the safety and efficacy of interventions in order to prevent harm in society.

2.2.1 Identification of costs

The cost of resource use is a core aspect of the economic evaluation of health interventions. The cost components measured are categorized as direct, indirect and intangible costs (Drummond et al. 2005). The direct costs are related to resources used, such as healthcare, medication and out-of-pocket expenses. Indirect costs, commonly referred to as productivity losses, are costs related to sickness absence and reduced work performance while at work

(Drummond et al. 2005). Some studies indicate that including productivity costs in economic evaluations provides valuable information for decision-makers (Jonsson 2009). Intangible costs include reduced quality of life and function that result from an illness (Lubeck 2003). Intangible costs are seldom included in economic evaluations of healthcare interventions (Tarricone 2006).

In order to calculate the cost associated with resource use in economic evaluations, a 'top-down' or a 'bottom-up' costing approach can be used (Drummond et al. 2005). The 'bottom-up' costing approach is preferred because it collects individual-level micro costs and more precise data from registers or diaries. The 'top-down' costing approach uses aggregated cost data which may not be precise. The first step in calculating cost is to determine which cost items should be included. This, in turn, depends on the perspective of the analysis. The second step is to measure the quantities of the resources used. The third step is to estimate cost using prices or the opportunity cost. Opportunity cost is the cost of the other alternative resource not used. Listed market prices are commonly used instead of the opportunity cost because of distortions in the market such as taxes and subsidies. Lastly, the quantities of resource used are multiplied by the price within the timeframe to obtain the direct cost associated with the illness.

2.2.2 Types of economic evaluation

Besides the cost component, it is the benefits (also referred to as outcome, consequences or health effects) of the intervention that determine which method should be used in economic evaluations. The available methods are cost-minimization analysis (CMA), cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA). Figure 2 gives a summarized description of the different types of economic evaluation (Drummond et al. 2005).

Figure 2. Types of economic evaluation

Types of economic evaluation	**Measuring and valuing costs**	**Identifying effects**	**Measuring and valuing effects**
Cost-minimization analysis (CMA)	Monetary units	Assumed to be equal	None
Cost-benefits analysis (CBA)	Monetary units	One or multiple outcomes	Monetary units
Cost-effectiveness analysis (CEA)	Monetary units	One or multiple outcomes combined in one measure	eg. prevented cases, life years saved etc.
Cost-utility analysis (CUA)	Monetary units	One or multiple outcomes combined in one measure	eg. QALYs, DALYs or other preference-based measures

All the methods include monetary unit as costs in the analysis, but the unit by which the benefits are measured can differ. A CMA compares only the cost of interventions because the method assumes that intervention benefits are equal. The CBA compares both the cost and the benefits of interventions in monetary terms. The CEA uses natural units as an outcome for example prevented cases and years of life saved. The possibility of comparing outcomes across diseases is eroded in the CEA because outcomes may not be commensurate. The preferred type of economic evaluation that allows comparison of outcomes across different interventions is the CUA (Drummond et al. 2005). The CUA is a special class of cost-effectiveness analysis in which the benefits are expressed in terms of Quality-Adjusted Life Years (QALYs).

2.2.3 Health-related quality of life (HRQL)

The core domains of health-related quality of life (HRQL) encompass the physical, social and emotional aspects of an individual in relation to his or her prior and current state of health (Santana and Feeny 2008). For the purpose of evaluating interventions there are a large number of HRQL assessment tools with a wide-ranging coverage of health dimensions. Disease-specific assessment tools cover dimensions that are relevant for a particular disease while the generic tools (health profiles and preference-based measures) are multidimensional, covering a number of domains such as pain and physical, social and emotional function.

There are two types of preference-based measure, namely the direct and the multi-attribute preference instruments. The direct preference-based measures assess preferences for health outcomes using various methods, including the visual analogue scale (VAS), the standard gamble (SG) and the time trade-off (TTO) (Drummond et al. 2005). The VAS measures preferences for ordinal ranking of state of health, with worst being '0' on a scale and perfect health being '100'. SG measures cardinal preferences for health outcomes by offering individuals two alternatives and asking them to make a choice or to "gamble''. Preference is measured on an interval scale on which perfect health for t years equals 1 and dead equals 0 (Drummond et al. 2005). The first alternatives consist of a desirable outcome, for example perfect health, for t years followed by death, with probability p, and a less desirable outcome, such as immediate death, with a probability of 1 - p. The probability p is varied until the individual is indifferent between the two alternatives, so that the preference score for state i is p, i.e., $u_i = p$. The TTO also measures cardinal preferences for state of health, whereby individuals are offered two alternatives: health state i for time t followed by dead and perfect health for time $x < t$ followed by dead. Here, time x is varied until the individual is indifferent between a short period of perfect health and a longer period of impaired health, such that the preference score for state i is $h_i = x/t$ (Drummond et al. 2005). The advantage of the direct preference-based approach is that it captures how the individual values positive intervention effects in relation to the negative side effects on their HRQL. However, a comparison of the methods has shown that SG yields higher HRQL values than the TTO, which in turn yields higher HRQL values than the VAS (Morimoto and Fukui 2002).

The most common of the multi-attribute preference instruments are the EuroQol EQ-5D (Kind et al. 2005, EuroQol group 1990), the six-dimensional health state short form (SF-6D) (Brazier, Roberts and Deverill 2002), the Health Utility Index (HUI2 and HUI3) (Feeny et al. 2002, Torrance et al. 1996) and the 15D (Sintonen 2001). These instruments use a multi-attribute health status classification system and a scoring system to value health status. For example, the EQ-5D-3L has five attributes (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) described with three levels per attribute which generate 243 possible states of health. Devlin and Krabbe (2013) added two additional levels to the three levels in the EQ-5D-3L, thereby making it EQ-5D-5L.

Beside the selection of a preference assessment method, there is also the issue of discrepancy between state of health evaluations of individuals experiencing the health state and of the general public. The discussion about whether preference values should come from individuals who are actually experiencing the state of health in question (experience-based values), or from individuals to whom the states of health are described (population or hypothetical values) has not yet been resolved (Aronsson et al. 2015). Those in favour of the use of experience-based values contend that hypothetical values from healthy individuals may fall short of the values that individuals experiencing the health state give (Burstrom et al. 2014). Empirical tests have shown that experience-based HRQL values tend to be higher than hypothetical values (Aronsson et al. 2015, Mann, Brazier and Tsuchiya 2009). Others argue that for publicly-funded healthcare, population-based values are more relevant than experience-based ones. In Sweden, the Dental and Pharmaceutical Benefits Agency (TLV) prefers experience-based values to hypothetical values for the same reason given by proponents of experience-based values (TLV 2003).

2.2.4 Quality-adjusted life year (QALY)

QALY is used in economic evaluations as a single generic measure primarily to correct an individual's life expectancy based on the levels of HRQL (Drummond and McGuire 2007, Drummond et al. 2005). Thus QALY takes both the quality and the quantity of life experienced over time into consideration. To calculate QALYs, an individual's life years are multiplied by the quality of life or preference values for the same years. The values derived are commonly referred to as QALY weights. QALY weights are values that describe states of health. The value for full health (or best imaginable state of health) is anchored at '1', while that for the state 'dead' is anchored at '0'. One year in perfect health thus equals 1 QALY, while one year in the state 'dead' equals 0 QALYs. Values for all other health states between '1' and '0' are determined on the basis of their position relative to full health and death. States of health 'worse than dead' are assigned values below zero. The use of QALYs as the unit of effectiveness ensures a standardized way of comparing the incremental cost per QALY gained for different health interventions in a so-called reference case analysis. A reference case analysis must include all relevant costs (regardless of who pays them), thus giving a societal perspective.

To decide which the more cost-effective intervention in a CEA is, the differences in cost and the differences in effect between the interventions under comparison are divided. The outcome, based on this comparison of costs and effects, is simplified by calculating the Incremental Cost-Effectiveness (Utility) Ratio (ICER or ICUR), which takes the form:

$$(Cost_2 - Cost_1) / (QALY_2 - QALY_1) = \Delta Cost / \Delta QALY = \text{ICER or ICUR}$$

The ICER can be interpreted as cost per gained QALY of intervention 2 compared to intervention 1. If the result of a CEA is cost-effective, the ICER means that the extra intervention costs are acceptable per unit of health outcome compared to other interventions or to some threshold value, which is assumed to be how much decision makers are willing to pay for the additional unit of effect. In order to determine if the effect of the intervention justifies the extra cost, the ICER is compared with societal willingness-to-pay between GBP 50,000 – 100,000 (Ryen and Svensson 2014). This value varies from country to country. Alternatively, the form of the ICER or ICUR can be rearranged to fit a net-benefit framework in order to make decisions based on the incremental net benefit in monetary terms (Drummond et al. 2005, Stinnett and Mullahy 1998).

2.3 Discrete Choice Experiments (DCEs)

DCEs are another approach to economic evaluation to provide measures of benefit valuation of alternative health interventions and for health-policy decision-making (Tockhorn-Heidenreich, Ryan and Hernández 2017). One of the useful pieces of information that DCEs can provide to economic evaluations for decision-making purposes is what and how important certain characteristics (attributes) are for individuals. DCEs can also be used to predict adherence rates and the acceptability of interventions (Louviere and Lancsar 2009).

DCEs are based on the assumption that health interventions can be described by their attributes and that how individuals value an intervention depends on these attributes (Gerard, Ryan and Amaya-Amaya 2008, Tockhorn-Heidenreich et al. 2017). This assumption draws on standard economic theory of value and consumer behaviour which assumes that individuals are rational decision makers and maximize their preferences subject to constraints (Gerard et al. 2008). Further assumptions are that there are a set of finite and mutually exclusive alternatives to choose from; that choice behaviour is inherently random (random utility model); and that individuals are capable of discriminating between alternatives (Manski 1977, McFadden 1986).

In health care, the techniques used to examine how individuals' value benefits are either revealed preference (RP) or stated preference (SP) (Gerard et al. 2008). The RP is an indirect means of exploring individual values of health benefits usually by observation. The SP technique involves asking individuals to state their preferences by presenting them with two or more competing options (choice sets or alternatives) that vary along several attributes (Ryan and Gerard 2003). Preferences are measured according to patterns of choosing or responding when presented with two or more alternatives and often require choices to be made (Martin et al. 2006). The key steps involved in conducting a DCE can be summarized as follows: conceptualizing the choice process; deciding on the appropriate levels for the attributes; constructing an experimental design (Reed Johnson et al. 2013); designing questionnaires and pilot testing choice questions (Bridges et al. 2011, Bech, Kjaer and Lauridsen 2011); sampling and sample size (de Bekker-Grob et al. 2015); data collection, coding of data (Bech and Gyrd-Hansen 2005); econometric analysis (Hauber et al. 2016); and interpretation of DCE results as depicted in Figure 3. Further details of these steps can be found elsewhere (Amaya-Amaya, Gerard and Ryan 2008, Bridges et al. 2011).

Figure 3. Key steps in conducting a DCE

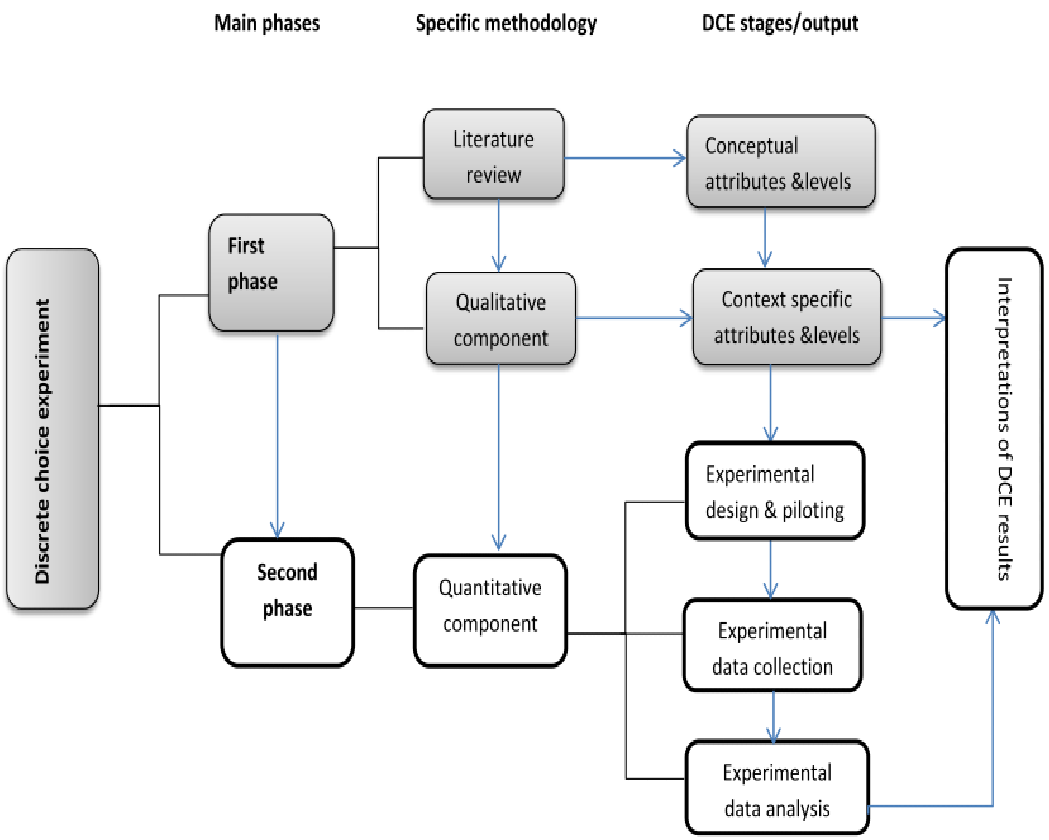


Figure 3 was reproduced with permission from Abiiro et al. (2014).

2.4 Productivity losses

Productivity losses are costs resulting from reduced work performance and replacement costs due to illness and disability in working adults (Krol, Brouwer and Rutten 2013). Studies of worker health interventions describe productivity loss as a work-related outcome that arises from two main causes (i.e. sickness absence or sickness presenteeism), measured and valued as costs to the employer (Uegaki et al. 2011).

The total number of work hours lost because of health problems is referred to as sickness absence (Hensing 2009). Sickness absence data is used as a measure of health among working individuals.

Sickness presenteeism is the phenomenon of being at work when ill (Johns 2010). Sickness presenteeism can also be defined as the decrease in on-the-job performance due to health problems (Schultz, Chen and Edington 2009). Sickness presenteeism can occur before and/or after the health problem have required a period of absence to cause reduced work performance. It is therefore important to take sickness presenteeism into account when evaluating interventions aimed at reducing sickness absence, because presenteeism mostly precedes future sickness absence. It can, however, in itself also cause reduced work capacity (Bergstrom et al. 2009, Cancelliere et al. 2011). Unfortunately, reduced productivity due to sickness presenteeism is often not included in economic evaluations, even though it can be an important cost for employers (Kigozi et al. 2017).

Sickness absence and sickness presenteeism are not the only causes of productivity loss. Earlier research has shown that work environment problems have a greater impact on productivity than does sickness presenteeism, even though they are not as prevalent (Karlsson et al. 2013).

2.5 Psychometric evaluations

In order to able to use a health outcome measurement instrument in research or clinical practice, its measurement properties such as reliability, validity and responsiveness, should be assessed adequately (Cohen, Swerdlik and Philips 2013, Mokkink et al. 2010).

The most commonly reported measurement properties in psychometric evaluations are reliability and validity (Cohen et al. 2013). A reliability test can be performed to assess how an instrument consistently measures a construct across individuals or to assess the stability in the responses of the same individuals. For instance, the consistency of repeated measures using the same device can be assessed with the Pearson correlation coefficient. This is often called test-retest reliability (Cohen et al. 2013, Bruton, Conway and Holgate 2000). The most commonly used index of test-retest reliability, the intra-class correlation (ICC), is the ratio of variance of measurements of a given target to the variance of all targets. This index reflects both the degree of consistency and stability in the responses of the same individuals.

The validity of a measure refers to the extent to which it measures what it is designed to measure (Cohen et al. 2013). There are different forms of validity and a range of validity tests that can be performed. The basic validation methods, which include content, face and factorial validation, provide an understanding of what is measured based on reason and evidence. Convergent validity tests whether constructs that should theoretically be related are in fact related, using two different measures (Koeske 1994). The concept of construct validity incorporates all other types of validity mentioned above (Cohen et al. 2013).

2.6 Valuing productivity costs

Productivity costs in economic evaluation can be estimated using the Human Capital (HC) approach or the Friction Cost method (FC) (Krol and Brouwer 2014). By the HC approach, productivity costs are estimated using the gross earnings of individuals independent of when the reduction (due to sickness absence or presenteeism) occurred (Krol and Brouwer 2014, Johannesson 1996). With the FC method, productivity costs are estimated on the basis of the time and resources used to find a replacement for the period the worker is absent (Koopmanschap et al. 1995, Krol and Brouwer 2014). The HC approach has an individual perspective while the FC method takes the employer perspective (van den Hout 2010). The HC approach is grounded in economic theory incorporating a broader view of lost productivity, particularly lost production in paid employment, lost non-work time and even sometimes the time of informal care-givers (Lensberg et al. 2013).

Notwithstanding the fact that the two methods produce different results (van den Hout 2010), it has been shown that productivity costs derived from the HC approach could significantly determine the cost-effectiveness of interventions. Thus, if it is accounted for in the economic evaluation of interventions it could provide additional information of benefit to decision-makers which might lead to different resource-allocation decisions (Krol et al. 2016). Despite the many arguments in favour of including productivity costs in economic evaluation of health interventions (Krol et al. 2016, Jonsson 2009), methods to value productivity costs are not without their difficulties. The minimum gross wage, which indirectly measures the value of marginal productivity (Lensberg et al. 2013), is used to calculate the costs related to production loss using the HC approach. However, the substantial differences at company-level due to factors such as company characteristics, regulations and contracts could influence wage and productivity differentials for workers with different characteristics and work dynamics (Fox 2009, Card, Lemieux and Riddell 2004, Koeniger, Leonardi and Nunziata 2007).

3 LITERATURE REVIEW OF PREVIOUS RESEARCH

This section presents short summaries of the current state of knowledge concerning the areas studied in this thesis. The summaries are based on the most recent systematic reviews. Where no reviews exist, original papers are used.

3.1 Effect and cost-effectiveness of yoga exercise for LBP

LBP affects many domains of the health of workers, such as self-efficacy, disability and pain. This impacts profoundly on their quality of life (Montazeri and Mousavi 2010). Accordingly, quality of life should be considered a clinically relevant back pain outcome in clinical trials when investigating LBP (Montazeri and Mousavi 2010, Rasmussen-Barr et al. 2012).

Many interventions for the prevention of LBP exist, but their effect are not long-lasting due to the high recurrence rate of LBP (Steffens et al. 2016). An overview of the effectiveness of prevention strategies for non-specific LBP concluded that exercise intervention could reduce the number of recurrent episodes of LBP (Steffens et al. 2016). Other interventions such as education were effective when combined with exercise to prevent LBP. Recent evidence continues to support the effectiveness of exercise interventions for the prevention of LBP (Chou et al. 2017, Steffens et al. 2016).

Yoga, a complementary and alternative (CAM) exercise strategy, has been used as secondary prevention measure for LBP. In a recent meta-analysis it was concluded that yoga can be recommended for chronic LBP (Cramer et al. 2013). Furthermore, research findings have shown that yoga provides moderate improvements for back-related function compared to other types of exercise or a do-nothing approach (Wieland et al. 2017). Previous studies (as shown in the reviews) have examined the effects of yoga for treating chronic LBP, but there is a lack of clear evidence whether it is possible to prevent LBP with yoga exercise.

Few studies have looked at improving work-related outcomes such as sickness absenteeism, sickness presenteeism and work performance using yoga. One recent study showed that yoga, compared with strength training or evidence-based advice, achieved no difference in effect on sickness absenteeism and sickness presenteeism (Bramberg et al. 2017). Little research has been carried out into the effectiveness of yoga for the quality of life of individuals with LBP and its cost-effectiveness (Chuang et al. 2012, Tekur et al. 2010). The cost-effectiveness of yoga for LBP with regard to quality of life should be investigated (Andronis et al. 2017).

3.2 Preferences for exercise interventions for LBP

Exercise interventions have been shown to be effective for prevention of recurrent LBP (Steffens et al. 2016). This notwithstanding, it may be challenging for individuals with LBP to adhere to the interventions which have been recommended to them to improve their state of health. Certain attributes of exercise interventions, such as type of exercise, exercise intensity and trainer-supervised exercise, were found to improve adherence to exercise for LBP (Jordan et al. 2010). Previous reviews of exercise interventions have also shown that supervised exercise that includes aerobic capacity, strengthening, endurance and coordination is effective in improving pain and reducing sickness absence in individuals with LBP (Tveito, Hysing and Eriksen 2004, Hayden et al. 2005). However, the findings are inconsistent with regard to whether exercise should be supervised or not, whether it should take the form of individually-designed sessions or group sessions, and what the optimal exercise intensity should be (Henchoz and Kai-Lik So 2008).

Financial incentives for exercise have also been shown to be effective in improving participation rates (Giles et al. 2014). There is however little evidence regarding the relative importance of incentives provided by the employer (as a paying vehicle) to motivate exercise among working adults with LBP.

Apart from the attributes of exercise interventions, individual characteristics (such as lifestyle and attitudes) have been found to be associated with acceptance and compliance rates for physical activity recommendations (Scheers, Philippaerts and Lefevre 2013). Further, it has been shown that determinants of involvement in physical activity include factors such as social support, stages of change, general physical health and prior adherence to physical activity (Wendel-Vos et al. 2007, Koeneman et al. 2011).

The similarity of the effects of exercise interventions with varying characteristics may also suggest that it is warranted to take individual preference into consideration in health policy decision-making (Tveito et al. 2004, Hayden et al. 2005). Research into individuals' preferences for exercise as secondary prevention for LBP is scarce. It is therefore imperative to examine the preferences of individuals with LBP to support adherence to exercise recommendations. This could enhance evidence-based practice and adherence to recommendations, and hence health outcomes (George and Robinson 2010, Preference Collaborative Review 2008).

3.3 Psychometric properties of instruments used to measure productivity loss

Objective measures of workers' productivity are often not available. To obtain such data one has to rely on subjective productivity measures, largely in the form of questionnaires (Krol and Brouwer 2014). Productivity loss instruments must therefore be reliable and valid.

There are many available instruments for measuring productivity loss due to sickness absence and sickness presenteeism. Each instrument has its own comparative advantages with respect to the aspects of productivity changes it measures, the recall period used, the strength of its psychometric properties, and whether it has the potential for estimating costs related to lost productivity for paid or unpaid work (Tang 2015). The methodological quality and extent of psychometric testing of productivity loss instruments have been criticised. In some tests, the choice of comparator instruments used for validity testing was not meaningful and the expected level of correlation between an instrument and its comparator(s) was not specified (Tang 2015, Mokkink et al. 2010). Furthermore, few longitudinal tests of validity and test-retest reliability have been performed on these instruments (Tang 2015).

According to Tang (2015), psychometrically-sound instruments for estimating costs related to productivity loss include the Health and Work Performance Questionnaire (HPQ), the Health and Labor Questionnaire, the Health-Related Productivity Questionnaire Diary, the Productivity and Disease Questionnaire, the Valuation of Lost Productivity (VOLP), the Work Productivity and Activity Impairment Questionnaire (WPAI) and the Work Productivity Short Inventory. There is a lack of production loss instruments in Swedish. Recently, a question from the WPAI was translated and initially tested in Swedish (Karlsson et al. 2013).

Productivity loss instruments capture reduced production resulting from health problems. However, in addition to health problems, reduced worker productivity may be caused by work-related factors such as the work environment (Brooks et al. 2010). There is a shortage of this type of measure. In order to fill this gap, an instrument that takes production loss due to work-related factors into account was developed in Swedish by modifying a question from the WPAI (Karlsson et al. 2013).

3.4 The wage multiplier approach

The shortcomings of using wages as an alternative to worker's marginal productivity have prompted the suggestion that new ways of valuing productivity costs are needed. The wage multiplier approach assumes that a period of sickness absence or presenteeism will invariably affect a worker's productivity and that the economic impact could be greater than the worker's wage. Thus, using the minimum gross wage to estimate productivity loss tends to underestimate it (Pauly et al. 2008, Pauly et al. 2002). Particularly in a teamwork setting, wage multipliers have been seen to reflect the real productivity cost to employers (Pauly et al. 2008, Zhang et al. 2015). Authors have argued that at company level, such as when an individual is instrumental in the tasks to be completed by a team or when outputs have been inordinately time-sensitive, or when it is difficult to find a replacement for a worker, productivity losses are usually greater than the minimum wage loss to the employer (Pauly et al. 2008, Zhang et al. 2015). Thus, the general model of Pauly et al. (2008) examines the magnitude and incidence of costs associated with sickness absence or presenteeism under different assumptions about the company's characteristics. In the event that compensation (i.e. recovering work loss after returning) occurs, using wages to estimate productivity costs becomes even more problematic. The influence of the compensatory work done by the returnee worker, by other colleagues or a temporary worker has been shown to be relevant for productivity costs (Knies et al. 2013). Insofar as these important company level characteristics are influential determinants of the value of productivity costs, it is the employer's perspective that might reflect the true value of productivity problems.

As in the case of sickness absence and presenteeism, one might hypothesize that productivity costs due to work environment-related problems may be higher than the wages. Wage multipliers that take into consideration the effect of problems related to the work environment on the total cost of lost productivity are needed. Previous wage multiplier estimates have estimated their productivity cost on the basis of responses from managers who may only have responded from a study-specific context (i.e. against a background of a particular economic system, including specific labour market and insurance regulations). Thus, country-specific wage multipliers which are applicable to the particular setting in question are needed.

4 AIMS

The overall purpose of this thesis is to contribute new knowledge which can help decision-making that support a sustainable working life. The targeted participants in the studies were thus employers or employees of working age. The frameworks used in the first part of this thesis aimed to investigate the cost-effectiveness of and preferences for secondary prevention interventions for LBP. The second part of the thesis focuses on the production loss measure validity test and on deriving wage multipliers to estimate productivity costs. The productivity costs associated with sickness absence, sickness presenteeism and work environment-related problems were investigated from the point of view of the employer.

The specific aims of the studies included in this thesis were as follows:
Study 1: To evaluate the cost-effectiveness of yoga as an early intervention for LBP compared with strength exercises and evidence-based advice.
Study 2: To examine specific attributes of exercise and their influence on individual preferences for exercise among working adults with LBP. The relative importance of these attributes for choice of exercise was also investigated.
Study 3: To examine the convergent validity and test-retest reliability of the Health-Related Production Loss (HRPL) and the Work Environment-Related Production Loss (WRPL) measures against the Health and Work Performance Questionnaire (HPQ).
Study 4: To derive wage multipliers to estimate the costs associated with productivity loss due to sickness absence, sickness presenteeism and work environment-related problems from a managerial perspective.

4.1 Outline of the thesis

This thesis is based on research articles published in, or submitted to, scientific peer-reviewed journals. Section 5 presents an overview of the methods considered in this thesis. It also presents the study design and sample, outcomes and statistical analyses used. Section 6 highlights the most important findings of the studies. Section 7 brings together the results and offers an extended discussion of the results and the implications of the findings.

5 METHODS

5.1 Study design and sample size

Study 1 was a randomized controlled trial (RCT) investigating kundalini yoga compared with strength exercise and evidence-based advice. The RCT was registered in the Clinical Trials protocol (NCT01653782). A block randomization design was used with the random allocation sequence computer-generated by a statistician. For each participant an opaque envelope was opened, in consecutive order, by an external research assistant not involved in the inclusion process. The allocation was concealed to all staff involved in the inclusion process. The participants did not know the detailed content of the different interventions. The yoga instructor and physiotherapist were not blinded. The research group assessing the study's outcome was blinded during the data collection and data analysis.

All participants received a minimal intervention comprising self-care advice and were encouraged to do home practice by performing the instructions given to them at least twice a week. Participants in the control group received only the minimal intervention. Those in the active intervention groups received kundalini yoga or strength exercises for six weeks. The 60-minute yoga classes were given by an experienced yoga instructor twice a week. The strength exercise used as an active comparator was an individually-tailored, supervised strength-training session (five times) led by an experienced physiotherapist. An extensive description of the content of the intervention given to participants can be found in Bramberg et al. (2017).

Subjects were recruited via the Occupational Health Services (OHS) in Stockholm County (n = 8) and by advertisements in the local media (n = 302). Participant recruitment and follow-up were conducted from April 2010 to June 2012. Those who responded to the invitation were screened using the Örebro Musculoskeletal Pain Screening Questionnaire (OMPSQ) (Linton and Halldén 1998) and were physically examined if they scored 90 points or more, i.e. if they fulfilled the requirements for the psychosocial risk profile. A cut-off score of ≥ 90 points was used because it has been shown that the risk of long-term sick leave due to back pain increases above this point (Linton and Boersma 2003). Further inclusion criteria were non-specific LBP with or without neck pain; 18–60 years of age; not on sick leave or on sick leave of less than eight weeks' duration; sufficient understanding of Swedish. The exclusion criteria were spinal pathology (e.g. tumours or spinal fractures); continuous ongoing sicklisting ≥ 8 weeks; comorbidities that could affect the ability to fully participate in the study (e.g. physical disability, severe mental illness); existing weekly yoga practice or strength training; and pregnancy. The calculated sample size was 40 participants per group to detect an effect of 25 % in changes in the primary endpoint sickness absence with 80 % power.

Study 2 used a DCE design to examine individual preferences for exercise as a secondary prevention for LBP. Study participants were patients consulting for LBP. They were recruited consecutively through a network of OHS providers and primary care clinics including physiotherapists, chiropractors and naprapaths who treat LBP. The recruitment took place in Sweden from March 2015 to March 2016. The inclusion criteria were: having LBP as the primary pain site; age range 18–65; currently not on sick leave for more than 14 days. The exclusion criteria were: back pain of specific origin or disease; prior spine surgery; severe comorbidities (e.g. physical disability, severe mental illness) that could affect the ability to perform exercises; insufficiency in Swedish.

The clinicians recruiting subjects for the study were given verbal and written instructions about what to tell subjects and how to determine inclusion or exclusion. A checklist was provided to be filled in by clinicians at inclusion. This was later sent to the research team. Based on Orme (2006), suggestions for estimating sample size in DCEs given by $n \times t \times a/c \geq 500$ were used. Where n = number of expected respondents to be estimated, t = number of choice tasks were given as 10, a = number of alternatives per task was given as two, and c = maximum number of levels was given as four. In this thesis, $n \geq (500 \times 4 / (2 \times 10)) = 100$. The sample size (n) in Study 2 was estimated to ≥ 100.

Study 3 is a cross-sectional study to test the convergent validity and test-retest reliability of the Health-related production loss (HRPL) and work environment-related production loss (WRPL) measures. Recruitment of subjects was conducted by sending information letters that described the study to a convenience sample who were part of a network of researchers and practitioners located in Stockholm County. To be included in the study, participants had to be working, not on sick leave and be in the age range 18–65. Those who fulfilled the inclusion criteria and consented to participate by responding to the e-mail represented the total sample in this study. In total, 161 individuals were invited, of whom n = 88 individuals agreed to participate. Data collection started in May 2014 and was completed in June 2014. Data were collected on two occasions to conduct the test-retest analysis. The interval between the test and retest assessment was one day (24 hours) to ensure that respondents were using the same time frame stated in the questions (7 days).

Study 4 was a cross-sectional study which estimated the effect and economic impact of sickness absenteeism, sickness presenteeism and work environment problems on reduced productivity. The following two inclusion criteria were applied: 1) manager in a medium to large company; 2) manager with operational responsibility for five to 50 employees. Medium and large companies were chosen so that as many workers with homogeneous job characteristics in teamwork supervised by the managers could be included. It was decided that a team should consist of at least five employees, while more than 50 employees were too many for a manager to properly assess their workers' productivity.

To obtain a range of job categories in the sample, at least 30 occupations with approximately 15 respondents in each job category were required to estimate the economic impact of reduced productivity caused by sickness absenteeism, sickness presenteeism and work environment problems. This sample estimate was based on suggestions from previous studies with similar study objectives and outcomes (Pauly et al. 2008).

The first group of managers (n = 24) working in a Swedish government agency was approached in 2014. Twenty of 24 (83 %) agreed to participate in the study. A second group of managers were contacted in the spring of 2015 using an experienced market research company (TNS Sifo). The recruitment of participants began by initially inviting a sample of (n = 3753) managers randomly selected from a large representative survey panel of Swedish managers. From this sample, 1,721 managers agreed to participate in the study. Of these, n = 738 managers fulfilled the inclusion criteria. In total, the study sample consisted of n = 758 managers.

5.2 Measures

Study 1

Data was collected using web-based, validated questionnaires and SMS text messages. To conduct the cost-effectiveness analysis, health-related quality of life (HRQL) was measured with the Swedish version of the EQ-5D-3L (EuroQol group 1990) at baseline and at each follow-up (6 weeks, 6 months and 12 months). The EQ-5D-3L measures five dimensions of the individual's state of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression (EuroQol group 1990, Kind et al. 2005). The participants' adherence to the recommendations was self-reported by SMS text messages sent out once a week for six weeks, and then once a month until the 12th month. The text message questions were: ‘‘How many times have you exercised in the past week? Answer with a number between 0 and 7’’ or ‘‘How many times have you exercised in the past four weeks? Answer with a number between 0 and 31”. Data for age, sex, marital status, education level, body mass index and pain were also collected at baseline.

The estimated direct costs of intervention were based on the number and type of activities in each intervention group, the amount of resources used and the duration of use of resources. The intervention costs were incurred regardless of the number of classes participants attended after allocation to an intervention group. All costs of resources were collected retrospectively (Appendix 1). The total cost of resources used in the trial reflects 2011/2012 prices, since the follow-up period occurred within this interval. The indirect cost of productivity losses was measured by sickness absence. It was measured by one self-report SMS-distributed question (Vos, Verhagen and Koes 2009). The question was ‘How many days in the past four weeks have you been absent from work because of illness? Answer with a number between 0 and 31’.

Study 2

Data was collected by web-based questionnaires administered once to each subject. The questionnaire consisted of two parts. The first contained questions about age, sex, educational level, annual income, number of children at home and job demands at work. The general health questionnaire was used to assess health status (Goldberg and Williams 1988). The stages of change with regard to exercising, readiness to exercise and attitudes towards physical exercise were also assessed (Marcus et al. 1992). Neck and back pain were assessed using parts of a questionnaire for classifying pain status (Von Korff, Dworkin and Le Resche 1990). Bothersome pain was assessed by the question ‘‘How many days during the past week has your LBP been bothersome, i.e. affected your daily activities or routines?’’ (Deyo et al. 1998). The responses were from 0 to 7.

The second part of the questionnaire was the choice task. The attributes and levels used to design the choice experiment were formulated after performing a literature search and a series of group discussions with physiotherapists, chiropractors and research experts on exercise. In the choice task, each individual was shown 10 pairs of exercise options (Appendix 2) and asked to choose the option they preferred for each pair. The choice tasks were created by using the conjoint survey design tool with a random and statistically efficient fractional orthogonal design, based on the usual design principles and practice in DCEs (Reed Johnson et al. 2013, Hainmueller, Hopkins and Yamamoto 2013).

Study 3

The health-related production loss (HRPL) starts with the question: ‘‘Over the past seven days, have you experienced any health-related problems while at work? Health problems refer to any physical or emotional problems or symptoms.’’ Response options are either ‘yes’ or ‘no’. It is followed by the question (Karlsson et al. 2013): ‘‘During the past seven days, how much did your health problems affect your performance while you were working? Think about days when you were limited in the amount or kind of work you could do, days when you accomplished less than you would like, or days when you could not do your work as carefully as usual. If health problems affected your performance only a little, choose a low number. Choose a high number if health problems affected your performance a great deal.’’ Response options ranged from 0 to 10, where 0 = ‘‘Health problems had no effect on my performance’’ and 10 = ‘‘Health problems completely prevented me from working.’’

To measure work environment problems and work environment-related production loss (WRPL), questions centred on whether an individual had experienced work environment-related problems (‘yes’ or ‘no’) and how these problems had affected their performance on a 0 – 10 scale, where 0 = ‘‘Work environment problems had no effect on my work’’ and 10 = ‘‘Work environment problems completely prevented me from working.’’ Work environment-related productivity loss is defined as any physical, psychological or social problems arising in the work environment that might impair work performance.

In the retest assessment, questions were rearranged in a different listed order to minimize recall of previous responses.

The HPQ work performance was chosen as a ‘‘golden standard’’ for the validity analysis. It is one of the most widely used productivity loss measures (Kessler et al. 2003). Individuals rate their overall performance on the days they worked in the previous seven days in their specific occupation and indicate their performance in relation to peers in a similar occupation. Response options ranged from 0 representing ‘worst possible performance’ and 10 representing ‘best possible performance’. Appendix 3 shows the questions used in Study 3 to measure production loss and work performance.

Study 4

A web-based questionnaire was used to assess job characteristics, productivity loss and the cost of productivity loss from the perspective of the manager.

To estimate the wage multipliers, managers were asked to assess the extent to which work performance was affected by job characteristics such as time sensitivity of output, teamwork, and ease of substitution in the event of sickness absence, sickness presenteeism and work environment-related problems. The response options were 1 to 5 on a Likert scale, where ‘1’ indicated that the employees work independently, work can easily be postponed, and it is easy to find a substitute worker of equal quality for an individual who is absent, present but sick or encountering work environment problems. A ‘5’ indicated that the employee works in a team, output is highly time sensitive or that it is difficult to substitute the worker.

Managers also assessed the extent to which a short, unexpected sickness absence (three days) affected the productivity of team workers. The impact of episodes of sickness presenteeism or work environment-related problems on team productivity was also assessed. Response options were on a 5-point Likert scale, where ‘1’ indicated that the company was not affected at all while a ‘5’ indicated a total shutdown.

Productivity loss, i.e. reduction of the affected worker’s performance in the event of acute or chronic illness or work environment-related problems compared to another worker with no such problems, was assessed by the managers. The response options for the productivity loss question scored from 0 to 10, where ‘0’ indicated that the worker’s performance was completely reduced while a ‘10’ indicated that the worker was not affected at all. The employer’s costs per day in excess of the worker’s wage for sickness absence, sickness presenteeism due to acute or chronic illness and work environment-related problems were also reported, either in monetary terms or as a percentage of the hourly wage.

All the questions except for the productivity loss and work environment-related questions were used in the Pauly et al. (2008) study. The questionnaire was translated into Swedish by two bilingual experts using the translation guidelines of Beaton et al. (2000). Experts familiar with the terminology of the subject covered by the questions suggested alternative wording where necessary. The questions were pre-tested and some modifications were proposed. The final version of the questions in Swedish translated back to English can be found in Appendix 4.

5.3 Statistical analyses

All analyses were performed using SPSS version 22 or STATA software version 12 (StataCorp 2011, IBM Corp 2013).

Study1

Incremental Cost-Effectiveness Ratio (ICER): An ICER was estimated from the ratio of costs and effects in the CEA framework that compared kundalini yoga to strength exercise and evidence-based advice interventions. The HRQL values were weighted with the time trade-off method to generate quality-adjusted life year (QALY) values anchored between 0 and 1, where 1 is a year lived in full health and 0 (zero) represents death. The Danish tariff was used since there were no time trade-off estimates for the Swedish population at the time (Wittrup-Jensen et al. 2009). The area under curve method was used for calculating QALYs for the entire trial period (Richardson and Manca 2004). A generalized linear model was used to analyze between-group comparisons of differences in QALY scores adjusted for baseline values since the groups may differ at baseline (Manca, Hawkins and Sculpher 2005). The net health benefit method was applied to determine the cost-effectiveness (Drummond et al. 2005).

Estimation of direct and indirect costs: The direct costs of resources were estimated according to the amount and duration of resources used. Productivity costs due to sickness absence were calculated using the human capital approach. Minimum earnings per day (supplied by Statistics Sweden) were assumed to be SEK 1,145 (EUR 132). This method is frequently used and is the one recommended by the Swedish pricing and reimbursement agency (TLV 2003). Discounting was not necessary as costs and consequences occurred within a year of recruiting participants.

Sensitivity analysis: To determine the effect of uncertainty in certain assumed parameters on the study findings, a one-way sensitivity analysis was performed by multiplying the mean daily wage used for calculating the cost of sickness absence by the median multiplier 1.28 from Pauly et al. (2008) to estimate the total absenteeism cost.

Intention-to-treat analysis was conducted including all individuals who were randomized. An interaction effect on the outcome HRQL was found between the number of days participants exercised per week and the intervention arm. An analysis was therefore performed to determine the lowest interaction point (Aiken and West 1991). The analysis showed that the number of times an individual exercised could be dichotomized into < 2 times per week (as non-adherers) and ≥ 2 times per week (as adherers), based on the lowest detected interaction point in the analysis. A generalized linear model was used to evaluate differences between intervention groups, controlling for age, baseline HRQL values and mean number of training days.

Study 2

The data about choices were analyzed using a conditional logit model based on the random utility model (Hauber et al. 2016). The random utility model assumes that each respondent will choose the alternative that provides the highest utility for all alternatives for each choice task (de Bekker-Grob, Ryan and Gerard 2012, Watson, Becker and de Bekker-Grob 2017). Assuming k exercise attributes and r individual characteristics, the conditional logit equation can be specified as:

$$U = \alpha + ß_1 X_1 + ß_2 X_2 + \ldots + ß_k X_k + \gamma_1 Z_1 + \gamma_2 Z_2 + \ldots + \gamma_r Z_r + \varepsilon$$

Where α is the alternative specific constant (ASC) that captures the average effect on utility of all factors that are not included in the model, i.e. the value of *U* when all X and all Z equals zero. X are attributes included in the DCE, β are the parameters (or coefficients) to account for the marginal utility of that attribute, Z are individual characteristics (age, sex, education, physical activity, etc.) and their corresponding coefficients and ε is the unobservable component arising from heterogeneity in tastes and any measurement errors or functional specification. All the attributes levels and background information about study participants were dummy-coded. Thus each p-value was a measure of the statistical significance of the difference between the estimated preference weight and the reference category (Bech and Gyrd-Hansen 2005).

The value that describes the relative importance of each attribute was given in percentages and reveals how respondents valued a specific exercise attribute in relation to the other studied attributes. The relative importance of the attributes was indicated by the difference between the highest and lowest attribute level parameters, divided by the total of the differences across all attributes.

Study 3

Using the Pearson Correlation and the Bland and Altman's test of agreement, the convergent validity of the production loss measures (HRPL and WRPL) were tested against the Health and Work Performance Questionnaire (HPQ) in working adults. Moderate to strong correlations were expected between the production loss measures and the HPQ work performance using data from the first assessment.

If the correlation coefficient was $r < 0.1$, it was interpreted as no correlation. A correlation coefficient of between 0.1 and 0.3 was interpreted as a weak correlation. A correlation coefficient of 0.3 – 0.5 was interpreted as moderate and one of ≥ 0.5 as strong (Cohen et al. 2013). Bland and Altman's plot (difference plot) was used to examine the agreement between the different measures of production loss and the HPQ work performance (Bland and Altman 1999).

Sub-group analysis: Sub-samples were drawn from the data for comparative analysis of associations between production loss and work performance. The sample was divided into four sub-groups based on employees' experience of health-related problems, work environment problems, a combination of health- and work environment-related problems and no problems at all at work. Due to the small sample in the sub-groups, a non-parametric bootstrap analysis was performed using 1,000 replications to obtain confidence intervals for the correlation coefficients (Efron 1981).

The test-retest reliability was determined using the Intraclass Correlation Coefficient (ICC). This is a standard method used to test the extent to which an instrument's scores (e.g. productivity loss values) remain the same when it is measured on different occasions for the same individuals. An averaged ICC coefficient < 0.50 was considered poor reliability; an ICC between 0.51 and 0.75 was seen as moderate, while coefficients > 0.75 were seen as a good test-retest reliability (Roach 2006, Terwee et al. 2007). The difference between the pair of the production loss values for the same subject was also tested using the Bland and Altman's test of repeatability (Bland and Altman 1999).

Study 4

The analysis was performed in accordance with Pauly et al. (2008), based on the assumption of there being a competitive market with perfect substitutability for all types of workers and different marginal productivity for each team of workers. Capital is held constant across companies in the production function that combines capital (K) and labour (L) to produce output Q. Further, it is assumed that if the available labour (L) is greater or equal to the labour requirement (L'), such that L > L', then output Q > 0 (Pauly et al. 2008). Thus, if we suppose that different jobs have different labour requirements, then wages (W) in jobs with L' = 1 will be equal to the marginal product value of labour added by that single individual in the job W (1).

Ordinal probit regression: The analysis was conducted in several steps. Firstly, an ordinal probit regression was used to predict the effect of job characteristics on team productivity. The job characteristics TW (teamwork), S (ease of substitution of workers) and TS (time sensitivity of output) were used as independent variables. The manager's estimation of the impact of sickness absence, sickness presenteeism and work environment-related problems on team productivity was used as the dependent variable in the equation:

$$\pi = \alpha + \beta_1,jTW + \beta_2,jS + \beta_3,jTS + \varepsilon_, \; j = 1, 2$$

where α is the intercept, π is the latent variable and ε is the error term. For both the ordinal dependent and independent variables, there were few responses at the tails on a 1 – 5 Likert scales. Thus the responses were categorized as follows: responses '1' and '2' were grouped into '1', response '3' stayed the same, but relabeled as '2' and responses '4' and '5' were grouped into '3'.

Secondly, the coefficients, β, were used to derive the mean predicted latent value π for each job category in the survey. In the final step, the predicted value of the latent variable is linked to the assessment of the cost, over and above wages, for sickness absence, sickness presenteeism and work environment problems (Y) respectively. Here, monetary costs in excess of the paid wage were transformed into percentages using the median salary data from Statistics Sweden. The large standard deviation in reported costs for sickness absence, sickness presenteeism and work environment problems were adjusted by rescaling the latent variable π. To do this, we firstly set Y equal to X if Y < X where X is the percentage reduction in the affected worker’s performance due to sickness absence, sickness presenteeism or work environment problems compared to another worker without any such problems. Secondly, the 90th percentile value of the difference (X-Y) was added to Y, which gives a rescaled Y as the wage multiplier. Thus the multipliers are based on adjusted values of Y so that the cost to the employer of the lost inputs of the affected worker is at least equal to their wage for the period in question. The wage multiplier was thus given as, $m(W) = 1 + c$, where c represents the additional costs for the employer, over and above the worker’s wage, (W) for health or work environment-related problems. The wage multiplier $m(W)$ would be equal to one if the employer does not have any extra costs for health or work environment-related problems in the workforce. Extra costs may be incurred if the employer has to deal with one or a combination of the following situations: a) other team members are not able to perform their work as expected; b) overtime has to be paid to a colleague to compensate for the increased workload; and c) delays in sales causing losses. In such cases, the wage multiplier will exceed one.

5.4 Ethical considerations

All the data used in this thesis were collected from human subjects. This thesis contains a discussion about certain decisions which were made to protect the rights of study participants. Before the start of the study, ethics applications were submitted to the regional ethics review board in Stockholm for approval. Permission, with the following reference numbers, was granted to perform the studies: Study 1: 2010/108-31/3, Study 2: 201472004-31/4, Study 3: 2014/225-32 and Study 4: 2013/1957-31/5).

6 RESULTS

6.1 Health-related quality of life among working adults with LBP

A total of 310 individuals were screened using the OMPSQ. Of these, 138 were excluded because they scored less than 90 points on the OMPSQ. The remaining 172 individuals were assessed for eligibility. Thirteen were excluded at the enrollment stage for spinal pathology or because they declined to participate. The 159 individuals who qualified for inclusion were randomized to one of the three intervention arms, resulting in 52 in the yoga group, 52 in the exercise group and 55 in the evidence-based advice group. Of the 159 participants who were allocated to the three groups, 119 (74.8%) responded to the questionnaires on all three occasions (yoga, n = 46; strength exercise, n= 36; evidence-based advice, n = 37). Women were 71 % of the participants. The mean age at baseline was 45.7 (SD 10.3). The mean pain intensity was 55.0 (SD 18.2).

As described in the statistics section, an interaction effect on the outcome HRQL was found between the number of days participants exercised per week and the intervention arms. Consequently, the analysis of HRQL took this interaction into consideration. Table 1 presents the analyses performed separately for non-adherers (i.e. exercised < 2 days per week) and adherers (i.e. exercised ≥ 2 days per week) on self-reported HRQL scores in the intervention arms. The results of the generalized linear model show that kundalini yoga had a significant ($p = 0.031$) higher effect on HRQL than evidence-based advice when participants trained twice a week or more. There was no significant difference ($p = 0.574$) in the effect of yoga on HRQL compared with exercise.

Table 1. Average HRQL scores (SD) for non-adherers and adherers in intervention groups

Period	**Yoga (n = 52)**		**Strength exercise (n = 52)**		**Evidence-based advice (n = 55)**	
Training days/week	< 2 times	≥ 2times	< 2 times	≥ 2times	< 2 times	≥ 2times
Baseline	0.72 (0.20)	0.71 (0.20)	0.76 (0.14)	0.74 (0.17)	0.70 (0.20)	0.73 (0.22)
6 weeks	0.64 (0.31)	0.80 (0.11)	0.81 (0.08)	0.78 (0.16)	0.74 (0.11)	0.70 (0.22)
6 months	0.77 (0.17)	0.77 (0.15)	0.81 (0.08)	0.76 (0.21)	0.72 (0.21)	0.70 (0.28)
12 months	0.73 (0.21)	0.79 (0.14)	0.75 (0.16)	0.79 (0.13)	0.73 (0.15)	0.75 (0.23)

HRQL- Health Related Quality of Life; SD- Standard deviation.

6.2 Cost-effectiveness of yoga for LBP

The average estimated direct cost of resources used (i.e. cost of initial examination, health personnel, materials etc.) was EUR 255 in the yoga group, EUR 461 in the strength exercise group and EUR 106 in the evidence-based advice group. The total direct cost (cost of initial examination, health personnel, materials etc.) represented the employer cost calculated according to market prices in the occupational health service. The indirect costs (i.e. productivity costs due to sickness absence) were on average EUR 1,627 in the yoga group, EUR 2,941in the exercise group and EUR 3,900 in the evidence-based advice group (Table 2). The opportunity costs of time for exercises were not included as a cost in the analysis because exercise sessions took place after work hours.

Table 2. Cost of resource use in intervention groups (EUR)

	Cost per intervention		
	Yoga (n = 52)	**Strength Exercise (n = 52)**	**Evidence -based advice (n = 55)**
Cost items			
Direct cost			
Physician assessment	3,588	3,588	3,795
Physician advice	-	-	1,898
Yoga trainer	9,568	-	-
Physiotherapists	-	19,061	-
Material/equipment	120	1,316	127
Total direct cost[1]	13,276	23,965	5,819
Mean direct cost	255	461	106
Indirect cost			
Productivity cost	84,591	152,907	214,529
Mean productivity cost	1,627	2,941	3,900
Total societal cost[2]	97,867	176,872	220,348
Mean societal cost	1,882	3,401	4,006
Sickness absence (days) after 1-year follow-up			
Training days/week			
< 2 times	24 (19)	31 (30)	22 (43)
≥ 2times	9 (15)	18 (50)	52 (108)

1. Total direct cost was employer cost.
2. Mean societal cost was used in economic evaluation from the societal perspective.

The results indicate that yoga costs EUR 150 more per individual than evidence-based advice from the employer perspective. However, yoga proved to be less costly per individual (EUR 206) compared with strength exercise (Table 3). From the societal perspective, yoga cost EUR 1519 and EUR 2124 less per individual compared with strength exercise and evidence-based advice respectively. The incremental effect of yoga demonstrated no significant improvement in HRQL compared with strength exercise or evidence-based advice.

Table 3. ICER of yoga compared to strength exercise and evidence-based advice on HRQL

Perspectives	Comparison	Incremental cost	Incremental HRQL (Mean, 95% CI)	ICER (cost per HRQL)
Employer perspective	Yoga vs. Evidence-based advice	150	0.036 (-0.033; 0.11)	4,984
	Yoga vs. Strength exercise	-206	0.023 (-0.05; 0.073)	Cost-effective
Societal perspective	Yoga vs. Evidence-based advice	-2,124	0.036 (-0.033; 0.11)	Cost-effective
	Yoga vs. Strength exercise	-1,519	0.023 (-0.05; 0.073)	Cost-effective

Costs are presented in Euros. HRQL is Health related quality of life; and CI is confidence interval. ICERs that are indicated cost-effective yielded negative incremental cost per individual.
In the published paper Aboagye et al. 2015, the HRQL was mistakenly labelled QALY. For QALY calculation see complementary analysis below.

6.2.1 Complementary CEA analysis

Health-related quality of life weights were used in the QALY calculations using the area under curve approach. An analysis correcting for the entire period over which the intervention affects the individual’s quality of life was carried out to assess the improvement in quality-adjusted life expectancy obtained through the intervention. Generalized estimation equation was used to evaluate differences between intervention groups, controlling for baseline HRQL values. Missing data were imputed in the follow-up using the previous values on HRQL carried forward for subjects with baseline values.

For the employer, yoga intervention yields an incremental cost-effectiveness ratio of EUR 2,343 per QALY compared with evidence-based advice. Without imputation, the incremental effect of yoga demonstrated a significant improvement in QALY compared with evidence-based advice but no significant difference in QALY compared with strength exercise. Yoga intervention was less costly for society than strength exercise and evidence-based advice. The incremental effect of yoga compared with strength exercise showed no significant difference in QALY with or without imputations. The results suggest that yoga intervention is likely to be cost-effective for the employer and for society compared with evidence-based advice.

Table 3.1. ICER of yoga compared with exercise and evidence-based advice on QALYs

	Employer perspective	**Societal perspective**	**Employer perspective**	**Societal perspective**
	Without imputation		With imputation	
Yoga vs. Evidence-based advice				
Incremental cost	150	-2,124	150	-2,124
Incremental QALY	0.064 (0.002; 0.12)	0.064 (0.002; 0.12)	0.035 (-0.017; 0.088)	0.035 (-0.017; 0.088)
ICER (cost per QALY)	2343	Cost-effective	4285	-
Yoga vs. Strength exercise				
Incremental cost	-206	-1,519	-206	-1,519
Incremental QALY	0.002 (-0.04; 0.043)	0.002 (-0.04; 0.043)	0.006 (-0.039; 0.05)	0.006 (-0.039; 0.05)
ICER (cost per QALY)	-	-	-	-

Costs are presented in Euros. QALY is quality-adjusted life years; and CI is confidence interval. ICERs that are indicated cost-effective yielded negative incremental cost per individual.

Using the incremental net benefit (INB) framework assuming a willingness to pay for a QALY worth EUR 11,500, yoga yielded a positive INB of EUR 1,542 and EUR 2,860 compared with strength exercise and evidence-based advice respectively. A one-way sensitivity analysis performed by multiplying the mean daily wage for the cost of sickness absence by the median multiplier 1.28 showed that yoga yielded a positive INB of EUR 1,909 and EUR 3,495 per QALY compared with strength exercise and evidence-based advice respectively.

6.3 Preferences for exercise among working adults with LBP

A total of 173 individuals were invited to participate in the survey. Six individuals were excluded for not fulfilling the inclusion criteria and nine declined participation. Another 45 of 158 subjects did not respond to the survey, which gives a response rate of 72 % (113 participants). One participant was excluded from the analysis for not completing the choice task. As a result, 112 participants were included in the analysis. Each participant provided responses to ten completed choices, resulting in 2,240 observations (i.e. 112 participants x 10 choices x 2 options for each choice).

In Study 2, by examining exercise preferences using the multi-criteria approach, it was shown that, all else being equal, workers' preferred exercise option was cardiovascular training, a group exercise with trainer supervision, exercise of high intensity and frequency of exercise twice per week. Proximity (defined as travel time) was not important for choice of exercise option (Table 4). At least one of each of the levels of the attributes contributed significantly to the choice of exercise, with the exception of travel time. Thus proximity (defined as travel time) was not important for choice of exercise type. The most popular type of incentive was the opportunity to exercise during working hours, which was preferred marginally to the wellness allowance incentive. A discount coupon for sporting goods was the least preferred type of incentive.

Table 4. Relative preference weight and the importance of the attributes for choice (n =112)

Attribute levels	**Coefficients**	**p-value**	**95% CI**
Cardiovascular training	1.10	**<0.001**	0.75 ; 1.46
Mindfulness-based training	0.18	0.262	-0.13 ; 0.49
Strength training	Set to 0		
Individual with supervision	0.45	0.136	-0.14 ; 1.05
Individual without supervision	0.92	0.093	-0.15 ; 1.99
Group with supervision	1.29	**<0.001**	0.75 ; 1.83
Group without supervision	Set to 0		
Low intensity	0.29	0.525	-0.61 ; 1.19
High intensity	1.51	**<0.001**	0.75 ; 2.27
Medium intensity	Set to 0		
Once a week (Frequency)	0.78	**<0.001**	0.42 ; 1.14
2 times/ week (Frequency)	1.66	**<0.001**	1.27 ; 2.06
3 times/ week	Set to 0		
Proximity (10 minutes)	0.34	0.407	-0.47 ; 1.15
Proximity (20 mintues)	0.53	0.254	-0.38 ; 1.44
Proximity (30 mintues)	Set to 0		
None (Incentives)	1.95	**<0.001**	1.07 ; 2.82
Wellness subsidies	2.71	**<0.001**	1.74 ; 3.69
Exercise at work	2.78	**<0.001**	1.87 ; 3.68
Discount coupon for sports goods	Set to 0		

Dependent variable: Choice of exercise option A or B; Set to 0 (Zero) = Reference category (different levels were used as the reference category for easy interpretation of coefficients)
Number of observations = 2240; Log likelihood = -1213.75

The relative importance of the attributes, demonstrated by the within attribute differences, showed that some of the differences were greater than others. The relative importance of 'exercise intensity' was on average 25 %; 'type of exercise' 19 %; 'frequency' 18 %; 'exercise and supervision levels' 17 %; and 'incentive' 17 %. The relative importance of 'proximity to place of exercise' was on average 4 %. Thus intensity of the exercise was the most important attribute, followed by type of exercise, frequency, level of supervision and incentive, in that order. Proximity (defined as travel time) was the least important attribute.

Individual characteristics also influenced preferences for exercise (Appendix 2). Exercise preferences varied between age groups with regard to the level of trainer supervision required, exercise intensity, travel time and financial incentive attributes. Active workers were more likely to prefer a higher frequency of exercise than their non-active counterparts. Working adults with LBP who had more than one child were more likely to prefer exercising close to home than their counterparts with no children. The results revealed no significant difference in variation in preferences by other participant characteristics such as sex, educational attainment, income, job demands or bothersome pain.

6.4 Convergent validity of the HRPL and WRPL measures against the HPQ work performance

The response rate in the first assessment of the 161 invited individuals was 55% (n = 88). The second assessment was sent out to n = 88 participants. The response rate in the second assessment was 67 % (n = 59). Of these, 50 % (n = 44) were able to complete the second assessment within one day after the first assessment and were included in the test-retest analysis.

In Study 3, the Pearson correlation (r) showed a moderate (r = 0.35) association between HRPL and the HPQ work performance and a weak association (r = 0.19) between WRPL and the HPQ work performance using only data from the first assessment of production loss (Table 5). For the sub-groups with health and/or work environment problems at work, the strength of association between the HRPL and WRPL and the HPQ work performance was moderately strong, as expected, with an r = 0.46 for those only reporting health problems and r = 0.31 for those only reporting work environment problems. All Pearson correlation tests between the HRPL and WRPL and the HPQ work performance were significant ($P < 0.05$).

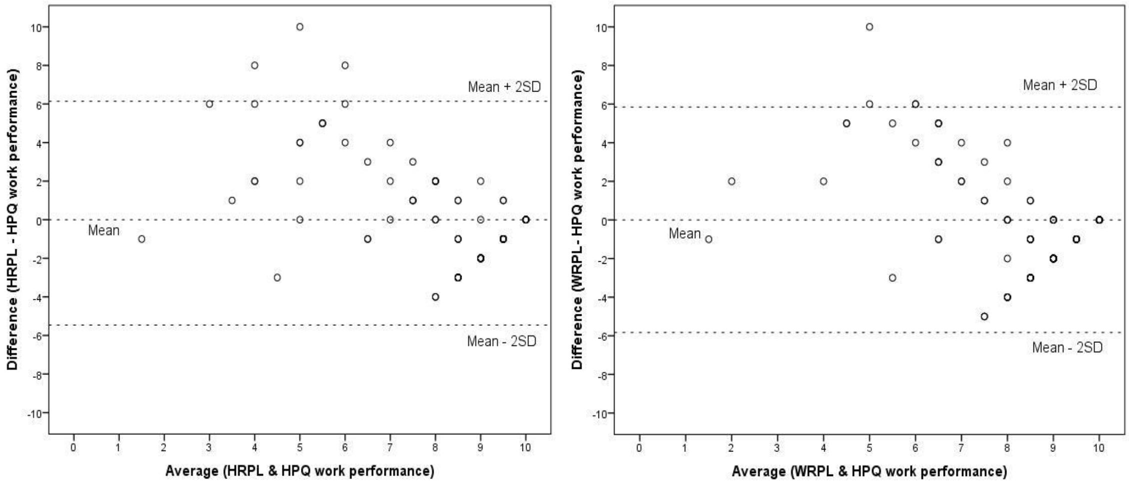
Table 5. Association between the HRPL and WRPL and the HPQ work performance

	Sample (n)	**Pearson's r**	**Bootstrapped (95% CI)**
Production loss (total sample)			
HRPL	85	0.35***	0.12 ; 0.53
WRPL	87	0.19**	0.01 ; 0.44
Production loss (Sub-groups)			
Only health-related problems	19	0.46**	0.05 ; 0.78
Only work environment problems	13	0.31**	0.01 ; 0.83
Health- and work environment-related problems	23	0.44**	0.09 ; 0.74
Not experienced any problems	31	0.11**	0.003 ; 0.41

HRPL: Health-related production loss; WRPL: Work environment-related production loss: and HPQ: Health and Work Performance Questionnaire
*** ≤ 0.01; ** ≤ 0.05

The Bland and Altman's plot (difference plot) and a test of difference demonstrated that there was an agreement between the HRPL and WRPL production loss measures and the HPQ work performance (Figure 4). The test of difference showed a small non-significance between the HRPL and HPQ work performance of 0.34 (95 % CI: - 0.28 to 0.97), and a negligible non-significant difference between the WRPL and HPQ work performance of - 0.03 (95 % CI: - 0.66 to 0.59). The results therefore suggest that there is convergence validity in the instruments. The scatter plots show that at higher averages of HRPL and the HPQ work performance (the same for WRPL and the HPQ work performance), the difference between the two measurements narrows. In other words, as the means increase, the differences between the HRPL and WRPL and work performance decrease over the range of measurement.

Figure 4. Bland-Altman plots for the HRPL and WRPL and HPQ work performance measures



HRPL: Health-related production loss; WRPL: Work environment-related production loss: and HPQ: Health and Work Performance Questionnaire

6.5 Test-retest reliability of HRPL and WRPL measures

The results revealed an ICC of 0.90; (95 % CI: 0.74 to 0.98) for the HRPL measure and 0.91for the WRPL measure (95 % CI: 0.79 to 0.98), which suggest stability in the measures. The Bland and Altman test of repeatability of the HRPL and WRPL measures showed that there was no statistically significant difference between the first and second assessment for the HRPL measure (0.23 (95% CI: -0.04 to 0.52)), and a negligible, not statistically significant difference for the WRPL measure (-0.04 (95% CI: -0.50 to 0.41)). These results support of the reliability of HRPL and WRPL measure.

6.6 The impact on teamwork productivity of health and work environment-related problems

The final sample included in Study 4 was of 758 managers from the first and second recruitments with 83 % and 46 % response rates respectively.

Table 6. Distribution of occupations in sample (n = 758)

Type of occupations	**n**	**%**
Engineers	46	6.07
Information technology (IT)	46	6.07
Carpenters, masons, construction workers, roofers	45	5.94
Doctors, registered nurses, pharmacists, dentists	44	5.80
Nurse's assistants and auxiliary nurses	44	5.80
Teachers (primary/ secondary level)	38	5.01
Insurance brokers, salesmen, purchasers and supply managers	34	4.49
Business administrators (private)	25	3.30
Media, communication, public relations, advertisers	23	3.03
Shop assistants and cashiers	22	2.90
Administrators (public service)	22	2.90
Workers in heavy industry and manufacturing	20	2.64
Analysts and investigators	18	2.37
Transport workers, couriers	18	2.37
Janitors/ real estate maintenance workers	17	2.24
Service and maintenance (machinery) workers	15	1.98
Chefs, maître d'hôtel, waiters	14	1.85
Crop producers, livestock breeders, fishermen	14	1.85
Professors and researchers (tertiary level)	13	1.72
Priests, deacons and pastors	13	1.72
Concrete casters, welders and tinsmiths	12	1.58
Social services workers	12	1.58
Military staff	10	1.32
Child care workers	10	1.32
Other**	65	8.58
Groups, n<10	118	16
Total	758	100

** Non-categorisable observations.

The results indicate that sickness absence, sickness presenteeism and work environment problems have significant impacts on teamwork productivity when different levels of job characteristics have been taken into consideration (Table 7). Especially for jobs characterized by a ‘‘high degree of teamwork’’ or ‘‘difficulty in finding substitutes’’, the likely effects of sickness absence, sickness presenteeism and work environment problems on teamwork productivity, even in the short-term, was high, all else being equal.

The values in Table 7 show the impact of health- and work environment-related problems on productivity considering job characteristics compared to the reference category (i.e. low time sensitivity, low degree of teamwork, and easy to substitute worker). The findings for the impact of health-related problems showed that if, for instance, team productivity is highly affected by sickness absence, the probability of the effect increases by 8.8 % and 13.9 % if it is either moderately difficult or very difficult to replace the worker relative to the reference category ‘’easy to substitute worker’’. If sickness absence mildly affects team productivity, the probability of the effect decreases by 9 % and 14 % if it is either moderately difficult or very difficult to replace the worker relative to this baseline job situation. This implies that a short-term sickness absence generally has little impact on teamwork productivity in this case.

Table 7. Predicting productivity loss due to sickness absence, sickness presenteeism and work environment problems by job characteristics (n=758).

	Sickness absence			Sickness presenteeism (SP)								
				SP (acute illness)			SP (chronic illness)			Work environment problems		
	Low effect	Moderate effect	High effect	Low effect	Moderate effect	High effect	Low effect	Moderate effect	High effect	Low effect	Moderate effect	High effect
Baseline probability	0.27	0.418	0.311	0.450	0.371	0.179	0.186	0.389	0.425	0.12	0.232	0.756
Time sensitivity												
2 Moderate	-0.179	0.047	0.132	-0.073	0.030	0.043	-0.083	-0.022	0.105	0.003	0.044	-0.047
3 High	-0.246	0.042	0.204	-0.090	0.035	0.054	-0.090	-0.026	0.115	0.004	0.054	-0.058
Team work												
2 Moderate	0.032	-0.001	-0.030	-0.035	0.015	0.020	0.004	0.001	-0.006	-0.010	-0.092	0.102
3 High	-0.101	-0.022	0.123	-0.152	0.048	0.104	-0.088	-0.053	0.141	-0.012	-0.123	0.134
Replace worker												
2 Moderate	-0.099	0.012	0.088	-0.088	0.034	0.054	-0.061	-0.021	0.082	-0.022	-0.192	0.214
3 Difficult	-0.145	0.007	0.139	-0.096	0.036	0.060	-0.080	-0.032	0.112	-0.027	-0.301	0.327

Baseline probability is the reference case, i.e. low time-sensitive output, low degree of teamwork, and easy to replace a worker.

Bold: significance at 5% significance level.

SP = Sickness presenteeism

6.7 Wage multipliers for sickness absence, sickness presenteeism and work environment problems

The wage multipliers (i.e. the 90th percentile of (X-Y)) in Table 8) are the cost in excess of the worker's wage by occupation. The results indicate that the costs to the employer of a short episode of sickness absence, sickness presenteeism or work environment problems exceed the worker's wage in most occupations. The median wage multiplier identified from this sample was 1.92 for sickness absence, 1.65 for sickness presenteeism (acute illness), 1.58 sickness presenteeism (chronic illness) and 1.70 for work environment problems. The result showed that absenteeism was costliest per employee, followed by work environment problems, sickness presenteeism due to acute illness and sickness presenteeism due to chronic illness.

Table 8. Wage multipliers by type of occupation, total observations (n = 758)

	n	Y rescaled Y ≥ X				Y rescaled w/90th percentile of (X-Y)			
		Sickness absence	SP (acute illness)	SP (chronic illness)	Work environment problems	Sickness absence	SP (acute illness)	SP (chronic illness)	Work environment problems
Engineers	46	1.63	1.44	1.08	1.44	2.44	2.13	1.58	2.11
Information technology (IT)	46	1.46	1.28	1.17	1.24	2.17	1.87	1.68	1.86
Carpenters, masons, construction workers, roofers	45	1.20	0.92	0.93	1.10	1.79	1.50	1.35	1.70
Doctors, nurses, pharmacists, dentists	44	1.33	1.15	1.18	1.20	1.98	1.66	1.70	1.79
Nurse's assistants and auxiliary nurses	44	0.99	1.03	1.07	0.92	1.61	1.54	1.59	1.45
Teachers (primary/ secondary level)	38	1.26	1.09	1.10	1.12	1.87	1.64	1.61	1.74
Insurance brokers, salesmen, purchasers & supply managers	34	1.15	0.82	0.89	1.04	1.90	1.65	1.34	1.64
Business administrators (private)	25	1.36	1.09	1.00	1.19	2.07	1.67	1.47	1.88
Media, communication, public relations, advertisers	23	1.40	1.31	1.22	0.94	2.13	1.83	1.73	1.46
Shop assistants and cashiers	22	1.26	0.99	0.92	0.95	1.89	1.58	1.33	1.51
Administrators (public)	22	1.04	0.96	1.05	1.42	1.66	1.52	1.51	2.14
Workers in heavy industry and manufacturing	20	0.96	1.07	1.06	1.19	1.54	1.55	1.57	1.78
Analysts and investigators	18	1.32	1.23	1.05	1.21	1.98	1.77	1.52	1.77
Transport workers, couriers	18	1.80	1.39	1.12	1.10	2.81	2.06	1.60	1.74
Janitors/ real estate maintenance workers	17	1.01	0.89	0.96	1.16	1.60	1.43	1.38	1.77
Service and maintenance (machinery) workers	15	1.69	1.22	1.05	1.44	2.54	1.92	1.53	2.17
Chefs, maître d'hôtel, waiters	14	0.88	1.02	1.18	1.07	1.40	1.43	1.66	1.66
Crop producers, livestock breeders, fishermen	14	1.21	0.93	0.83	1.06	1.85	1.54	1.23	1.67
Professors & researchers (tertiary level)	13	1.09	0.93	1.17	1.09	1.67	1.47	1.72	1.65
Priests, deacons & pastors	13	1.29	1.31	1.14	1.00	1.93	1.84	1.59	1.43
Concrete casters, welders, & tinsmiths	12	0.89	0.85	0.84	0.99	1.50	1.39	1.24	1.54
Social services workers	12	1.30	0.98	1.06	1.14	1.98	1.64	1.59	1.64
Military staff	10	1.90	1.67	1.35	1.06	2.90	2.27	1.84	1.64
Child care workers	10	1.37	1.25	1.12	0.87	2.11	1.76	1.57	1.32
Other**	65	1.29	1.17	1.10	1.26	1.92	1.75	1.62	1.88
Groups with n < 10	118								
Mean	758	1.28	1.12	1.07	1.13	1.97	1.70	1.54	1.72
Median	758	1.29	1.09	1.07	1.10	1.92	1.65	1.58	1.70

**Non-categorised observations; Y rescaled w/90th percentile of (X-Y) is wage multipliers. Y is costs per day to the company in excess of the worker's wage due to sickness absence, sickness presenteeism (acute or chronic) and work environment-related problems. X is percentage reduction in the affected worker's performance.

7 DISCUSSION

The considerable impact that LBP can have on working life has given rise to the need for effective interventions aimed at preventing the recurrence of disabling back pain and productivity costs. In order to illustrate the benefits of different types of intervention by comparing their value, it is important to assess LBP interventions from a number of perspectives. In the first part of this thesis, value assessment frameworks were used to investigate the cost-effectiveness and preference values for exercise as a secondary prevention intervention for LBP among working adults. Reduced productivity cost could also demonstrate the benefits of interventions and affect employer decisions about health intervention options to invest in. However, productivity costs are often ignored because of the challenges in obtaining reliable productivity cost estimates. The second part of this thesis focused on production loss measure validity test and deriving wage multipliers to estimate employers' productivity costs associated with sickness absence, sickness presenteeism and work environment-related problems. The results from studies 1 to 4 could be of importance for healthcare providers, employers and health policy makers.

7.1 Comparing findings with previous research

7.1.1 Cost-effectiveness of yoga intervention for LBP

Using the CEA framework, our findings indicate that yoga improved the quality of life of individuals with LBP more than evidence-based advice and to a similar degree to strength exercise among those who exercised at least twice a week. For the employer, the CEA showed that yoga improves quality of life and costs less per individual than strength exercise but more per individual than evidence-based advice. Yoga was also more likely to be cost-effective than evidence-based advice for the employer and for society because of the positive INB in relation to society's willingness-to-pay for yoga. Sensitivity analysis suggests that yoga is more cost-effective compared with strength exercise or evidence-based advice. The variable that influenced the cost-effectiveness of yoga most substantially was the difference in cost of productivity loss due to sickness absence.

In line with the current evidence, this finding suggests that kundalini yoga, compared with exercise interventions, is associated with moderate improvements in the quality of life of individuals with LBP (Wieland et al. 2017, Chou et al. 2017). Our results further support the conclusion made by Andronis et al. (2017), that that group exercise in the form of yoga may be a cost-effective intervention for LBP, although further investigation is warranted.

Compared to previous studies, the incremental cost estimate per QALY gained from the yoga intervention was considerably lower than for strength exercise and evidence-based advice (Chuang et al. 2012). This might be attributed to the substantial differences in participant characteristics and study design (in terms of cost measured, resource type used and valuation techniques).

7.1.2 Preferences for exercise among working adults with LBP

Our findings show that the most preferred exercise for working adults with LBP was cardiovascular training of high intensity, performed in a group with trainer supervision once to twice a week. The most preferred financial incentive was the opportunity to exercise during working hours. The relative importance values indicate that the intensity of exercise was the most important attribute, followed by type of exercise, exercise frequency, level of trainer supervision and incentive. Travel time to the exercise location was subordinate to all other attributes when choosing a form of exercise. The relative importance values imply that a change from one level of exercise attribute to another (e.g. from individual with supervision training to group with supervision training, or from cardiovascular training to mindfulness-based training) could considerably influence exercise preferences. Individual characteristics such as age and number of children at home also influenced exercise preferences. Individual characteristics (sex, educational attainment, income, job demands, physical activity level and bothersome pain) did not significantly influence preferences for exercise among working adults with LBP.

Previous research on sedentary adults found that the most preferred physical activity involved minimal travel time and one weekly activity session which is in part similar to the findings in this study (Farooqui et al. 2014). Here, however, older working adults with LBP preferred a longer travel time. This is probably explained by differing levels of time sensitivity between the older and the younger adults, most of whom had children at home. One important difference between our study and that of Farooqui et al. (2014) is that travel time was used as an incentive and not as an attribute. This makes comparison difficult. Using the employer as paying vehicle for incentives is a novel contribution made by the present study. Previous studies to test individual preferences or aiming to improve worker participation in exercise, have commonly included all types of incentive except ones which involve payment on the part of the employer (Giles et al. 2014). Previous research that has looked into the role of incentives in motivating physical exercise, suggests that the use of rewards could encourage healthy behaviours. The findings of this thesis support this (Finkelstein et al. 2008, Giles et al. 2014). Our results show that the offer of vouchers for sports goods should not be recommended as an incentive.

7.1.3 Validity and test-retest reliability of an at-work production loss measure

The association between the HRPL and WRPL against work performance assessed with the HPQ was moderately strong, which implies that the production loss measures are valid. The test-retest reliability suggests stability in both measures, which implies that these measures can be used to obtain reliable production loss estimates.

Previous research has shown that the construct validity and responsiveness of the HRPL tested against health-related outcomes was moderately strongly correlated with the health outcomes (Lohela Karlsson et al. 2015). The results of our study strengthen the validity of the production loss measures.

Previous studies have found that there is substantial variability in methods of measuring production loss, but the associations between different production loss instruments have been shown to be moderate, which is consistent with the findings of this thesis (Braakman-Jansen et al. 2012, Zhang et al. 2010, Beaton et al. 2010).

The Bland-Altman plots suggest that although there might be agreement between the production loss measures and the HPQ work performance, the agreement gets higher as the average production loss increases. This lack of agreement over the entire range of the scale suggests that there may be important conceptual differences between the measures with regard to work performance or the manner in which productivity loss is defined and measured. Further research may be necessary, using new methods, to investigate the definition of reduction in work performance in order to properly determine the comparability of instruments.

7.1.4 The cost of productivity loss to the employer

This is one of the few studies to support the assertion that productivity loss is most likely to be underestimated when valued according to team worker wages. One of the contentious issues of the CEA framework is how to value productivity costs in economic evaluations. Our findings are further evidence that productivity loss due to sickness absence, sickness presenteeism and work environment-related problems exceeds workers' wages in teamwork productivity when job characteristics have been taken into account (Krol et al. 2012, Pauly et al. 2008, Zhang et al. 2017). Short-term sickness absence considerably impacts team productivity in terms of costs per employee compared with work environment-related problems and sickness presenteeism (either due to acute or chronic illness). This result corroborates findings in the literature which suggest that the multiplier approach could be used to adjust wages to reflect the true cost of productivity loss from sickness absence and sickness presenteeism (Zhang et al. 2017, Pauly et al. 2008). Our results also highlight the importance of including sickness presenteeism and the cost of work environment-related problems in economic evaluations from a company's perspective. It is worth noting that when the wage multipliers are compared, the magnitude of the multipliers was high in this thesis compared to that of previous studies (Pauly et al. 2008). One explanation might be differences in social insurance systems and work contracts in the study setting which could influence wage and productivity differentials.

7.2 General strengths and limitations

A number of methodological strengths and limitations may have important implications for how we interpret the results of this thesis.

Study 1

One of the main strengths of the outcome (Study 1) is that it is a randomized control trial. A limitation of the cost-effectiveness analysis is the number of participants lost to follow-up, primarily in the control group, or the number of participants with missing data points. This could introduce bias and loss of power (Dumville, Torgerson and Hewitt 2006). To compensate for drop-outs and non-respondents, imputation from baseline values was performed.

To calculate the sample size for an economic evaluation based on a randomized clinical trial, a clinically relevant difference in the primary effect measure also becomes the basis for the economic evaluation (Briggs 2000, Korthals-de Bos et al. 2004). However, in this thesis, power calculation was based on work ability and not health-related quality of life. This makes it difficult to estimate whether the sample size in the economic evaluation may be a concern since the outcome in the CEA is a ratio.

This thesis uses hypothetical value sets based on the Danish general population to calculate QALYs. Preference values can be obtained in different ways, either by questioning those experiencing the health problem or from hypothetical values from individuals who have not necessarily experienced the health problem (Burstrom et al. 2014). However, the use of experience-based value sets to calculate QALYs has been recommended (Burstrom et al. 2014, Aronsson et al. 2015). Some studies also argue that generic instruments such as the EQ-5D questionnaire are not as sensitive to changes in health-related quality of life as disease-specific instruments (Whynes et al. 2013). This may explain the small effects detected in the target group in this thesis. Further, as the preferences for health states may differ from country to country, the Danish value sets may not be relevant for the present study population.

A strength of the present study is that it had access to written records about the resources used in the interventions. The indirect costs of health-related productivity in the economic evaluation from the employer’s perspective may be large because of the approach that was used to estimate it. The uncertainty around the cost estimates including the ICER is lacking and this has implications for how we interpret the results of the economic evaluation and for decisions about how resources should be optimally allocated (Uegaki et al. 2011).

Study 2

The DCE design can provide an immediate understanding of the results in terms of policy implementation. One of the key strengths of a DCE is the ability to include attributes that currently exist as well as those that could potentially influence the choice of interest, in this case, employer incentives for exercise. Of particular concern in the DCE study was the fact that the sample size was too small to conduct a stratified sub-group analysis. This notwithstanding, the results were quite stable across the investigated stratified sub-groups. Although the sample used in the thesis was selected in line with the survey objective and the inclusion criteria, differences between the study sample and the target population that were not captured by the background variables may limit the generalizability of the results. The ‘‘forced-choice’’ question format may be a limitation, but it was chosen to find out how respondents “trade-off” the different attributes of the options presented (Street and Burgess 2007, Reed Johnson et al. 2013).

In experiments about choice, a validity test can also be conducted to determine whether respondents choose the correct alternative in a dominated task (i.e. where one choice is clearly favourable at all the attribute levels). The design technique in this study randomly allocated the levels to create choice pairs of exercise options in which one option could not per se be inferior to the other. A rational validity test that included a dominant option could not be achieved because of the nature of the category attribute levels and scenarios created. In this thesis, an opt-out option was not included because choosing not to be active would be against the best evidence.

Study 3

One issue of concern about using a self-report measure of productivity is the tendency of employees not to report accurately about the impact of their health status on on-the-job performance (Severens et al. 2000, Johns 2010). The non-existence of objective productivity data in general adds to a common problem with research into productivity loss. Specifically, there is no “gold standard” for measuring and valuing sickness presenteeism.

In the validation study, we were not able to collect a sample representing the Swedish working population or a sample big enough to perform sub-group analysis on factors such as age and sex.

Our study is one of the few test-retest reliability studies of a production loss questionnaire. Two other instruments which have been assessed are the WPAI and the VOLP (Zhang et al. 2011, Bushnell et al. 2006). Production loss is a variable outcome which depends on one’s state of health and changes in the work environment. In previous studies the time between the first and the second assessment for test-retest reliability has been one and two weeks respectively. To properly assess the stability of an instrument, it is essential to measure under similar conditions. Our study therefore used a one-day interval restrictively (excluding those who responded after 24 hours) to ensure that the responses were all within the same week. This may have introduced the risk of the respondent remembering the response from the day before. By re-arranging the questions (in the second assessment) we tried to reduce that risk.

Study 4

This study looks at the impact of an unexpected, short-term absence (three days) of a worker in a teamwork production process. Thus, no conclusions can be drawn about the impact of long-term sickness absence. Apart from the concept and face validity tests after the questionnaire was translated into Swedish, additional testing has not yet been conducted. In the study by Pauly et al. (2008), telephone interviews were used to help managers interpret the questions. This did not take place in our study, thus introducing the risk of concepts being misinterpreted.

7.3 Implications for health policy/ actions

The results of this study suggest that yoga may be a cost-effective strategy for secondary prevention among working individuals with LBP, provided that patients comply with the intervention recommendations of exercising twice a week. From a decision-making standpoint, the finding indicate that health gains from yoga are marginally beneficial than those from strength exercise and evidence-based advice from the societal perspective. If one looks at the positive net health benefits in terms of value for the investment in interventions, then yoga exercise is clearly the optimal choice.

Despite the evidence about the positive outcomes for health of taking individual preferences into consideration in interventions (Preference Collaborative Review 2008), there are also concerns that the many individual preferences could stand in the way of appropriate treatment. The relative preference information that the DCE provides is possibly easy to understand and potentially helpful for policy implementation. If we assume that exercise design and financial incentives influence adherence to exercise, and that adherence in turn influences health outcomes, it is clearly advisable for these specific exercise attributes to be taken into consideration by providers/ employers when recommending/ providing exercise for working adults with LBP. This would enhance individual satisfaction and adherence to LBP exercise interventions. Above all, to improve participation in physical activity, the incentives included should be properly examined since their opportunity cost to the individual could matter as much as other independent exercise attributes.

Our findings suggest that the effect of health-related and work environment-related problems on co-workers' productivity (i.e. the multiplier effect) is significant. Our study proposes that calculating productivity cost using wage multipliers might give signals for employers about the extent of their cost and also to promote the inclusion of productivity costs in economic evaluations of worker health interventions whenever it is appropriate. Furthermore, the impact of sickness absence, sickness presenteeism and work environment problems on work performance could pose a significant cost to employers. The economic impact of productivity changes in paid labour due to sickness absence, sickness presenteeism and work environment problems implies that reducing the costs of health and work environment-related productivity loss could yield productivity gains for employers. This, in turn, would incentivize employers to invest in preventive occupational safety and health measures.

7.4 Suggestions for future research

The present study indicates that yoga intervention has only a small and slightly greater positive effect on health-related quality of life compared with other interventions. Previous research has shown that a range of different types of exercise have similar effect. However, even when we compared exercise interventions with evidence-based advice, only a small effect was found. This may be due to limitations in the study design, such as small sample size, insensitive measures of the outcome of interest or the natural course of LBP. Further studies with larger sample sizes and sensitive outcome measures are called for. Only a few studies have been conducted into the cost-effectiveness of exercise interventions, including yoga for LBP. The inconsistent findings about the cost-effectiveness of exercise programmes may indicate that further research is needed.

Future research should use the multi-criteria approach to examine the incentives for employers to implement and promote worker health. Since incentives have been found to encourage individuals to engage in exercise, the affordability of incentive use and their ability to encourage sustained physical activity among working adults should be examined in model programmes. Further, in order to improve adherence to exercise among people with LBP, future research into preferences should use extra questions to ask participants to indicate their level of certainty about actually doing what they say they will do (stated preference). This will clarify the extent of the bias caused by, for example, forced choice of exercise options, and will provide more information about the relevance and influence of certain attributes. It would also be interesting to include in LBP randomized trials preferred interventions to evaluate intervention effects.

The psychometric evaluation of self-reported productivity loss values presents some challenges. There is a lack of ‘‘gold standard’’ measures that accurately estimate real-life productivity. There is therefore a need to evaluate productivity loss instruments towards objective measures (criterion validity studies) to predict productivity loss while at work.

In paid work, sickness absence, sickness presenteeism and work environment problems have a considerable impact on productivity. There are many challenges involved in measuring and valuing productivity cost. Some of the major problems are the wide range of different methodologies for valuing cost (e.g. the HC and FC methods); the wide range of measuring instruments; and the difficulty of identifying the amount of loss. Decision-makers do not normally include productivity losses in economic evaluations of interventions because of concerns about the validity of productivity cost estimates. More work is required to reach a consensus about how to improve approaches and methodologies in valuing productivity costs.

Productivity costs are often ignored in economic evaluations because of the challenges in obtaining reliable productivity cost estimates. This is contentious because it makes it difficult for employers to make investment decisions about interventions to improve worker’s health based on productivity reductions. The inclusion of productivity costs in economic analyses significantly demonstrates the relative cost-benefits of interventions, which in turn might lead to different decisions about resource allocation. The direct cost savings of a workplace intervention could even disappear as a consequence of indirect costs associated with increased sickness absence and presenteeism. While it may be that effective interventions could yield savings in the short-term, their effects on indirect costs are less clear. Thus, a clear understanding of productivity costs as well as the changes in benefit is needed to make rational decisions when paying for interventions to improve worker’s health.

8 CONCLUSION

Based on the findings in this thesis, the following conclusions were reached.

a) Kundalini yoga can be considered a cost-effective early intervention for the prevention of LBP, but further investigation is warranted.

b) When examining preferences for exercise to prevent LBP, the characteristics that mattered most were the individuals' age and exercise attributes such as type of exercise, frequency, level of supervision and incentives. This implies that providers and employers could improve participation in exercise for working adults with non-specific LBP by focusing on the exercise components which are most attractive.

c) The validity test of the HRPL & WRPL suggests that the measures have convergent validity and good stability. This finding may have implications for improving methods of assessing production loss as an outcome, which represents an important cost for employers.

d) In a range of occupations, the economic impact of sickness absence, sickness presenteeism and work environment related problems on team productivity exceeds the cost of wages, when job characteristics have been taken into account. This implies that reducing the cost of health and work environment related productivity loss would yield productivity gains for employers. The proposed wage multipliers can be used to calculate the cost of health and work environment-related productivity loss, which represents an important cost for employers.

9 ACKNOWLEDGEMENTS

The authors are grateful to all the study participants in the individual projects. We are also grateful to the staff of the occupational health services centres particularly for their contribution to participant recruitment. I would like to thank all those who contributed in one way or the other from the analysis and interpretation of data to the writing of the article include in the thesis. To my main supervisor Professor Irene Jensen, thank you for your confidence in me and guiding me through this interesting path of learning. Thank you for all your support that finally helped me to make it to the finish line. To all my co-supervisors (Malin Lohela Karlsson (PhD), and Associate Professors Iben Axén and Lydia Kwak), I really appreciate your understanding and support through the whole process. All supervisors have been a solid support to me at all times.

To Laura Hartman, my mentor, thanks for the guidance in all matters we discussed and all the meetings you made time for out of your busy schedules. You are everything one could look for in a good mentor.

I acknowledge the contribution of those who gave expert and technical support in all the studies, particularly Jan Hagberg (PhD), statistician at the Division of Intervention and Implementation Research for Worker Health, Institute of Environmental Medicine for all his support with statistical analysis. Professor Lennart Bodin was always supportive whenever he was contacted for advice on statistical analysis. Arthur Camara, a software engineer at Axiom Zen, was instrumental in visualizing the questions for the preference elicitation. Anton Strezhnev who is part of the group that designed the software for designing a DCE at Harvard University, gave much guidance on how to use the conjoint survey design tool.

I am thankful for all the help I received from colleagues who work at IIR and for the pleasant environment they created in which I could thrive. I will say it has been many special years working with you all. Special thanks go to Eva Nilsson for all your help and to those (Elisabeth B. Brämberg, Carl Strömberg, Gunnar Bergström, Caroline Lornudd and Christina Björklund) who read and gave comments on the final draft of the thesis. To fellow doctoral students (Bozana Johansson, Camilla Martinsson, and previously a doctoral student but now graduated Med. dr. Andreas Eklund), I can say that I really enjoyed working with you. Thank you for the time we spent to discuss, review and the assistance I got from you all. This work was supported by financial support from FORTE and AFA Insurance.

10 REFERENCES

Abiiro, G. A., G. Leppert, G. B. Mbera, P. J. Robyn & M. De Allegri (2014) Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. *BMC Health Serv Res,* **14,** 235.

Ahonen, G. 2015. The Economy of Sustainable Careers During the Work Life Course: A Case from Finland. ed. B. R. In: Vuori J., Price R. (eds) Sustainable Working Lives. Aligning Perspectives on Health, Safety and Well-Being. Dordrecht, Netherlands: Springer.

Aiken, L. S. & S. G. West. 1991. *Multiple regression: Testing and interpreting interactions.* . Newbury Park, London: Sage.

Amaya-Amaya, M., K. Gerard & M. Ryan. 2008. Discrete choice experiments in a nutshell. In *Using Discrete Choice Experiments to Value Health and Health Care,* eds. M. Ryan, K. Gerard & M. Amaya-Amaya. Dordrecht, The Netherlands.: Springer.

Andronis, L., P. Kinghorn, S. Qiao, D. G. Whitehurst, S. Durrell & H. McLeod (2017) Cost-Effectiveness of Non-Invasive and Non-Pharmacological Interventions for Low Back Pain: a Systematic Literature Review. *Appl Health Econ Health Policy,* **15,** 173-201.

Aronsson, M., M. Husberg, A. Kalkan, N. Eckard & J. Alwin (2015) Differences between hypothetical and experience-based value sets for EQ-5D used in Sweden: Implications for decision makers. *Scand J Public Health,* **43,** 848-54.

Axen, I. & C. Leboeuf-Yde (2013) Trajectories of low back pain. *Best Pract Res Clin Rheumatol,* **27,** 601-12.

Balague, F., A. F. Mannion, F. Pellise & C. Cedraschi (2012) Non-specific low back pain. *Lancet,* **379,** 482-91.

Beaton, D. E., C. Bombardier, F. Guillemin & M. B. Ferraz (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976),* **25,** 3186-91.

Beaton, D. E., K. Tang, M. A. Gignac, D. Lacaille, E. M. Badley, A. H. Anis & C. Bombardier (2010) Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. *Arthritis Care Res (Hoboken),* **62,** 28-37.

Bech, M. & D. Gyrd-Hansen (2005) Effects coding in discrete choice experiments. *Health Econ,* **14,** 1079-83.

Bech, M., T. Kjaer & J. Lauridsen (2011) Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Econ,* **20,** 273-86.

Bergstrom, G., L. Bodin, J. Hagberg, T. Lindh, G. Aronsson & M. Josephson (2009) Does sickness presenteeism have an impact on future general health? *Int Arch Occup Environ Health,* **82,** 1179-90.

Bevan, S., T. Quadrello, R. McGee, M. Mahdon, A. Vavrovsky & L. Barham. 2009. Fit for work? Musculoskeletal disorders in the European workforce. London: The Work Foundation.

Bland, J. M. & D. G. Altman (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res,* **8,** 135-60.

Braakman-Jansen, L. M., E. Taal, I. H. Kuper & M. A. van de Laar (2012) Productivity loss due to absenteeism and presenteeism by different instruments in patients with RA and subjects without RA. *Rheumatology (Oxford),* **51,** 354-61.

Bramberg, E. B., G. Bergstrom, I. Jensen, J. Hagberg & L. Kwak (2017) Effects of yoga, strength training and advice on back pain: a randomized controlled trial. *BMC Musculoskelet Disord,* **18,** 132.

Brazier, J., J. Roberts & M. Deverill (2002) The estimation of a preference-based measure of health from the SF-36. *J Health Econ,* **21,** 271-92.

Bridges, J. F., A. B. Hauber, D. Marshall, A. Lloyd, L. A. Prosser, D. A. Regier, . . . J. Mauskopf (2011) Conjoint analysis applications in health-a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health,* **14,** 403-13.

Briggs, A. (2000) Economic evaluation and clinical trials: size matters - The need for greater power in cost analyses poses an ethical dilemma. *British Medical Journal,* **321,** 1362-1363.

Brooks, A., S. E. Hagen, S. Sathyanarayanan, A. B. Schultz & D. W. Edington (2010) Presenteeism: critical issues. *J Occup Environ Med,* **52,** 1055-67.

Bruton, A., J. H. Conway & S. T. Holgate (2000) Reliability: What is it and how is it measured? *Physiotherapy,* **86,** 94-99.

Burstrom, K., S. Sun, U. G. Gerdtham, M. Henriksson, M. Johannesson, L. A. Levin & N. Zethraeus (2014) Swedish experience-based value sets for EQ-5D health states. *Qual Life Res,* **23,** 431-42.

Bushnell, D. M., M. C. Reilly, C. Galani, M. L. Martin, J. F. Ricci, D. L. Patrick & C. R. McBurney (2006) Validation of electronic data capture of the Irritable Bowel Syndrome--Quality of Life Measure, the Work Productivity and Activity Impairment Questionnaire for Irritable Bowel Syndrome and the EuroQol. *Value Health,* **9,** 98-105.

Cancelliere, C., J. D. Cassidy, C. Ammendolia & P. Cote (2011) Are workplace health promotion programs effective at improving presenteeism in workers? A systematic review and best evidence synthesis of the literature. *BMC Public Health,* **11,** 395.

Card, D., T. Lemieux & W. C. Riddell (2004) Unions and wage inequality. *Journal of Labor Research,* **25,** 519-562.

Chapman, J. R., D. C. Norvell, J. T. Hermsmeyer, R. J. Bransford, J. DeVine, M. J. McGirt & M. J. Lee (2011) Evaluating common outcomes for measuring treatment success for chronic low back pain. *Spine (Phila Pa 1976),* **36,** S54-68.

Chou, R., R. Deyo, J. Friedly, A. Skelly, R. Hashimoto, M. Weimer, . . . E. D. Brodt (2017) Nonpharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med,* **166,** 493-505.

Chuang, L. H., M. O. Soares, H. Tilbrook, H. Cox, C. E. Hewitt, J. Aplin, . . . D. J. Torgerson (2012) A pragmatic multicentered randomized controlled trial of yoga for chronic low back pain: economic evaluation. *Spine (Phila Pa 1976),* **37,** 1593-601.

Cohen, R., M. Swerdlik & S. Philips. 2013. *Psychological Testing and Assessment. An Introduction to Tests and Measurements* California: Mayfield Publishing Company.

Cramer, H., R. Lauche, H. Haller & G. Dobos (2013) A systematic review and meta-analysis of yoga for low back pain. *Clin J Pain,* **29,** 450-60.

de Bekker-Grob, E. W., B. Donkers, M. F. Jonker & E. A. Stolk (2015) Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *Patient,* **8,** 373-84.

de Bekker-Grob, E. W., M. Ryan & K. Gerard (2012) Discrete choice experiments in health economics: a review of the literature. *Health Econ,* **21,** 145-72.

Devlin, N. J. & P. F. M. Krabbe (2013) The development of new research methods for the valuation of EQ-5D-5L. *The European Journal of Health Economics,* **14,** 1-3.

Deyo, R. A., M. Battie, A. J. H. M. Beurskens, C. Bombardier, P. Croft, B. Koes, . . . G. Waddell (1998) Outcome measures for low back pain research - A proposal for standardized use. *Spine,* **23,** 2003-2013.

Doshi, J. A. & R. J. Willke (2017) Advancing High-Quality Value Assessments of Health Care Interventions. *Value Health,* **20,** 181-184.

Drummond, M. & A. McGuire. 2007. *Economic evaluation in health care merging theory with practice*. Oxford: Oxford University Press.

Drummond, M., M. Sculpher, G. Torrance, B. O'Brien & G. Stoddart. 2005. *Methods for the economic evaluation of health care programmes*. Oxford Oxford University Press.

Dumville, J. C., D. J. Torgerson & C. E. Hewitt (2006) Reporting attrition in randomised controlled trials. *BMJ,* **332,** 969-71.

Efron, B. (1981) Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika,* **68,** 589-599.

Ekman, M., S. Jonhagen, E. Hunsche & L. Jonsson (2005) Burden of illness of chronic low back pain in Sweden: a cross-sectional, retrospective study in primary care setting. *Spine (Phila Pa 1976),* **30,** 1777-85.

EU-OSHA. 2010. OSH in figures: Work-related musculoskeletal disorders in the EU — Facts and figures. Luxembourg: European Agency for Safety and Health at Work.

EuroQol group (1990) EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy,* **16,** 199-208.

Farooqui, M. A., Y. T. Tan, M. Bilger & E. A. Finkelstein (2014) Effects of financial incentives on motivating physical activity among older adults: results from a discrete choice experiment. *BMC Public Health,* **14,** 141.

Feeny, D., W. Furlong, G. W. Torrance, C. H. Goldsmith, Z. Zhu, S. DePauw, . . . M. Boyle (2002) Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care,* **40,** 113-28.

Finkelstein, E. A., D. S. Brown, D. R. Brown & D. M. Buchner (2008) A randomized study of financial incentives to increase physical activity among sedentary older adults. *Preventive Medicine,* **47,** 182-187.

Fox, J. T. (2009) Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching. *Journal of Labor Economics,* **27,** 83-126.

George, S. Z. & M. E. Robinson (2010) Preference, expectation, and satisfaction in a clinical trial of behavioral interventions for acute and sub-acute low back pain. *J Pain,* **11,** 1074-82.

Gerard, K., M. Ryan & M. Amaya-Amaya. 2008. Introduction. In *Using Discrete Choice Experiments to Value Health and Health Care,* eds. M. Ryan, K. Gerard & M. Amaya-Amaya. Dordrecht, The Netherlands: Springer.

Giles, E. L., S. Robalino, E. McColl, F. F. Sniehotta & J. Adams (2014) The effectiveness of financial incentives for health behaviour change: systematic review and meta-analysis. *PLoS One,* **9,** e90347.

Goldberg, D. & P. Williams. 1988. A user's guide to the General Health Questionnaire. Windsor, UK: NFER-Nelson.

Hainmueller, J., D. J. Hopkins & T. Yamamoto (2013) Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political Analysis,* **22,** 1-30.

Hansson, T. & I. Jensen (2004) Swedish Council on Technology Assessment in Health Care (SBU). Chapter 6. Sickness absence due to back and neck disorders. *Scand J Public Health Suppl,* **63,** 109-51.

Hauber, A. B., J. M. Gonzalez, C. G. Groothuis-Oudshoorn, T. Prior, D. A. Marshall, C. Cunningham, . . . J. F. Bridges (2016) Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value Health,* **19,** 300-15.

Hayden, J. A., M. W. van Tulder, A. Malmivaara & B. W. Koes (2005) Exercise therapy for treatment of non-specific low back pain. *Cochrane Database Syst Rev***,** CD000335.

Henchoz, Y. & A. Kai-Lik So (2008) Exercise and nonspecific low back pain: a literature review. *Joint Bone Spine,* **75,** 533-9.

Hensing, G. (2009) The measurements of sickness absence – a theoretical perspective. *Norsk Epidemiologi* **19** 147-151.

Hoy, D., L. March, P. Brooks, F. Blyth, A. Woolf, C. Bain, . . . R. Buchbinder (2014) The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis,* **73,** 968-74.

Hubertsson, J., M. Englund, U. Hallgarde, U. Lidwall, S. Lofvendahl & I. F. Petersson (2014) Sick leave patterns in common musculoskeletal disorders--a study of doctor prescribed sick leave. *BMC Musculoskelet Disord,* **15,** 176.

IBM SPSS Statistics for Windows. 22.0, IBM Corp, Armonk, NY.

Johannesson, M. (1996) The willingness to pay for health changes, the human-capital approach and the external costs. *Health Policy,* **36,** 231-244.

Johns, G. (2010) Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior,* **31,** 519-542.

Jones, M. A., G. Stratton, T. Reilly & V. B. Unnithan (2005) Biological risk indicators for recurrent non-specific low back pain in adolescents. *Br J Sports Med,* **39,** 137-40.

Jonsson, B. (2009) Ten arguments for a societal perspective in the economic evaluation of medical innovations. *European Journal of Health Economics,* **10,** 357-359.

Jordan, J. L., M. A. Holden, E. E. Mason & N. E. Foster (2010) Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database Syst Rev*, CD005956.

Joud, A., I. F. Petersson & M. Englund (2012) Low back pain: epidemiology of consultations. *Arthritis Care Res (Hoboken),* **64,** 1084-8.

Karlsson, M. L., G. Bergstrom, C. Bjorklund, J. Hagberg & I. Jensen (2013) Measuring production loss due to health and work environment problems: construct validity and implications. *J Occup Environ Med,* **55,** 1475-83.

Kessler, R. C., C. Barber, A. Beck, P. Berglund, P. D. Cleary, D. McKenas, . . . P. Wang (2003) The world health organization health and work performance questionnaire (HPQ). *Journal of Occupational and Environmental Medicine,* **45,** 156-174.

Kigozi, J., S. Jowett, M. Lewis, P. Barton & J. Coast (2017) The Estimation and Inclusion of Presenteeism Costs in Applied Economic Evaluation: A Systematic Review. *Value in Health*.

Kind, P., R. Brooks, R. Rabin & Eds. 2005. *EQ-5D concepts and methods: A development history*. Dordrecht, Netherlands Springer.

Knies, S., A. Boonen, M. J. Candel, S. M. Evers & J. L. Severens (2013) Compensation mechanisms for lost productivity: a comparison between four European countries. *Value Health,* **16,** 740-4.

Koeneman, M. A., M. W. Verheijden, M. J. Chinapaw & M. Hopman-Rock (2011) Determinants of physical activity and exercise in healthy older adults: a systematic review. *Int J Behav Nutr Phys Act,* **8,** 142.

Koeniger, W., M. Leonardi & L. Nunziata (2007) Labor market institutions and wage inequality. *Industrial & Labor Relations Review,* **60,** 340-356.

Koeske, G. F. (1994) Some Recommendations for Improving Measurement Validation in Social Work Research. *Journal of Social Service Research,* **18,** 43-72.

Koopmanschap, M. A., F. F. Rutten, B. M. van Ineveld & L. van Roijen (1995) The friction cost method for measuring indirect costs of disease. *J Health Econ,* **14,** 171-89.

Korthals-de Bos, I., M. van Tulder, H. van Dieten & L. Bouter (2004) Economic evaluations and randomized trials in spinal disorders: principles and methods. *Spine (Phila Pa 1976),* **29,** 442-8.

Krol, M. & W. Brouwer (2014) How to estimate productivity costs in economic evaluations. *Pharmacoeconomics,* **32,** 335-44.

Krol, M., W. Brouwer & F. Rutten (2013) Productivity costs in economic evaluations: past, present, future. *Pharmacoeconomics,* **31,** 537-49.

Krol, M., W. B. Brouwer, J. L. Severens, J. Kaper & S. M. Evers (2012) Productivity cost calculations in health economic evaluations: correcting for compensation mechanisms and multiplier effects. *Soc Sci Med,* **75,** 1981-8.

Krol, M., J. Papenburg, S. S. Tan, W. Brouwer & L. Hakkaart (2016) A noticeable difference? Productivity costs related to paid and unpaid work in economic evaluations on expensive drugs. *Eur J Health Econ,* **17,** 391-402.

Lambeek, L. C., M. W. van Tulder, I. C. Swinkels, L. L. Koppes, J. R. Anema & W. van Mechelen (2011) The trend in total cost of back pain in The Netherlands in the period 2002 to 2007. *Spine (Phila Pa 1976),* **36,** 1050-8.

Lamers, L. M., W. J. Meerding, J. L. Severens & W. B. F. Brouwer (2005) The relationship between productivity and health-related quality of life: An empirical exploration in persons with low back pain. *Quality of Life Research,* **14,** 805-813.

Leboeuf-Yde, C., N. Lemeunier, N. Wedderkopp & P. Kjaer (2013) Evidence-based classification of low back pain in the general population: one-year data collected with SMS Track. *Chiropr Man Therap,* **21,** 30.

Lensberg, B. R., M. F. Drummond, N. Danchenko, N. Despiegel & C. Francois (2013) Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. *Clinicoecon Outcomes Res,* **5,** 565-73.

Lidwall, U. 2011. Vad kostar olika sjukdomar i sjukförsäkringen? Stockholm: Socialförsäkringsrapport 2011:4.

Linton, S. & K. Halldén (1998) Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain,* **14,** 209-15.

Linton, S. J. & K. Boersma (2003) Early identification of patients at risk of developing a persistent back problem: The predictive validity of the Orebro Musculoskeletal Pain Questionnaire. *Clinical Journal of Pain,* **19,** 80-86.

Lohela Karlsson, M., H. Busch, E. Aboagye & I. Jensen (2015) Validation of a measure of health-related production loss: construct validity and responsiveness - a cohort study. *BMC Public Health,* **15,** 1148.

Louviere, J. J. & E. Lancsar (2009) Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Econ Policy Law,* **4,** 527-46.

Lubeck, D. (2003) The costs of musculoskeletal disease: health needs assessment and health economics. *Best Practice & Research Clinical Rheumatology,* **17,** 529-539.

Manca, A., N. Hawkins & M. J. Sculpher (2005) Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ,* **14,** 487-96.

Mann, R., J. Brazier & A. Tsuchiya (2009) A comparison of patient and general population weightings of EQ-5D dimensions. *Health Econ,* **18,** 363-72.

Manski, C. (1977) The Structure of random utility models. *Theory and Decision,* **8,** 29-254.

Marcus, B. H., S. W. Banspach, R. C. Lefebvre, J. S. Rossi, R. A. Carleton & D. B. Abrams (1992) Using the Stages of Change Model to Increase the Adoption of Physical Activity Among Community Participants. *American Journal of Health Promotion,* **6,** 424-429.

Marsh, K., I. J. M, P. Thokala, R. Baltussen, M. Boysen, Z. Kalo, . . . I. T. Force (2016) Multiple Criteria Decision Analysis for Health Care Decision Making--Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health,* **19,** 125-37.

Martin, T. L., C. T. Yu, G. L. Martin & D. Fazzio (2006) On Choice, Preference, and Preference for Choice. *Behav Anal Today,* **7,** 234-241.

McFadden, D. (1986) The choice theory approach to market research. *Marketing Science,* **5,** 275–297.

Mokkink, L. B., C. B. Terwee, D. L. Patrick, J. Alonso, P. W. Stratford, D. L. Knol, . . . H. C. de Vet (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res,* **19,** 539-49.

Montazeri, A. & S. J. Mousavi. 2010. *Quality of Life and Low Back Pain*. New York: Springer

Morimoto, T. & T. Fukui (2002) Utilities measured by rating scale, time trade-off, and standard gamble: review and reference for health care professionals. *J Epidemiol,* **12,** 160-78.

Murray, C. J., T. Vos, R. Lozano, M. Naghavi, A. D. Flaxman, C. Michaud, . . . Z. A. Memish (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet,* **380,** 2197-223.

Orme, B. 2006. Getting started with conjoint analysis: strategies for product design and pricing research. Madison (WI): Research Publishers LLC.

Pauly, M. V., S. Nicholson, D. Polsky, M. L. Berger & C. Sharda (2008) Valuing reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. *Health Econ,* **17,** 469-85.

Pauly, M. V., S. Nicholson, J. Xu, D. Polsky, P. M. Danzon, J. F. Murray & M. L. Berger (2002) A general model of the impact of absenteeism on employers and employees. *Health Econ,* **11,** 221-31.

Persson, J., L. Bernfort, C. Wahlin, B. Oberg & K. Ekberg (2015) Costs of production loss and primary health care interventions for return-to-work of sick-listed workers in Sweden. *Disabil Rehabil,* **37,** 771-6.

Preference Collaborative Review, G. (2008) Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ,* **337,** a1864.

Punnett, L., A. Pruss-Utun, D. I. Nelson, M. A. Fingerhut, J. Leigh, S. Tak & S. Phillips (2005) Estimating the global burden of low back pain attributable to combined occupational exposures. *Am J Ind Med,* **48,** 459-69.

Rasmussen-Barr, E., M. Campello, I. Arvidsson, L. Nilsson-Wikmar & B. O. Ang (2012) Factors predicting clinical outcome 12 and 36 months after an exercise intervention for recurrent low-back pain. *Disabil Rehabil,* **34,** 136-44.

Reed Johnson, F., E. Lancsar, D. Marshall, V. Kilambi, A. Muhlbacher, D. A. Regier, . . . J. F. Bridges (2013) Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health,* **16,** 3-13.

Richardson, G. & A. Manca (2004) Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health Econ,* **13,** 1203-10.

Roach, K. E. (2006) Measurement of health outcomes: Reliability, validity and responsiveness. *Proceedings of the American Academy of Orthotists and Prosthetists,* P8-P12.

Ryan, M. & K. Gerard (2003) Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics and Health Policy,* **2,** 55-64.

Ryen, L. & M. Svensson (2014) The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. *Health Econ*.

Santana, M. J. & D. Feeny. 2008. The Importance of Measuring Health-related Quality of Life. Alberta, Canada: The Institute of Health Economics (IHE).

Savigny P, Kuntze S, Watson P, Underwood M, Ritchie G, Cotterell M, . . . Walsh D. 2009. Low Back Pain: early management of persistent non-specific low back pain. London: National Collaborating Centre for Primary Care and Royal College of General Practitioners.

Scheers, T., R. Philippaerts & J. Lefevre (2013) Compliance with different physical activity recommendations and its association with socio-demographic characteristics using an objective measure. *BMC Public Health,* **13,** 136.

Schultz, A. B., C. Y. Chen & D. W. Edington (2009) The Cost and Impact of Health Conditions on Presenteeism to Employers A Review of the Literature. *Pharmacoeconomics,* **27,** 365-378.

Severens, J. L., J. Mulder, R. J. F. Laheij & A. L. M. Verbeek (2000) Precision and accuracy in measuring absence from work as a basis for calculating productivity costs in The Netherlands. *Social Science & Medicine,* **51,** 243-249.

Sintonen, H. (2001) The 15D instrument of health-related quality of life: properties and applications. *Ann Med,* **33,** 328-36.

Stata Statistical Software: Release 12., StataCorp LP, College Station, TX.

Steffens, D., C. G. Maher, L. S. Pereira, M. L. Stevens, V. C. Oliveira, M. Chapple, . . . M. J. Hancock (2016) Prevention of Low Back Pain: A Systematic Review and Meta-analysis. *JAMA Intern Med,* **176,** 199-208.

Sterud, T. & T. Tynes (2013) Work-related psychosocial and mechanical risk factors for low back pain: a 3-year follow-up study of the general working population in Norway. *Occupational and Environmental Medicine,* **70,** 296-302.

Stinnett, A. A. & J. Mullahy (1998) Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making,* **18,** S68-80.

Street, D. J. & L. Burgess. 2007. The Construction of Optimal Stated Choice Experiments: Theory and Methods. Hoboken: Wiley.

Sundelin, G., K. Harms-Ringdahl, B. Öberg, S. O. Hansson, O. Hägg, U. Lundberg, . . . C. Hall (2015) The significance of occupational exposure for low back pain – a systematic review. *Physiotherapy,* **101,** e1456-e1457.

Tang, K. (2015) Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence. *Pharmacoeconomics,* **33,** 31-48.

Tarricone, R. (2006) Cost-of-illness analysis. What room in health economics? *Health Policy,* **77,** 51-63.

Tekur, P., S. Chametcha, R. N. Hongasandra & N. Raghuram (2010) Effect of yoga on quality of life of CLBP patients: A randomized control study. *Int J Yoga,* **3,** 10-7.

Terwee, C. B., S. D. M. Bot, M. R. de Boer, D. A. W. M. van der Windt, D. L. Knol, J. Dekker, . . . H. C. W. de Vet (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology,* **60,** 34-42.

TLV. 2003. General guidelines for eco¬nomic evaluations from the Pharmaceutical Benefits Board. Tandvårds-och Läkemedelsförmånsverket, (LFNAR 2003:2).

Tockhorn-Heidenreich, A., M. Ryan & R. Hernández. 2017. Discrete Choice Experiments. ed. P. H. H. In: Facey K., Single A. (eds) Patient Involvement in Health Technology Assessment, 121-133. Singapore: Adis.

Torrance, G. W., D. H. Feeny, W. J. Furlong, R. D. Barr, Y. Zhang & Q. Wang (1996) Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Med Care,* **34,** 702-22.

Tveito, T. H., M. Hysing & H. R. Eriksen (2004) Low back pain interventions at the workplace: a systematic literature review. *Occupational Medicine-Oxford,* **54,** 3-13.

Uegaki, K., M. C. de Bruijne, A. J. van der Beek, W. van Mechelen & M. W. van Tulder (2011) Economic evaluations of occupational health interventions from a company's perspective: a systematic review of methods to estimate the cost of health-related productivity loss. *J Occup Rehabil,* **21,** 90-9.

van den Hout, W. B. (2010) The value of productivity: human-capital versus friction-cost method. *Ann Rheum Dis,* **69 (Suppl 1),** 89-91.

Watson, V., F. Becker & E. de Bekker-Grob (2017) Discrete Choice Experiment Response Rates: A Meta-analysis. *Health Econ,* **26,** 810-817.

Wendel-Vos, W., M. Droomers, S. Kremers, J. Brug & F. van Lenthe (2007) Potential environmental determinants of physical activity in adults: a systematic review. *Obes Rev,* **8,** 425-40.

Whynes, D. K., R. A. McCahon, A. Ravenscroft, V. Hodgkinson, R. Evley & J. G. Hardman (2013) Responsiveness of the EQ-5D health-related quality-of-life instrument in assessing low back pain. *Value Health*, **16,** 124-32.

Wieland, L. S., N. Skoetz, K. Pilkington, R. Vempati, C. R. D'Adamo & B. M. Berman (2017) Yoga treatment for chronic non-specific low back pain. *Cochrane Database Syst Rev,* **1,** CD010671.

Wieser, S., B. Horisberger, S. Schmidhauser, C. Eisenring, U. Brugger, A. Ruckstuhl, . . . U. Muller (2011) Cost of low back pain in Switzerland in 2005. *Eur J Health Econ,* **12,** 455-67.

Wittrup-Jensen, K. U., J. Lauridsen, C. Gudex & K. M. Pedersen (2009) Generation of a Danish TTO value set for EQ-5D health states. *Scand J Public Health,* **37,** 459-66.

Von Korff, M., S. F. Dworkin & L. Le Resche (1990) Graded chronic pain status: an epidemiologic evaluation. *Pain,* **40,** 279-91.

Woolf, A. D. & B. Pfleger (2003) Burden of major musculoskeletal conditions. *Bull World Health Organ,* **81,** 646-56.

Vos, C. J., A. P. Verhagen & B. W. Koes (2009) The Ability of the Acute Low Back Pain Screening Questionnaire to Predict Sick Leave in Patients with Acute Neck Pain. *Journal of Manipulative and Physiological Therapeutics,* **32,** 178-183.

Zhang, W., N. Bansback, J. Kopec & A. H. Anis (2011) Measuring time input loss among patients with rheumatoid arthritis: validity and reliability of the Valuation of Lost Productivity questionnaire. *J Occup Environ Med,* **53,** 530-6.

Zhang, W., M. A. M. Gignac, D. Beaton, K. Tang, A. H. Anis & C. A. N. W. Pr (2010) Productivity Loss Due to Presenteeism Among Patients with Arthritis: Estimates from 4 Instruments. *Journal of Rheumatology,* **37,** 1805-1814.

Zhang, W., H. Sun, S. Woodcock & A. Anis (2015) Illness related wage and productivity losses: Valuing 'presenteeism'. *Soc Sci Med,* **147,** 62-71.

Zhang, W., H. Sun, S. Woodcock & A. H. Anis (2017) Valuing productivity loss due to absenteeism: firm-level evidence from a Canadian linked employer-employee survey. *Health Econ Rev,* **7,** 3.

APPENDIX 1: COSTS OF RESOURCE USE AND PRODUCTIVITY LOSS (STUDY 1)

Items	Unit cost\ charges (SEK)	Sources	Allocation
Yoga intervention	1600	Based on the trial	Cost per individual
Strength exercise intervention	600	Based on the trial	Participation fee
Direct healthcare cost			
Specialist	1050	Reference cost (County Council 2012)	30 min per visit
Physician assessment	1575	Reference cost (County Council 2012)	15 min visit
Physiotherapist	520	Reference cost (County Council 2012)	per hour visit
Physiotherapy re-visit	320	Reference cost (County Council 2012)	3.25 hour visit
Primary care (Patientavgifter)			
Physician assessment	200	Reference cost (Vårdguiden 2013)	15 min visit
Fee-for-service (specialist)	350	Reference cost (Vårdguiden 2013)	30 min per visit
Fee-for-service (physiotherapist)	100	Reference cost (Vårdguiden 2013)	per visit
Physiotherapy re-visit	100	Reference cost (Vårdguiden 2013)	3.25 hour visit
Occupational health service			
Physician assessment	600		30 min per visit
Physiotherapist	750		4.25 hour consultation
Evidence-based advice (Physician)	300	50% physician cost at OHS	15 min visit
Material cost/ productivity lost			
Back book	20	Based on the trial	per unit price
Exercise ball	200	Based on the trial	per unit price
Days off work	1143/ day	Statistics Sweden 2012/2013	median monthly wage

APPENDIX 2: CHOICE TASK & SUB-GROUP ANALYSIS (STUDY 2)

Frågor om val av träningsform

Ryggsbesvär är ett mycket vanligt problem för många människor. Träning har vetenskapligt visat sig vara effektivt för att förebygga återkommande ryggsbesvär, därför är det viktigt att ta reda på vilken typ av träning som människor föredrar.

Nedan ser du en uppställning över vilka träningsprogram som används i enkäten. Läs noga igenom tabellen innan du besvarar frågorna på nästa sida. Du kommer därefter att få ett antal olika träningsbeskrivningar där du skall välja vilket alternativ som passar dig. Du kan gå tillbaka och titta på dessa förklaringar av träningsprogram när du vill.

Träningsprogrammen beskrivs genom följande egenskaper

Egenskaper	**Vad som ingår i egenskaperna**	**Beskrivning av egenskaper**
Typ av träning	Styrketräning Konditionsträning Mindfulness-baserad träning	Den form av träning som ska genomföras. Exempel på typ av träning; löpning, styrka, yoga, dans och liknande.
Träningsutformning	Individuellt med instruktör Individuellt utan instruktör Grupp med ledare Grupp utan ledare	Beslut att träna ensam eller i grupp, samt en träning som är handledd eller inte.
Intensitet	Låg Medel Hög	Graden av ansträngning som krävs för att utföra träningen eller den fysiska aktiviteten.
Frekvens	En gång i veckan Två gånger i veckan Tre gånger i veckan	Hur många gånger träningen kan utföras per vecka.
Närheten till träningsplats	10 minuter 20 minuter 30 minuter	Närhet till platsen (t ex. gym, simhall, lokaler) där du kan utföra träningen regelbundet.
Incitament till träning	Ingen ersättning Rabattkupong till sportaffär Friskvårdsbidrag Träning på arbetstid (1h per vecka)	Beskriver vilken ersättning du skulle föredra från arbetsgivare som uppmuntran/stöd till din träning.

Nedan följer tio olika alternativ av träningsprogram. Läs noga igenom dessa och svara sedan på de efterföljande frågorna. För varje val kryssar du endast det alternativ du föredrar, **antingen** alternativ A **eller** alternativ B.

Kombinationerna av egenskaper i valen nedanför kan tyckas ovanliga och kan vara svåra att svara på, men välj ändå ett alternativ för varje fråga, A eller B. Det finns inget rätt eller fel svar, utan det är dina personliga åsikter som är viktiga.

Val 1

	Alternativ **A**	Alternativ **B**
Typ av träning	Styrketräning	Styrketräning
Träningsutformning	Individuellt utan instruktör	Individuellt utan instruktör
Intensitet	Medel	Medel
Frekvens	Två gånger i veckan	En gång i veckan
Närheten till träningsplats	10 minuter	10 minuter
Incitament till träning	Träning på arbetstid	Träning på arbetstid
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 2

	Alternativ **A**	Alternativ **B**
Typ av träning	Mindfulness-baserad träning	Styrketräning
Träningsutformning	Grupp utan ledare	Individuellt med instruktör
Intensitet	Låg	Hög
Frekvens	En gång i veckan	Två gånger i veckan
Närheten till träningsplats	10 minuter	30 minuter
Incitament till träning	Träning på arbetstid	Friskvårdsbidrag
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 3

	Alternativ **A**	Alternativ **B**
Typ av träning	Konditionsträning	Styrketräning
Träningsutformning	Grupp utan ledare	Individuellt utan instruktör
Intensitet	Låg	Medel
Frekvens	En gång i veckan	Två gånger i veckan
Närheten till träningsplats	10 minuter	20 minuter
Incitament till träning	Ingen ersättning	Friskvårdsbidrag
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 4

	Alternativ **A**	Alternativ **B**
Typ av träning	Mindfulness-baserad träning	Konditionsträning
Träningsutformning	Grupp med ledare	Individuellt utan instruktör
Intensitet	Medel	Hög
Frekvens	Två gånger i veckan	Två gånger i veckan
Närheten till träningsplats	30 minuter	20 minuter
Incitament till träning	Ingen ersättning	Rabattkupong till sportaffär
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 5

	Alternativ **A**	Alternativ **B**
Typ av träning	Konditionsträning	Styrketräning
Träningsutformning	Individuellt utan instruktör	Grupp med ledare
Intensitet	Medel	Låg
Frekvens	En gång i veckan	Tre gånger i veckan
Närheten till träningsplats	20 minuter	20 minuter
Incitament till träning	Träning på arbetstid	Friskvårdsbidrag
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 6

	Alternativ **A**	Alternativ **B**
Typ av träning	Mindfulness-baserad träning	Mindfulness-baserad träning
Träningsutformning	Individuellt utan instruktör	Grupp med ledare
Intensitet	Hög	Låg
Frekvens	Två gånger i veckan	En gång i veckan
Närheten till träningsplats	10 minuter	20 minuter
Incitament till träning	Rabattkupong till sportaffär	Friskvårdsbidrag
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 7

	Alternativ **A**	Alternativ **B**
Typ av träning	Styrketräning	Konditionsträning
Träningsutformning	Grupp med ledare	Grupp med ledare
Intensitet	Låg	Hög
Frekvens	En gång i veckan	En gång i veckan
Närheten till träningsplats	20 minuter	30 minuter
Incitament till träning	Friskvårdsbidrag	Rabattkupong till sportaffär
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 8

	Alternativ **A**	Alternativ **B**
Typ av träning	Mindfulness-baserad träning	Konditionsträning
Träningsutformning	Individuellt med instruktör	Individuellt utan instruktör
Intensitet	Låg	Medel
Frekvens	Tre gånger i veckan	En gång i veckan
Närheten till träningsplats	20 minuter	10 minuter
Incitament till träning	Friskvårdsbidrag	Träning på arbetstid
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 9

	Alternativ **A**	Alternativ **B**
Typ av träning	Mindfulness-baserad träning	Styrketräning
Träningsutformning	Grupp utan ledare	Individuellt med instruktör
Intensitet	Hög	Hög
Frekvens	En gång i veckan	Två gånger i veckan
Närheten till träningsplats	10 minuter	30 minuter
Incitament till träning	Ingen ersättning	Ingen ersättning
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Val 10

	Alternativ **A**	Alternativ **B**
Typ av träning	Konditionsträning	Konditionsträning
Träningsutformning	Individuellt med instruktör	Grupp med ledare
Intensitet	Låg	Hög
Frekvens	Två gånger i veckan	Tre gånger i veckan
Närheten till träningsplats	10 minuter	30 minuter
Incitament till träning	Friskvårdsbidrag	Rabattkupong till sportaffär
Vilket träningsprogram skulle du föredra?	**[ ]**	**[ ]**

(Kryssa endast i en ruta)

Avslutande frågor

28. I frågorna ovan fick du välja mellan olika träningsalternativ. Finns det några andra faktorer eller egenskaper, förutom de som du fått ta ställning till, som du anser vara viktiga för att välja att träna?

……
……
……
……
…………………………………………

29. Om du har några ytterligare kommentarer till svarsalternativen eller enkäten så kan du gärna skriva dem här.

……
…………
……
……
……
……………………………

Sub-group analysis: Exercise preferences influenced by individual characteristics, age group

Attribute levels	**Coefficient [Age ≤ 44] n = 58**	**Coefficient [Age > 44] n = 54**	**Difference [Coefficients]**	**p-value**	**95% CI**
Cardiovascular training	1.49	0.79	0.71	0.055	-0.02 ; 1.44
Mindfulness-based training	0.19	0.16	0.032	0.922	-0.61 ; 0.67
Strength training	Set to 0	Set to 0	-		
Individual with supervision	-0.08	0.88	-0.96	0.126	-2.19 ; 0.27
Individual without supervision	2.79	-0.82	3.62	**<0.001**	1.42 ; 5.81
Group with supervision	1.59	1.09	0.50	0.364	-0.59 ; 1.59
Group without supervision	Set to 0	Set to 0	-		
Low intensity	1.87	-1.13	3.00	**<0.001**	1.16 ; 4.84
High intensity	2.89	0.32	2.58	**<0.001**	1.02 ; 4.15
Medium intensity	Set to 0	Set to 0	-		
Once a week (Frequency)	0.67	.97	-0.30	0.412	-1.03 ; 0.42
2 times/ week (Frequency)	1.62	1.86	-0.24	0.559	-1.05 ; 0.57
3 times/ week	Set to 0	Set to 0	-		
Proximity (10 minutes)	-0.92	1.57	-2.49	**0.003**	-4.14 ; 0.83
Proximity (20 mintues)	-0.99	1.96	-2.95	**<0.001**	-4.81 ; 1.09
Proximity (30 mintues)	Set to 0	Set to 0	-		
None (Incentives)	3.17	0.95	2.22	**0.015**	0.43 ; 4.02
Wellness subsidies	4.05	1.67	2.38	**0.020**	0.37 ; 4.39
Exercise at work	4.08	1.75	2.33	**0.014**	0.47 ; 4.19
Discount coupon for sports goods	Set to 0	Set to 0	-		

Number of observation (Age ≤ 44) = 1160; Number of observation (Age > 44) = 1080; Set to 0 (Zero) = Reference category

Sub-group analysis: Exercise preferences influenced by individual characteristics, physical activity level

Attribute levels	Coefficient [Inactive] n = 37	Coefficient [Active] n = 74	Difference [Coefficients]	p-value	[95% Conf. Interval]
Cardiovascular training	1.03	1.13	-0.09	0.815	-0.87; 0.68
Mindfulness-based training	0.09	0.19	-0.11	0.745	-0.80; 0.57
Strength training	Set to 0	Set to 0			
Individual with supervision	1.38	0.12	1.25	0.062	-0.06; 2.57
Individual without supervision	0.67	0.91	-0.24	0.845	-2.60; 2.13
Group with supervision	1.68	1.15	0.53	0.372	-0.63; 1.68
Group without supervision	Set to 0	Set to 0			
Low intensity	0.28	0.23	0.05	0.958	-1.93; 2.03
High intensity	1.11	1.66	-0.55	0.512	-2.20; 1.09
Medium intensity	Set to 0	Set to 0			
Once a week (Frequency)	1.54	0.51	1.03	**0.015**	0.20; 1.86
2 times/ week (Frequency)	2.54	1.39	1.15	**0.012**	0.25; 2.06
3 times/ week	Set to 0	Set to 0			
Proximity (10 minutes)	0.97	0.26	0.71	0.441	-1.10; 2.53
Proximity (20 mintues)	1.45	0.32	1.13	0.269	-0.87; 3.13
Proximity (30 mintues)	Set to 0	Set to 0			
None (Incentives)	1.79	1.97	-0.18	0.853	-2.09; 1.73
Wellness subsidies	2.15	2.90	-0.75	0.484	-2.86; 1.35
Exercise at work	2.73	2.76	-0.03	0.973	-1.99; 1.92
Discount coupon for sports goods	Set to 0	Set to 0			

Number of observations (inactive) = 740; Number of observations (active) = 1480; Set to 0 (Zero) = Reference category

Sub-group analysis: Exercise preferences influenced by individual characteristics, number of children at home

Attribute levels	Coefficient [Children = 0] n = 48	Coefficient [Children ≥ 1] n = 58	Difference [Coefficients]	p-value	[95% Conf. Interval]
Cardiovascular training	0.84	1.49	-0.65	0.091	-1.40; 0.10
Mindfulness-based training	0.12	0.27	-0.15	0.658	-0.81; 0.51
Strength training	Set to 0	Set to 0			
Individual with supervision	0.73	0.03	0.70	0.279	-0.57; 1.97
Individual without supervision	-0.73	2.37	-3.10	**0.007**	-5.37; -0.84
Group with supervision	1.19	1.49	-0.31	0.599	-1.44; 0.83
Group without supervision	Set to 0	Set to 0			
Low intensity	-1.12	1.37	-2.49	**0.009**	-4.39; -0.59
High intensity	0.59	2.56	-1.97	**0.016**	-3.58; -0.36
Medium intensity	Set to 0	Set to 0			
Once a week (Frequency)	0.88	0.76	0.12	0.755	-0.63; 0.87
2 times/ week (Frequency)	1.94	1.58	0.36	0.393	-0.47; 1.19
3 times/ week	Set to 0	Set to 0			
Proximity (10 minutes)	1.36	-0.46	1.83	**0.035**	0.13; 3.53
Proximity (20 mintues)	1.84	-0.55	2.39	**0.014**	0.48; 4.31
Proximity (30 mintues)	Set to 0	Set to 0			
None (Incentives)	1.05	2.97	-1.91	**0.042**	-3.76; -0.07
Wellness subsidies	2.13	3.74	-1.61	0.126	-3.68; 0.45
Exercise at work	2.14	3.73	-1.59	0.103	-3.51; 0.32
Discount coupon for sports goods	Set to 0	Set to 0			

Number of observations = 960 for those with no child; Number of obs = 1160 for those with more than one child; Set to 0 (Zero) = Reference category

APPENDIX 3: Enkät för att undersöka hur ohälsa och arbetsmiljöproblem påverkar anställdas prestation (STUDY 3)

Sida 1.
Bakgrundsfrågor samt frågor om dina arbetstider:

1. Kön
Man
Kvinna

2. Ålder:____

3. Arbetar vanligtvis?
Dag (mellan kl 06-18)
Kväll (mellan kl 18-22)
Natt (mellan kl 22-06)
Skift
Efter turlista (tjänstgöringsschema)
Annan arbetstid

4. Hur många timmar arbetar du i genomsnitt per vecka? (räkna ej in övertid)
Ange i hela timmar (tim/vecka):_______

Sida 2.
Frågor om sjukfrånvaro och prestation

5. Hur många arbetsdagar under de senaste sju dagarna har du sammanlagt varit borta från arbetet på grund av sjukskrivning (sjukanmälan)?
*Ange i antal hela arbetsdagar:*_____
(Exempelvis om du arbetar 50 % och varit sjukskriven 2 dagar så blir det omräknat 1 hel arbetsdag).

6. På en skala från 0 till 10, där 0 är den sämsta möjliga arbetsprestation hos någon med din typ av jobb och 10 motsvarar prestationen hos en högpresterande individ, hur skulle du skatta den sedvanliga prestationen hos majoriteten av de som arbetar i liknande jobb som ditt?

0	1	2	3	4	5	6	7	8	9	10
Sämsta möjliga prestation										Bästa möjliga prestation

7. Om du använder samma 0 till 10 skala, hur skulle du skatta din prestation överlag under de dagar du arbetade de senaste sju dagarna?

0 Sämsta möjliga prestation	1	2	3	4	5	6	7	8	9	10 Bästa möjliga prestation

Sida 3.
Frågor om hur ohälsa och arbetsmiljöproblem påverkar din prestation

8. Har du under de senaste sju dagarna upplevt hälsoproblem men ändå valt att gå till arbetet?
Med hälsoproblem avses alla eventuella fysiska eller känslomässiga problem eller symptom.
Ja
Nej

9. Under de senaste sju dagarna, i vilken utsträckning påverkade dina hälsoproblem din prestation medan du arbetade?

0 Hälsoproblemen hade ingen påverkan på mitt arbete	1	2	3	4	5	6	7	8	9	10 Hälsoproblemen hindrade mig fullständigt från att arbeta

10. Har du under de senaste sju dagarna upplevt problem i din arbetsmiljö? *Med arbetsmiljöproblem avses alla eventuella fysiska, psykologiska eller sociala problem som kan uppstå i arbetsmiljön.*
Ja
Nej

11. Under de senaste sju dagarna, i vilken utsträckning påverkade arbetsmiljörelaterade problem din prestation medan du arbetade?

0 Arbetsmiljöproblemen hade ingen påverkan på mitt arbete	1	2	3	4	5	6	7	8	9	10 Arbetsmiljöproblemen hindrade mig fullständigt från att arbeta

APPENDIX 4: QUESTIONS ON COST OF PRODUCTIVITY LOSS FOR MANAGERS (STUDY 4)

Q1. Characteristics of the jobs

(1) Ease of substitution for absence: 'If a worker is absent for the entire day because of illness, how easy is it to find a co-worker or outside temp worker to pick up the most important responsibilities of the sick worker?' Please use a scale 1-5 where '1' represents 'very easy to replace worker with similar quality' and '5' is 'impossible to replace worker with similar quality'.

(2) Ease of substitution for presenteeism: 'How easy is it for a co-worker or an outside temp worker to pick up the most important responsibilities of the worker who is at work but sick?' Please use a scale 1-5 where '1' is 'very easy to replace worker with similar quality' and '5' is 'impossible to replace worker with similar quality'

(3) Ease of substitution for work environment problems: 'If instead a worker were to experience a work environment problem for the entire day, how easy is it for a co-worker or outside temp worker to pick up the most important responsibilities of the sick worker?' Please use a scale 1-5 where '1' represents 'very easy to replace worker with similar quality' and '5' is 'impossible to replace worker with similar quality'.

(4) Time sensitivity: 'How time sensitive is this worker's output?' Please use a scale 1-5 where '1' refers to work that can be 'easily postponed' and '5' refers to situations where the work 'cannot be postponed without severe consequences'.

(5) Teamwork: 'How important is this worker to the function of his/her team?' Please use a scale 1-5 where '1' means 'the team can work independently of each other and function as usual when a worker is absent, is present but sick or is experiencing work environment problems' and '5' refers to situations where 'the team cannot function when the worker is absent, is present for work but sick or is experiencing work environment problems'.

Q2. Managers' estimates of the impact of an episode of presenteeism or a work-environment problem on the affected worker's input

(6) 'Compared to a worker who is perfectly healthy, what, on average, would be the reduction in the affected worker's daily work performance if he/she were to have a temporary acute [or chronic] condition or experience a work environment problem?' Please use a scale of 0 to 10, where '0' indicates that the worker's 'performance would be completely reduced' while a '10' indicates that the worker would 'not be affected at all'.

Q3. Managers' estimates of the impact of an absence, episode of presenteeism or work environment problem on output

(7) Impact of a 3-day absence: 'Consider a situation in which a [type of worker] becomes unexpectedly ill and misses 3 days of work. What impact would this 3-day absence have on the output or work of the absent worker's team [or the other people the manager supervises if the absent worker does not work in a team]?' Please use a scale of 1 to 5, where '1' is 'no effect at all' and '5' is 'total shutdown'.

(8) Impact of an episode of presenteeism (separate questions for acute and chronic health conditions): 'What impact would the presence of this sick worker have on the output or work of the sick worker's team [or the other people the manager supervises if the worker does not work in a team]?' Please use a scale of 1 to 5, where '1' is 'no effect at all' and '5' is 'total shutdown'.

(9) Impact of a work environment problem: 'Consider a situation in which a [type of worker] unexpectedly experiences a work environment problem. What impact would the existence of this work environment problem have on the output or work of the sick worker's team [or the other people the manager supervises if the worker does not work in a team]?' Please use a scale of 1 to 5, where '1' is 'no effect at all' and '5' is 'total shutdown'.

Q4. Scaling questions

(10) Absences: 'Overall, how much do you think an absence by this worker costs the company, in terms of additional costs the company incurs or sales lost due to the absence? Do not include any payments made to the absent worker. Earlier you said that these workers are paid about SEK___ per day. Please try to estimate, as best you can, how much an absence of this type of worker costs the company in terms of their daily wage.' [The answer can be given as a percentage of the worker's daily wage or a specific amount.]

(11) Presenteeism (separate questions for acute and chronic health conditions): 'Earlier you said that these workers are paid about SEK___ per day. Overall, how much do you think it costs the company in terms of additional costs the company incurs or sales lost due to a temporary acute condition [or a chronic condition] for one day, compared to when the person is not sick? Costs include the value of the lost productivity, covering for the affected worker, any negative impact the illness has on the productivity of other workers you supervise, any sales lost due to reduced productivity, and any expenses to accommodate the worker's condition.' [The answer can be given as a percentage of the worker's daily wage or a specific amount.]

(12) Work environment problem: 'Earlier you said that these workers are paid about SEK___ per day. Overall, how much do you think it costs the company in terms of additional costs the company incurs or sales lost due to work environment problems for one day, compared to when the person is not facing such problems? Costs include the value of the lost productivity, covering for the affected worker, any negative impact the problem may have on the

productivity of other workers you supervise, any sales lost due to reduced productivity, and any expenses to accommodate the worker’s condition.’ [The answer can be given as a percentage of the worker’s daily wage or a specific amount.]