KU LEUVEN

DOCTORAL SCHOOL OF
BIOMEDICAL SCIENCES
Faculty of Medicine

To be seen, or not to be seen?

Evaluating periodic occupational health screening

Jonas Steel

Supervisors:
Prof. dr. Jeroen Luyten
Prof. dr. Lode Godderis

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Biomedical Sciences (PhD)

July 2021

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Public Health and Primary Care

KU LEUVEN

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

To be seen, or not to be seen?

Evaluating periodic occupational health screening

Jonas STEEL

Jury:	
Supervisor:	Prof. dr. Jeroen Luyten
Co-supervisor:	Prof. dr. Lode Godderis
Chair committee:	Prof. dr. Kathleen Freson
Chair defence:	Prof. dr. Pascal Borry
Jury Members:	Prof. dr. Erik Schokkaert
	Prof. dr. Frank Hendrickx
	Prof. dr. Dirk De Ridder
	Dr. Bart Curvers
	Prof. dr. Lutgart Braeckman (UGent)
	Prof. dr. Emile Tompa (McMaster University)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences

July 2021

© 2021 KU Leuven – Faculty of Medicine
Uitgegeven in eigen beheer, Jonas Steel, Kapucijnenvoer 35, B-3000 Leuven

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

Beglückt darf nun dich, o Heimat,
ich schauen und grüssen froh deine
lieblichen Auen;
nun lass ich ruhn den Wanderstab,
weil [meinen Werten] getreu ich
[promoviert] hab!
Wagner 1845, Tannhäuser

Had je in 2009 aan mijn 18-jarige zelf verteld dat ik in 2021 een doctoraat in de Biomedische wetenschappen zou behalen, dan was de kans klein dat ik je geloofd had. Ik was nog maar net gestart aan mijn studies filosofie aan het HIW, en was nog ver verwijderd van praktische beslommeringen als nuttigheid, verloning, jobkansen, of toekomstplannen. Vier à vijf jaar later was die kans waarschijnlijk nog maar weinig gegroeid. Ik gooide mezelf op de Master Beleidseconomie en Master economie, en genoot met volle teugen van vernieuwde inzichten in een erg ander domein dan wat ik tot dan toe had mogen leren kennen. Vandaag, elf jaar na de start van mijn universitaire avonturen, en met omzwervingen in nog een heel ander terrein achter de kiezen, kan ik tevreden terugblikken.

Het was geen makkelijke tocht, de dalen waren in het begin van mijn doctoraat soms erg diep, en ik dacht —trouwens zoals meer dan de helft van de doctorandi in Vlaanderen —meermaals dat ik maar beter van carrière zou veranderen. Gelukkig leer je jezelf ook beter en beter kennen, en leerde ik al snel dat opgeven niet aan mij besteed is. Met een portie doorzettingsvermogen, West-Vlaams werk-ethos, een snuifje koppigheid, en vooral dankzij de eindeloze ondersteuning en hulp van anderen, kan ik vandaag terugkijken op een bijzonder leerrijke en boeiende tocht. Ik neem dan ook graag de tijd om enkele van deze mensen te bedanken.

Jeroen, we startten in 2016 samen aan een nieuw avontuur: jij was net aangesteld

als prof op het LIGB, en ik was je eerste doctoraatsstudent. Je toonde jezelf al snel een doorgedreven academicus, die er steeds naar streefde om overal het beste uit te halen. Je kon altijd tijd voor me maken, en je deur stond altijd voor me open. Je leerde me opnieuw te schrijven (ook al dacht ik dat ik dat al onder de knie had), resultaats-en output gericht te denken, en bracht verbeteringen teweeg die mijn resultaten altijd dat stukje hoger wisten te tillen. Bedankt dus voor je geduld en de leerrijke samenwerking.

Lode, als Jeroen de manager van het team was, was jij zonder twijfel de coach. Je leerde me als geen ander mezelf kennen, hielp me om mijn sterktes in de verf te zetten en verder te ontwikkelen, en mijn zwakkere kanten bij te schaven. Je talent om te enthousiasmeren, mijn leergierigheid aan te wakkeren, mijn neus in de juiste richting te zetten, en je warme empathische communicatie maakten het een plezier om met je samen te werken. Ik ben dan ook verheugd dat we onze samenwerking nog verder kunnen doorzetten in een andere context. Bedankt voor je vertrouwen.

Het interdisciplinaire karakter van mijn doctoraat maakte dat ik het plezier had om niet enkel in het LIGB, maar ook in omgeving en gezondheid tal van toffe collega's te hebben. Evelyn, Miet, Lina, Mathijs, Erinn, Astrid VM, Elias, Jonas, Fien, Astrid VW, Deborah, Ellen, Dorothea, Kaat, Isabelle, M-C, Lotje, Sofie, Elke, Anke, Florence en alle andere collega's op het LIGB en in de return-to-work groep, bedankt voor de leuke koffiebabbels, Kruidtuin-lunches, en behulpzame gesprekken. Het was een plezier om met jullie samen te werken, en ik wens jullie alle succes toe in jullie projecten!

Bart, Chris, Marie-Noëlle, Ludo en Ann, de BBvag, IDEWE, Liantis, en Mensura, vanaf dag één stelden jullie je praktische kennis en ideeën tot mijn beschikking. Bedankt voor het vertrouwen! Bedankt ook aan alle werknemers, verpleegkundigen, arbeidsartsen, de deelnemende ziekenhuizen (AZ Nikolaas, Heilig Hart Leuven, Heilig Hart Tienen, en AZ Vesalius) en ondersteundende administratieve krachten voor jullie hulp en medewerking bij de Delphi, de focusgroep, de validiteitsstudies, en de effectiviteitsstudie. Jullie hulp was van onschatbare waarde bij het verzamelen van de data!

Prof. Schokkaert, Prof. Hendrickx, Prof. De Ridder, Dr. Curvers, Prof. Braeckman, en Prof. Tompa, hartelijk dank voor uw input doorheen dit doctoraatstraject en voor uw feedback op het finale manuscript. Het was een plezier om een sterk multidisciplinaire jury te hebben, en ik wil elk van u dan ook bijzonder bedanken voor uw bijdrage.

Tot slot wil ik graag mijn familie, vrienden en gezin bedanken. Zonder jullie steun had ik vandaag nooit zover gestaan. De ridders en schaapjes konden steeds mijn batterijen opladen om er terug in te vliegen, ook al gooide een

welbepaalde pandemie soms roet in de plannen. An, bedankt voor je geduld voor de bij wijlen lange dagen, je bemoedigende woorden en steun, en voor alle zalige ontspanning en vrolijke momenten de afgelopen jaren. Floortje en Lana, jullie zijn de mooiste geschenken die ik ooit heb mogen ontvangen, en ik ben ongelooflijk trots dat ik jullie de afgelopen jaren heb mogen zien opgroeien.

Abstract

The doctoral thesis focuses on periodic occupational health screening: a regular medical examination of employees with exposure to occupational risks at the workplace. These periodic health screenings enable the occupational physicians to be in regular contact with a significant proportion of the working population (at least 70% in Belgium) in a preventive medical setting, with the objective of early detection of health problems, advice towards risk-free behaviour, timely risk reductions or referrals to other care personnel. However, since a large number of employees needs to be seen yearly, it is also one of the most time consuming activities of the occupational physician. This is more and more difficult to reconcile with the shortage and persistent drop of occupational healthcare personnel.

To this end, the first part of this PhD summarises the development and validation of an alternative form of periodic occupational health screenings: an electronic screening tool with an algorithm for selective follow-up. The tool was developed and validated using qualitative and quantitative methods, and its content validity, internal consistency, structural validity, and criterion validity was assessed and improved in multiple phases. Next, the PhD outlines a randomized controlled trial of periodic health screening in the Belgian hospital sector, wherein the effectiveness of the electronic screening tool was evaluated vis-à-vis a face-to-face health screening in the short term. In the control arm of the experiment, employees all had a face-to-face periodic health examination with the occupational physician. In the intervention arm, employees completed an electronic health screening survey, after which the worst scoring 20% (i.e. the employees in the worst state of health according to the algorithm) were referred to the occupational physician for a consultation.

As a secondary objective, the thesis explores several key methodological topics at the crossroads of occupational health and safety, its efficient organization and economic evaluation. It describes methodological challenges of economic evaluations in OHS, with a special focus on health-related productivity.

Beknopte samenvatting

De doctoraatsthesis focust op periodieke gezondheidsbeoordelingen: een regelmatige medische beoordeling van werknemers die blootsgesteld worden aan beroepsrisico's op het werk. Deze periodieke medische onderzoeken zorgen ervoor dat de arbeidsarts regelmatig in contact staat met een significant deel van de werkende bevolking (ten minste 70% in België) in een preventieve medische setting, met het doel om gezondheidsproblemen vroeg te detecteren, advies te geven met betrekking tot risico-vrij gedrag, vermindering van risico, of verwijzingen te doen naar ander zorgverstrekkend personeel. Gezien een groot deel van de werknemers jaarlijks gezien moeten worden, slorpt het ook veel tijd op van de arbeidsarts in vergelijking met zijn/haar andere taken. Meer en meer is dit moeilijk te verzoenen met een tekort en daling van het aantal arbeidsgeneeskundig personeel, en het gebrek aan bewijs voor haar effectiviteit.

Daarom ontwikkelde het eerste deel van dit doctoraat een alternatieve vorm van periodieke gezondheidsbeoordelingen: een elektronische screeningtool met een algoritme dat de werknemers selectief naar de arbeidsarts toeleidt. De tool werd ontwikkeld en gevalideerd met behulp van kwalitatieve en kwantitatieve methodes, en haar inhoudsvaliditeit, interne consistentie, structurele validiteit, en criterium validiteit werd in meerdere fases beoordeeld en verbeterd. Daarna brengt dit doctoraat verslag uit van een gerandomiseerd gecontrolleerd experiment rond periodieke gezondheidsbeoordeling in de Belgische ziekenhuissector. Hierin werd de effectiviteit van het elektronische triage instrument geëvalueerd vis-à-vis een persoonlijke consultatie. In de controlearm van het experiment kregen alle werknemers een periodieke gezondheidsbeoordeling van de arbeidsarts. In de interventiearm vulden werknemers een online vragenlijst in, waarna de slechtst scorende 20% (i.e. de 20% met de slechtste gezondheid volgens het algoritme) werden doorverwezen naar de arbeidsarts voor een consultatie.

Als secundair objectief, onderzoekt de thesis verschillende methodologische sleutel elementen op de kruising tussen arbeidsgeneeskunde, haar efficiënte

organisatie en economische evaluatie. Het beschrijft de methodologische valkuilen en uitdagingen van economische evaluaties in arbeidsgeneeskunde, met een speciale focus op de schatting van gezondheids-gerelateerde productiviteit.

Contents

Abstract v

Beknopte samenvatting vii

Contents ix

List of Abbreviations xv

List of Symbols xvii

List of Figures xvii

List of Tables xix

1 INTRODUCTION AND RESEARCH OBJECTIVES 1

2 DEVELOPMENT AND VALIDATION OF A PERIODIC HEALTH SCREENING TOOL 11

2.1 Introduction 12
2.2 Methods 14
2.2.1 Questionnaire and algorithm development 15
2.2.2 Criterion validation study 16
2.2.3 Structural validity and internal consistency study 17
2.3 Results 18
2.3.1 Questionnaire and algorithm development 18
2.3.2 Criterion validation study 19
2.3.3 Structural validity and internal consistency study 21
2.4 Discussion 23

3 EFFECTIVENESS OF PERIODIC HEALTH SCREENING 31

3.1 Introduction 32

3.2 Methods . . . 35
3.2.1 Experimental design . . . 35
3.2.2 Statistical analysis . . . 37
3.3 Results . . . 39
3.3.1 Sample . . . 39
3.3.2 Generalised Mixed Models for final and intermediate outcomes . . . 39
3.3.3 Post-hoc tests . . . 40
3.4 Discussion . . . 50

4 METHODOLOGICAL CHALLENGES OF ECONOMIC EVALUATIONS IN OCCUPATIONAL HEALTH AND SAFETY 57
4.1 Methodological challenges in economic evaluation of occupational health and safety interventions . . . 58
4.1.1 Introduction . . . 58
4.1.2 Methods . . . 60
4.1.3 Three Challenges of Economic Evaluation in Occupational Health and Safety . . . 60
4.1.4 Discussion . . . 68
4.1.5 Conclusions . . . 70
4.2 Productivity estimation in economic evaluations of occupational health and safety interventions: a systematic review . . . 71
4.2.1 Introduction . . . 72
4.2.2 Methods . . . 76
4.2.3 Results . . . 77
4.2.4 Discussion . . . 86
4.3 Suggestions for a reference case . . . 104

5 DISCUSSION 107

A APPENDIX 117
A.1 Appendix to Chapter 2 . . . 117
A.1.1 Scoring algorithm . . . 117
A.1.2 Overview of categories and dimensions . . . 118
A.1.3 Descriptive statistics Study 1 . . . 120
A.1.4 Descriptive statistics Study 2 . . . 121
A.1.5 Questionnaire . . . 123
A.2 Appendix to Chapter 3 . . . 130
A.2.1 Generalised Mixed Effects Models . . . 130
A.2.2 Descriptive statistics by wave . . . 134
A.2.3 Descriptive statistics by group . . . 140
A.2.4 Sample quality . . . 143
A.2.5 Full mixed models, including covariates . . . 144

Bibliography **151**

CURRICULUM - Jonas Steel **179**

List of publications **181**

List of Abbreviations

AUC Area Under the Curve. 17, 23

AUDIT-C Alcohol Use Disorder Identification Test-Consumption. 22

B Binomial. xvii, 38, 43–47, 49

BA (uncontrolled) Before–After study. 79

c Control Group. 46, 47

CBA Controlled Before–After study. 79

CE compensation Expenses. 72, 73, 78, 79

CEA Cost-Effectiveness Analysis. 77, 79

CEA LT Cost-Effectiveness Analysis with Lost-Time in natural units. 72, 73, 78, 79

CFA Confirmatory Factor Analysis. 17, 22

CFI Comparative Fit Index. 17, 22

COPSOQ COpenhagen PsychoSOcial Questionnaire. 21, 22

COSMIN COnsensus-based Standards for the selection of health status Measurement INstruments. 16, 24

CUA Cost-Utility Analysis. 77, 79

DAM Decision-Analytic Modelling. 79

DLR Diagnostic Likelihood Ratio. xix, 20

EQ-5D Euroqol 5D. 21, 37, 38, 40, 41, 45, 46, 48

ESPPW External Services for Prevention and Protection at Work. 4

FCA Friction Cost Approach. 72–75, 78–80, 84, 88

FN False Negatives. xix, 20

FP False Positives. xix, 20

GHQ General Health Questionnaire. 21, 37, 38, 40, 41, 45, 46

GLMM Generalised Linear Mixed Effects models. 37

GP Generalised Poisson. xvii, 38, 43, 45–47, 49

HCA Human Capital Approach. 72–75, 78–80, 88

HCT Historical Controlled Trial. 79

HIS Health Interview Survey. 21, 22

HLQ Health and Labour Questionnaire. 80, 82

HLS Health Literacy (Survey). 37, 40, 45, 46

i Intervention Group. 46, 47

ICC Intraclass Correlation Coefficient. 41–44

iPCQ iMTA Productivity Cost Questionnaire. 21

IRR Incidence Rate Ratio. 40, 43

ISPPW Internal Services for Prevention and Protection at Work. 4

MA Multiplier Approach. 72–74

NB Negative Binomial. xvii, 38, 43, 46, 47, 49

NFR Need for Recovery. 21, 37, 38, 40, 42, 45, 46

NMQ Nordic Musculoskeletal Questionnaire. 22, 37, 38, 40, 45, 46

NOSQ Nordic Occupational Skin Questionnaire. 22

NPV Negative Predictive Value. xix, 20

OC Ordinal Continuous. 38, 43

OHS Occupational Health and Safety. 2, 57–62, 64–77, 87, 88, 113

OP Occupational Physician. 4

OR Odds Ratio. 40, 43, 44

PCS Prospective Cohort Study. 79

PHS Periodic Health Screenings. 3

PPV Positive Predictive Value. xix, 20

PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses. 15, 18

PRODISQ PROductivity and DISease Questionnaire. 80, 82

QQ Quantity and Quality questionnaire. 82

RCS Retrospective Cohort Study. 79

RCT (cluster) Randomized Controlled Trial. 79

RMSEA Root Mean Squared Error of Approximation. 17, 22

ROC Receiver Operating Curves. 17

SRMR Standardized Root Mean square Residual. 17, 22

USP US Panel approach. 72–74, 78

WHO HPQ World Health Organization Health and Work Performance Questionnaire. 80, 82

WOCCQ WOrking Conditions and Control Questionnaire. 21

List of Figures

2.1 Methodology 14
2.2 PRISMA flow diagram 18
2.3 ROC curves for the ordinal scoring mechanism against the assessment by the occupational physician. A=How necessary is this PHS?, B=Sum of 4 questions: how well do you think the employee stands up against physical demands of his work? How well do you think the employee stands up against mental demands of his work? Workability (1-10), and workability by category. For panel A, the top two answers ("4" & "5 very necessary" = 1) are compared to the answers below ("0 not at all necessary" & "1" & "2" & "3" =0). In panel B, the 20% poorest scoring individuals on the questions are cases (=1). 20
2.4 Plot of the second-order factor model (standardized estimates, also available in Table 2.3.) 25

3.1 Study Design. PHS=periodic health screening. [source COVID-19 data: Dong, Du, and Gardner, 2020] 36
3.2 Predicted effect plots of mixed models with 95% confidence intervals, blue=control group, orange=intervention group. x-axis=measurement moment (1,2,3) 48
3.3 Predicted effect plots of mixed models with 95% confidence intervals, blue=control group, orange=intervention group. x-axis=measurement moment (1,2,3); GP = Generalised Poisson; B=Binomial; NB=Negative Binomial 49

4.1 Costs and effects of occupational health and safety interventions, by cost and effect category and by perspective. [Blue = intervention benefits or effects; Orange = intervention costs, OH = occupational health, HRM = Human Resource Management]. 63
4.2 Study selection process and outcome 76

List of Tables

2.1 Youden analyses for different case definitions by the physician: 20% worst scores, 31% worst scores, 53% worst cases, and a criterion with employee mental health assessment. AUC=Area Under ROC Curve, PPV=Positive Predictive Value, NPV=Negative Predictive Value, DLR=Diagnostic Likelihood Ratio, FP=False Positives, FN=False Negatives, CI=confidence interval 20

2.2 Questionnaire sources. Variables included in the screening tool (agreed upon by the Delphi panel and focus group), questionnaire source used per variable, and their validity and reliability. /=no suitable questionnaire identified; no validity or reliability tests available 21

2.3 CFA (standardized) estimates of the second-order factor model with robust standard errors. Estimates reflect factor loadings of each manifest variable on the latent variable. sleep, burnout and stress are reversed scales. All estimates are significant at 0.01 confidence 26

3.1 Implementation of EU Directive 89/391/EEC in the EU (and UK); n=28; based on (Colosio et al., 2017) 33

3.2 Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals 41

3.3 Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals. 43

3.4 Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals. 44

3.5 Posthoc estimations with contrasts of within (one group over time) and between (same time different groups) group with 95% confidence intervals 45
3.6 Posthoc estimations with contrasts of within (one group over time) and between (same time different groups). Non-inferiority tests 46

4.1 Definition, advantages and disadvantages of productivity estimation methods. [QALY=quality-adjusted life year] 73
4.2 Characteristics of the included studies (N=90) and the interventions they studied. An extended table is available in an online annex. 79
4.3 Inclusion and measurement details of absenteeism, presenteeism, and friction periods by the studies included in this article (N=90). 81
4.4 Main results on the composition, type and source of price weights, and inclusion of an elasticity of labor to output, for studies that included a valuation of lost time (N=69). 85
4.5 Main results on sensitivity analyses, discounting, and inflation of productivity loss by the included studies (N=90). 86

A.1 The 12 categories and 107 dimensions identified in the systematic review and presented to the Delphi panel, and the percentage of panel members that deemed the dimension 'important' or 'very important' to include in the screening tool in round 3. Numbers in bold indicate more than 80% consensus in the Delphi, italic text indicates the 15 variables added in round 2 of the Delphi. B) Green groups indicate categories brought up by the focus group; Blue groups were prompted first. %IMP = percentage of panel members that deemed the dimension 'important' or 'very important' to include in the screening tool in round 3. 118
A.2 Descriptive statistics Study 1 to assess criterion validity (n=58) 120
A.3 Descriptive statistics Study 2 to assess internal consistency and structural validity (n=776). Numeric variables state mean and standard deviation (sd), categorical variables state count and proportion. 121
A.4 Questionnaire. The translation of the Dutch questions are for information purposes only, and are, due to copyright, not always the official versions of the questionnaires. The English questions were not validated in our article. Request permission for the use of the official versions of these questions by contacting the authors of the original questionnaires. 123
A.5 Summary Statistics by wave 134
A.6 Summary Statistics by wave, complete cases 137

A.7 Summary Statistics by group (Wave 1) 140
A.8 Summary Statistics by group (Wave 2) 141
A.9 Summary Statistics by group (Wave 3) 142
A.10 Comparison of sample estimates in time 1 for hospital B C, and D
to population data from hospital C, B, and D 143
A.11 Regressions mixed methods: group + time + interaction, and covariates 144
A.12 Regressions mixed methods: group + time + interaction, and covariates 145
A.13 Regressions mixed methods: group + time + interaction, and covariates 146
A.14 Regressions mixed methods: group + time + interaction, and covariates 147
A.15 Regressions mixed methods: group + time + interaction, and covariates. 148

Chapter 1

INTRODUCTION AND RESEARCH OBJECTIVES

This chapter was partially based on:
Steel, J., Godderis, L., & Luyten, J. (2017). *Research report: Economic evaluation in occupational health and safety.* Belgische Beroepsvereniging voor Arbeidsgeneesheren
Steel, J., Luyten, J., & Godderis, L. (2018). *Occupational health: The global evidence and value.* Society of Occupational Medicine (SOM); London
Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals

Work, health, and well-being: an intricate connection

Given the fact that we spend a major part of our lives at work, it should come as no surprise that work, health, and well-being are intricately linked. On the one hand, evidence points out that work is an important source of value and meaning (Roe & Ester, 1999; Ros, Schwartz, & Surkiss, 1999), and it generally has positive effects on health. Employment is generally the main source of income and material well-being, it meets psychosocial needs, and is often central to individual identity and social status (Waddell & Burton, 2006). This is also demonstrated by the average positive effects of return-to-work after

illness or injury (although health selection also plays a partial role) (Désiron, Donceel, Godderis, Van Hoof, & de Rijk, 2015), the benefits of employment, when possible, for sick or disabled people (Waddell & Burton, 2006), or the negative effects of worklessness and unemployment for about 90-95% of the population (Waddell & Burton, 2006).

On the other hand, the workplace can be a hazardous environment, and the nature and quality of work can have important health impacts. Employees can be confronted with a wide range of hazards: dangerous chemical substances (e.g. irritants, carcinogenic substances, or flour dust), biological substances (e.g. needle accidents, or contaminated bodily fluids), physical hazards (e.g. noise, cold, heat, or heavy lifting) or psychosocial hazards (e.g. lack of support from superiors or colleagues) (World Health Organization, 1995). In a 2017 project, the economic burden for all fatal and non-fatal work-related injuries and illnesses was estimated at a loss of 3.9% of all work-years globally, equivalent to a cost of approximately € 2680 billion (or US $ 2966 billion) for low-, middle-, and high-income countries (European Agency for Safety and Health at Work [EU-OSHA], 2017; Heuvel et al., 2017). This roughly corresponds to 4% of the global Gross domestic product (GDP) worldwide, which is comparable to the entire GDP of the United Kingdom (The World Bank Group, 2017). Globally, the International Labour Organisation estimates that every 15 seconds one person dies of an occupational accident or illness, and 153 persons have an occupational accident (Takala et al., 2014; Van Sprundel et al., 2016). Future challenges also include dealing with an aging working population (Burton & World Health Organization, 2010), and upcoming health hazards, such as nanomaterials or e-waste (Lucchini & London, 2014).

To make matters worse (at least if you tend towards egalitarianism), this positive and negative impact of work is not equally distributed across social classes: work (quality) and worklessness plays a role in health inequalities, social justice and fairness (Waddell & Burton, 2006). Risk of unemployment and the impact of worklessness is higher for vulnerable populations (Acheson et al., 1998; Waddell & Burton, 2006). Job characteristics are also often worse for lower income classes, for instance: blue-collar occupations are both more physically demanding and more often characterized by low job control (Ravesteijn, Kippersluis, & van Doorslaer, 2018). Average health and life expectancy also display a gradient by occupation, which can only partly (Ravesteijn et al., 2018 mention 60%) be attributed to selection effects such as reverse causation, e.g. health constraining occupational choice, and third factors, e.g. education, genetics, or health and risk preferences (Ravesteijn et al., 2018).

Occupational health and safety (OHS) –the medical field that focuses on workers' health prevention, well-being, and re-integration –helps to diminish this burden. It tries to ensure that work-related exposures do not result in health impairments

or illnesses, or reduced functioning because of health problems. OHS is thereby primarily focused on prevention. Primary prevention by using risk analyses to detect, limit, or avoid the impact of health risks (e.g. through vaccinations); secondary prevention by limiting the health impact of exposure to health risks (e.g. using health screening); and tertiary prevention by limiting the impact of illnesses and diseases, for instance by working on their return-to-work (Van Sprundel et al., 2016).

However, as in many other healthcare disciplines, OHS needs to reconcile improving health with pressurized resources. If funds are employed better by investing them where they matter most, or by reducing costs where possible, we maximize their contribution to reducing OHS burdens. To this end, this PhD discusses several topics where occupational health and safety, its efficient organisation, and economic evaluation overlap.

On the one hand, we aim to make a methodological contribution to this area, by discussing where and how OHS differs from other disciplines, and how this affects the economic evaluation of OHS interventions (Chapter 4). Several publications have highlighted that many economic evaluation studies in OHS are lacking sufficient scientific quality (Luyten, Steel, & Godderis, 2017; Tompa, Dolinschi, de Oliveira, & Irvin, 2009; Uegaki et al., 2010; van Dongen et al., 2012). To some extent, this is because the measurement and attribution of the effectiveness of OHS interventions is inherently more challenging than it is in, e.g., the domain of pharmaceutical evaluation. However, some researchers have also observed that many studies diverge from the 'reference case' (e.g., as described by Drummond (Drummond, Sculpher, Claxton, Stoddart, & Torrance, 2015)) in how economic evaluation should be executed and reported: Under-reporting of the employed methodology and context, not including all relevant costs and consequences, incorrectly valuing them, insufficiently accounting for uncertainty or justifying assumptions (Hamberg-van Reenen, Proper, & van den Berg, 2012; Niven, 2002; Tompa, Dolinschi, & de Oliveira, 2006; Uegaki et al., 2010). We study this topic in depth in Chapter 4 of this PhD, and give special attention to one of the key aspects in economic evaluations of OHS: health-related productivity, its relation to absenteeism and presenteeism, and its estimation.

Apart from these general methodological contributions, the rest of the PhD-manuscript tries to bring the domains of efficient organisation and OHS closer together in one specific setting: periodic health screenings. And tries to make suggestions as to how they can be improved.

Periodic health screenings (PHS) are an important tool of the occupational physician, especially for primary and secondary prevention. These face-to-face health screenings serve to timely detect potential work-related health impairments and diseases (e.g. occupational asthma or burnout), or changes in

general health that affect workplace functioning (e.g. musculoskeletal problems or depression). In the EU (and UK), these routine medical examinations are offered to nearly all workers on an annual basis, as an implementation of Article 14 of the Directive 89/391/EEC, which deals with surveillance of the health of workers at regular intervals, appropriate to the risks at work (EU, 1989b). However, the access and organisation of these services can differ across countries (Colosio et al., 2017; EU, 1989b). EU-OSHA's European Survey of Enterprises on New and Emerging Risks (ESENER), a survey among European workplaces, indicates that 65% of establishments arrange regular medical examinations to monitor the health of employees, while this depends on the company size (58% for micro and 89% for large enterprises) (EU, 2017).

The screenings often take a job-specific risk analysis as a starting point, determining who is eligible for periodic health screening. It is usually performed by the occupational physician (OP), sometimes (e.g. in Belgium) working in an internal or external service for prevention and protection at work (ISPPW or ESPPW). All employees that are eligible for occupational screening[1] are invited by the occupational physician to attend a regular screening. Depending on the country, several biometrics are measured (weight, length, Body Mass Index, blood pressure), along with spirometrics (long function), a vision test, and an optional blood and urine test, all usually performed by an occupational nurse or physician. During the clinical investigation, the occupational physician normally investigates the general health status and systems of the employee (heart and blood, airways, musculoskeletal, nose, throat, ear, skin, belly organs, thyroid, etc.). This includes an anamnesis with questions about new health burdens or changes in occupational risks, follow-up questions on previous complaints, medical advice, referral to a healthcare provider (eg, General Practitioner (GP) or specialist), or booking another appointment with an occupational health specialist (occupational physician, occupational nurse, etc.) (Colosio et al., 2017; EU, 1989b). The PHS can lead to both individual (employee) and collective (worksite) measures (International Labour Office, 1998; J. Sluiter, Weel, & Hulshof, 2013b).

However, there are as of yet no indications for the relative effectiveness of PHS compared to alternatives (Krogsbøll, Jørgensen, & Gøtzsche, 2013): Based on current evidence, it remains an open question whether targeted occupational screenings with risk-selection are beneficial (MacAuley, 2012). As they have been implemented early in the 20th century, the question is also whether these face-to-face periodic health screenings are still effective in a new company

[1] In Belgium, a royal decree determines the personnel types that are eligible for periodic health screenings: personnel with safety functions, jobs with heightened vigilance, work that involves physical, biological or chemical agents or tasks that are an ergonomic or mental burden ("Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017", n.d.).

context, with new risks in the population. Another common challenge of such workers' health surveillance systems is that they are demanding in terms of working hours. In Belgium for instance, about 70% of the employees is regularly seen by the occupational physician, thus taking up a significant portion of their time (Lambrechts et al., 2015). In addition, occupational physicians are proving more and more difficult to recruit, and their numbers are difficult to increase (All Party Parliamentary Group on Occupational Safety and Health, 2016; Mairiaux, Chanteux, Schleich, & Delwick, 2010; Torrance & Heron, 2017a). In turn, this has contributed to a diverse coverage and quality of workers' health surveillance across countries, thus undermining universal access (Colosio et al., 2017; EU, 1989b). Therefore, the question rises whether it is possible to organize at least part of these screenings in an alternative way, that is less labour-intensive.

Since general screenings among healthy individuals do not seem very effective (Krogsbøll et al., 2013; MacAuley, 2012), and a less time-intensive alternative could help alleviate dwindling occupational physicians numbers, a more targeted approach might be more promising. One possible alternative is to use a screening survey that selectively refers employees to the occupational physician when it detects indications of functioning and/or (work-related) health problems. Apart from the fact that using surveys for triage could save occupational physicians' time, it is a priori also an attractive option for employers and employees. The former might gain more value for money, the latter might prefer a survey's practical ease (an online survey can be completed everywhere), and evidence points out that some health problems are more easily disclosed in surveys than in face-to-face consultations (J. S. Steel, Godderis, & Luyten, 2021b). Several countries already use surveys as surveillance systems to monitor occupational health, but not for the purpose of individual triage and screening (Cherry, Meyer, Holt, Chen, & McDonald, 2000; Valenty et al., 2015a). Belgium has taken legal steps in the direction of implementing triage instruments, and allows some employees to be followed up at two-yearly intervals with a medical questionnaire in between, which serves to identify employees in need of closer follow-up ("Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017", n.d.).

Although several occupational health surveys are available, these are not suitable to reliably perform the required two tasks of periodic health screening (systematic detection and selective referral) for a range of health risks in a broad worker population. Available surveys either focus on other goals (e.g. incidence estimation), on specific health dimensions (e.g. mental health), or on specific populations (e.g. nurses, ambulance workers, or rail safety workers) (Ruitenburg, Frings-Dresen, & Sluiter, 2016). To our knowledge, psychometrically validated triage questionnaires that can be used for broad populations while detecting a

range of health problems and estimating whom a physician should see, do not exist.

Objectives of the PhD

The PhD first summarizes the development and validation of an alternative form of periodic occupational health screenings: an electronic screening tool with an algorithm for selective follow-up. The tool is developed and validated using qualitative and quantitative methods, and its content validity, internal consistency, structural validity, and criterion validity is assessed and improved in multiple phases.

Next, the PhD reports a randomized controlled trial of periodic health screening in the Belgian hospital sector, wherein the effectiveness of the electronic screening tool is evaluated vis-à-vis a face-to-face health screening in the short term. In the control arm of the experiment, employees all have a face-to-face periodic health examination with the occupational physician. In the intervention arm, employees complete an electronic health screening survey, after which the worst scoring 20% (i.e. the employees in the worst state of health according to the algorithm) will be referred to the occupational physician for a consultation. Ideally, this PhD would have included a full economic evaluation of the electronic screening tool. Given the limited timing, this was not warranted within the current project.

As a secondary objective, the thesis explores several key methodological topics at the crossroads of occupational health and safety, its efficient organization and economic evaluation. It describes the methodological pitfalls and challenges of economic evaluations in OHS, with a special focus on the estimation of health-related productivity.

References

Acheson, D., Barker, D., Chambers, J., Drever, F., Earwicker, R., Graham, H., ... Whitehead, M. (1998). *Independent inquiry into inequalities in health report*. The Stationery Office, London.

All Party Parliamentary Group on Occupational Safety and Health. (2016). Occupational medical workforce crisis. the need for action to keep the uk workforce healthy. All Party Parliamentary Group on Occupational Safety and Health London.

Burton, J., & World Health Organization. (2010). *Who healthy workplace framework and model: Background and supporting literature and practices* (Report No. Report No.: 9241500247). World Health Organization.

Cherry, N., Meyer, J., Holt, D., Chen, Y., & McDonald, J. (2000). Surveillance of work-related diseases by occupational physicians in the uk: Opra 1996–1999. *Occupational Medicine*, *50*(7), 496–503.

Colosio, C., Mandic-Rajcevic, S., Godderis, L., van der Laan, G., Hulshof, C., & van Dijk, F. (2017). Workers' health surveillance: Implementation of the directive 89/391/eec in europe. *Occupational Medicine*, *67*(7), 574–578.

Désiron, H. A., Donceel, P., Godderis, L., Van Hoof, E., & de Rijk, A. (2015). What is the value of occupational therapy in return to work for breast cancer patients? a qualitative inquiry among experts. *European journal of cancer care*, *24*(2), 267–280.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Univ Pr.

EU. (1989b). Council directive 89/391/eec of 12 june 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work. *Official Journal of the European Communities*, *183*, 1–8.

EU. (2017). Commission staff working document, ex-post evaluation of the european union occupational safety and health directives (refit evaluation). *Official Journal of the European Communities*, *SWD 10*, 21.

European Agency for Safety and Health at Work. (2017). The economics of occupational safety and health - the value of osh to society. Web Page.

Hamberg-van Reenen, H. H., Proper, K. I., & van den Berg, M. (2012). Worksite mental health interventions: A systematic review of economic evaluations. *Occup Environ Med*, *69*(11), 837–45. doi:10.1136/oemed-2012-100668

Heuvel, S., Zwaan, L., Dam, L. v., Oude Hengel, K., Eekhout, I., van Emmerik, M., ... Wilhelm, C. (2017). *Estimating the costs of work-related accidents and ill-health: An analysis of european data sources* (Report No. 9292409972). European Agency for Safety and Health at Work (EU-OSHA).

International Labour Office. (1998). *Technical and ethical guidelines for workers' health surveillance (osh no. 72)*. International Labour Office (ILO).

Krogsbøll, L. T., Jørgensen, K., & Gøtzsche, P. C. (2013). General health checks in adults for reducing morbidity and mortality from disease. *JAMA*, *309*(23), 2489–2490. doi:10.1001/jama.2013.5039

Lambrechts, M.-C., Ketterer, F., Symons, L., Mairiaux, P., Peremans, L., Remmen, R., ... Godderis, L. (2015). The approach taken to substance abuse by occupational physicians: A qualitative study on influencing factors. *Journal of occupational and environmental medicine*, *57*(11), 1228–1235.

Lucchini, R. G., & London, L. (2014). Global occupational health: Current challenges and the need for urgent action. *Ann Glob Health*, *80*(4), 251–6.

Luyten, J., Steel, J., & Godderis, L. (2017). Economic evaluation of occupational health services: Necessary, challenging and promising. *Occup Environ Med*, *74*(11), 847–848. doi:10.1136/oemed-2017-104495

MacAuley, D. (2012). The value of conducting periodic health checks. British Medical Journal Publishing Group.

Mairiaux, P., Chanteux, A., Schleich, E., & Delwick, B. (2010). *Financement des missions de prévention des services externes de prévention et de protection au travail*. SPF Emploi, travail et concertation sociale.

Niven, K. J. (2002). A review of the application of health economics to health and safety in healthcare. *Health Policy*, *61*(3), 291–304. doi:10.1016/S0168-8510(01)00224-X

Ravesteijn, B., Kippersluis, H. v., & van Doorslaer, E. (2018). The wear and tear on health: What is the role of occupation? *Health economics*, *27*(2), e69–e86.

Roe, R. A., & Ester, P. (1999). Values and work: Empirical findings and theoretical perspective. *Applied psychology*, *48*(1), 1–21.

Ros, M., Schwartz, S. H., & Surkiss, S. (1999). Basic individual values, work values, and the meaning of work. *Applied psychology*, *48*(1), 49–71.

Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017. (n.d.).

Ruitenburg, M. M., Frings-Dresen, M. H., & Sluiter, J. K. (2016). How to define the content of a job-specific worker's health surveillance for hospital physicians? *Saf Health Work*, *7*(1), 18–31. doi:10.1016/j.shaw.2015.08.004

Sluiter, J., Weel, A., & Hulshof, C. (2013b). Leidraad preventief medisch onderzoek van werkenden. *NVAB, Utrecht*.

Steel, J., Godderis, L., & Luyten, J. (2017). *Research report: Economic evaluation in occupational health and safety*. Belgische Beroepsvereniging voor Arbeidsgeneesheren.

Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals.

Steel, J., Luyten, J., & Godderis, L. (2018). *Occupational health: The global evidence and value*. Society of Occupational Medicine (SOM); London.

Steel, J. S., Godderis, L., & Luyten, J. (2021b). Disclosure in online vs. face-to-face occupational health screenings: A cross-sectional study in belgian hospital employees. *International Journal of Environmental Research and Public Health*, *18*(4), 1460. doi:10.3390/ijerph18041460

Takala, J., Hämäläinen, P., Saarela, K. L., Yun, L. Y., Manickam, K., Jin, T. W., ... Lim, S., et al. (2014). Global estimates of the burden of injury and illness at work in 2012. *Journal of occupational and environmental hygiene*, *11*(5), 326–337.

The World Bank Group. (2017). Databank - gdp (current us dollars). Web Page.

Tompa, E., Dolinschi, R., & de Oliveira, C. (2006). Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil*, *16*(3), 375–400. doi:10.1007/s10926-006-9035-2

Tompa, E., Dolinschi, R., de Oliveira, C., & Irvin, E. (2009). A systematic review of occupational health and safety interventions with economic analyses. *J Occup Environ Med*, *51*(9), 1004–23. doi:10.1097/JOM.0b013e3181b34f60

Torrance, I., & Heron, R. (2017a). Occupational health should be part of the nhs. British Medical Journal Publishing Group.

Uegaki, K., de Bruijne, M. C., Lambeek, L., Anema, J. R., van der Beek, A. J., van Mechelen, W., & van Tulder, M. W. (2010). Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. *Scand J Work Environ Health*, *36*(4), 273–88. doi:10.5271/sjweh.3017

Valenty, M., Homère, J., Lemaitre, A., Plaine, J., Ruhlman, M., Cohidon, C., & Imbernon, E. (2015a). Surveillance programme for uncompensated work-related diseases in france. *Occupational Medicine*, *65*(8), 642–650.

Van Sprundel, M. (, Van Regenmortel, A. (, Muylaert, K. (, Borgers, E., Cambré, L., De Schryver, A., ... Vanthielen, K. (2016). *Arbeid en gezondheid: Een delicaat evenwicht?* Antwerpen: Intersentia.

van Dongen, J. M., Proper, K. I., van Wier, M. F., van der Beek, A. J., Bongers, P. M., van Mechelen, W., & van Tulder, M. W. (2012). A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health*, *38*(5), 393–408. doi:10.5271/sjweh.3275

Waddell, G., & Burton, A. K. (2006). Is work good for your health and well-being?

World Health Organization. (1995). *Global strategy on occupational health for all: The way to health at work, recommendation of the second meeting of the who collaborating centres in occupational health, 11-14 october 1994, beijing, china*. World Health Organization.

Chapter 2

DEVELOPMENT AND VALIDATION OF A PERIODIC HEALTH SCREENING TOOL

This chapter is accepted and awaiting publication:
Steel, J. S., Godderis, L., & Luyten, J. (2021a). Development and validation of a triage tool to follow-up workers' health. *Occup. Med. (Lond)*, *(under review)*

This first part of the PhD summarizes the development and validation of an alternative form of periodic occupational health screenings: an electronic screening tool with an algorithm for selective follow-up. The tool is developed and validated using qualitative and quantitative methods, and its content validity, internal consistency, structural validity, and criterion validity is assessed and improved in multiple phases.

Abstract

Background As work and health are closely interlinked, it is important to carefully monitor employees. However, limited resources restrict in-depth follow-up.

Aims To develop a low-cost screening instrument for employees' overall health status, that can be used across industries and that allows triaging workers to in-depth health surveillance in case of indications of health or functioning problems.

Methods We developed a new questionnaire-based algorithm built on multiple predictors to assess the need for further follow-up. We carried out a systematic review, Delphi panel (n=9), and focus group (n=5) to determine the predictors, tested for language pitfalls in a pilot study, and evaluated the questionnaire's validity in two separate studies. Study 1 (n=60) analysed the discriminatory power of the instrument by comparing it to the assessment of an occupational physician in a sample of employees from diverse occupational settings. Study 2 (n=869) appraised the factor structure and internal consistency of the screening tool in a sample of employees from the hospital sector.

Results Risk factors, current physical and mental health, functioning, absenteeism, job satisfaction, and lifestyle were identified as the most relevant predictors. Study 1 showed the survey had good criterion validity (AUC = 0.72). Study 2 (N=869, 28% response) demonstrated the internal consistency (Cronbach's alpha = 0.94), and a factor analysis confirmed a second-order factor structure with adequate model fit (CFI=0.96, RMSEA=0.04 and SRMR=0.07).

Conclusion This questionnaire can be used to triage workers for occupational health follow-up and can, additionally, be useful to describe the epidemiology of work-related illness.

Keywords survey development, validity, reliability, periodic health screening, psychometrics

2.1 Introduction

Work has a significant influence on health, making it important to regularly follow-up employees. Many countries rely on face-to-face periodic health screenings by occupational physicians or nurses to timely detect potential work-related health impairments and diseases (e.g. occupational asthma or burnout), or changes in general health that affect workplace functioning (e.g. musculoskeletal problems or depression) (Colosio et al., 2017). In the EU, these

routine medical examinations are organised regularly for all workers (14 out of 27 EU countries) or only when there is exposure to specific occupational hazards (12 out of 27 EU countries) (Colosio et al., 2017). A common challenge of such workers' health surveillance systems is that they are demanding in terms of working hours of an increasingly pressurized group of occupational physicians (All Party Parliamentary Group on Occupational Safety and Health, 2016; Mairiaux et al., 2010; Torrance & Heron, 2017b). In turn, this has contributed to a diverse coverage and quality of workers' health surveillance across countries, thus undermining universal access (Colosio et al., 2017; EU, 1989a).

Therefore, the question rises whether it is possible to organize at least part of these screenings more efficiently by using an easy-to-use, low-cost screening instrument that can systematically detect work and health problems of employees, so that they can be selectively referred to the occupational physician when there are indications of functioning and/or (work-related) health problems. Some countries have already taken legal steps in the direction of implementing such triage instruments. Belgium, for instance, no longer requires a yearly medical examination by an occupational physician. Some employees will be followed up at two-yearly intervals with a medical questionnaire in between, which serves to identify employees in need of closer follow-up (2006).

Although several occupational health surveys are available, these are not suitable to reliably perform the required two tasks of periodic health screening (systematic detection and selective referral) for a range of health risks in a broad worker population. Available surveys either focus on other goals (e.g. incidence estimation (Cherry et al., 2000; Valenty et al., 2015b)), on specific health dimensions (e.g. the Nurses Work Functioning Questionnaire (Gärtner, Nieuwenhuijsen, Van Dijk, & Sluiter, 2011), General Health Questionnaire (Goldberg, 1978), and Warwick-Edinburgh mental well-being scale (WEMWBS) (Tennant et al., 2007) focus on mental health), or on specific populations (e.g. nurses (Boezeman, Nieuwenhuijsen, & Sluiter, 2016; Gärtner, Ketelaar, et al., 2011; Noben et al., 2014), ambulance workers (Van Schaaijk, Boschman, Frings-Dresen, & Sluiter, 2017), or rail safety workers (Boschman, Hulshof, Frings-Dresen, & Sluiter, 2016)) (Ruitenburg et al., 2016). To our knowledge, psychometrically validated triage questionnaires that can be used for broad populations while detecting a range of health problems and estimating whom a physician should see do not exist.

The aim of the present article is therefore to report on the development and psychometric validation of a new questionnaire-based algorithm for the periodic screening (systematic detection and selective referral) of all employee types, based on a list of the most relevant predictors and frequently occurring health problems.

2.2 Methods

As depicted in Figure 2.1, and advised by De Vet et al. (De Vet, Terwee, Mokkink, & Knol, 2011), we used a systematic review, Delphi panel and focus group to determine which variables should be included in the instrument; tested a first version in a pilot study; and then evaluated the reliability and criterion and structural validity of this questionnaire in two separate studies. These studies received ethical approval from the Ethical Committee of UZ/KU Leuven under number S62259 and OBC number MP008817.

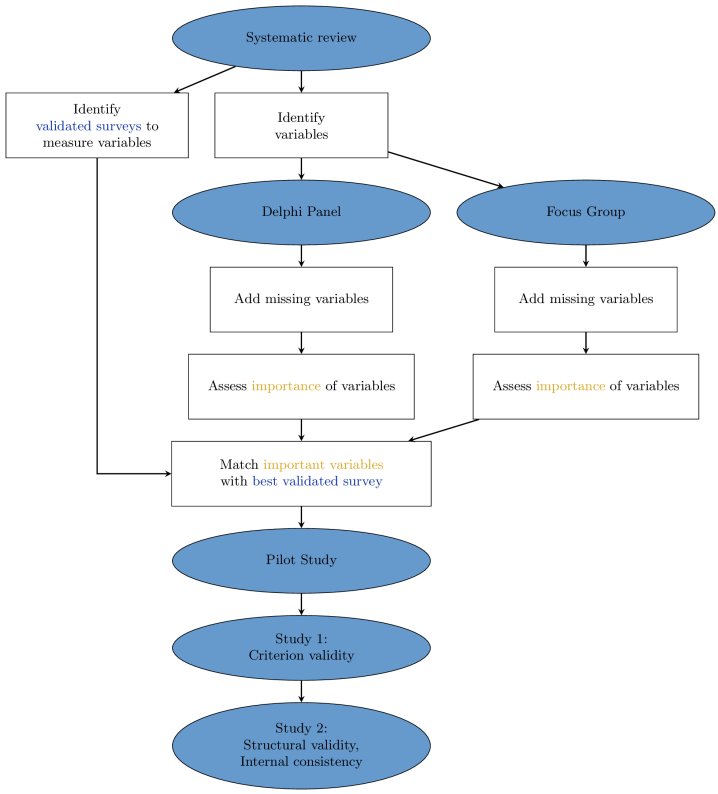


Figure 2.1: Methodology

2.2.1 Questionnaire and algorithm development

For the systematic review, literature searches were performed to identify existing reviews of instruments to measure occupational health, occupational risks, work ability, or health promotion. We refer to the Prospero protocol (ID CRD42019117218) for additional information, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). Two researchers carried out the systematic review screening and eligibility assessment independently. One researcher extracted measurement instruments and psychometric properties (validation and reliability as assessed by the systematic review) from the included systematic reviews, and the topics of the instruments (e.g. job satisfaction, stress, or pain) were categorized. This served to identify relevant variables for the questionnaire, and to identify the best available psychometrically validated and reliable surveys measuring these variables.

Next, the categories and dimensions from the review were used in a three-round online Delphi panel of nine experienced occupational physicians (three of which were employed in the hospital sector, the others worked in a range of industries) to achieve expert consensus on which of the identified dimensions should be linked to a psychometrically validated questionnaire and included in the survey. Before the start of the panel, the process and goal of the Delphi panel and the instrument we aimed to develop was thoroughly explained to the participants. In the first round, participants could add new dimensions to those identified in the review, which were analysed through content analysis, and presented back to the panel in a second round. In the second round, all dimensions were listed with their definitions, and the participants indicated the importance of each dimension in light of the survey's goal on a 5-point Likert scale. The third round presented the individual and group results to the participants, and offered the panel members a chance to adapt their opinion. If $\gtrsim 80\%$ of the panel members scored a dimension with one of the two highest Likert responses, we defined this as consensus on the importance of that dimension (Hasson, Keeney, & McKenna, 2000).

Subsequently, a focus group was conducted among healthcare workers from an assisted-living centre (three nurses, an occupational therapist, and a logistics employee, who all receive yearly occupational screening) to add missing categories from the target population's point of view, and assess their priorities. Participation was on voluntary basis. After an introductory discussion, participants were asked which subjects they would include in the questionnaire, which was the most important, and whether all important aspects had been addressed in the conversation. Categories identified in the systematic review were prompted if the participants did not mention them.

Finally, the categories and dimensions identified as most important by the focus group and Delphi panel were matched with psychometrically validated surveys (identified in the systematic review). Preference was given to short surveys that had been used in a broad employee population and had good psychometric properties, following the COSMIN taxonomy (Mokkink et al., 2010). The resulting instrument was translated to Dutch, no backwards translation was performed. It was consequently tested in a pilot study (n=5) among five cleaning staff members to detect language pitfalls, and the translation and wording was validated by a professional translation team (Leuven Language Institute). For the algorithm, two different scoring systems were applied to the questionnaire responses to calculate the individual scores, which indicated the need to contact the occupational physician: (1) an unweighted aggregation of the answers, and (2) an ordinal scoring system (Appendix A.1.1).

2.2.2 Criterion validation study

Study 1 assessed the criterion validity ("the degree to which the scores of a measurement instrument are an adequate reflection of a gold standard [a perfect or standard way to measure the outcome]" (De Vet et al., 2011)) of the tool tested in the pilot study. It was conducted in three different occupational medical centres in Belgium, which offer occupational health services for various industries. On nine days (three per centre) between May and June 2019 all employees that were attending a periodic health screening with the occupational physician completed the questionnaire prior to the consultation, and their algorithm scores (both scoring system (1) and (2)) were calculated.

By surveying the occupational physician after each consultation, we then measure the physician's assessments for the need for follow-up. We construct two different indicators to reflect that it is generally difficult to find a perfect (or "gold") criterion, and to assess how changing the criterion (the assessment by the physician) affects the results. One question was used as a direct indication for the need for follow-up: 1) the necessity of the consultation considering the problems (s)he detected (six options, not necessary at all - very necessary). Its top two answers were defined as cases (need for follow-up) while the bottom four answers were defined as controls (no need for follow-up). Questions 2 to 5 were used as an indirect reflection of the need for follow-up: 2) physical; and 3) mental burdens of his work (five options, very badly - very good); 4) the workability of the employee on a numeric scale (0 - 10); and 5) workability by category (six options, unable to work - optimal workability). Questions 2 to 5 were aggregated to a single score: a low (high) score indicates a low (high) need for follow-up, and cut-off at the top 20%.

Receiver Operating Curves (ROC) and the associated Area Under the Curve (AUC) were then used to assess sensitivity and specificity (the algorithms' ability to correctly classify an individual as someone who needs follow-up, or not, respectively) at these cut-off points (De Vet et al., 2011). Youden's index was calculated to search for the optimal cut-off point of the algorithm, it equals sensitivity + specificity - 1, and ranges from 0 (no discrimination of true or false cases) to 1 (perfect test, no false positive or false negatives).

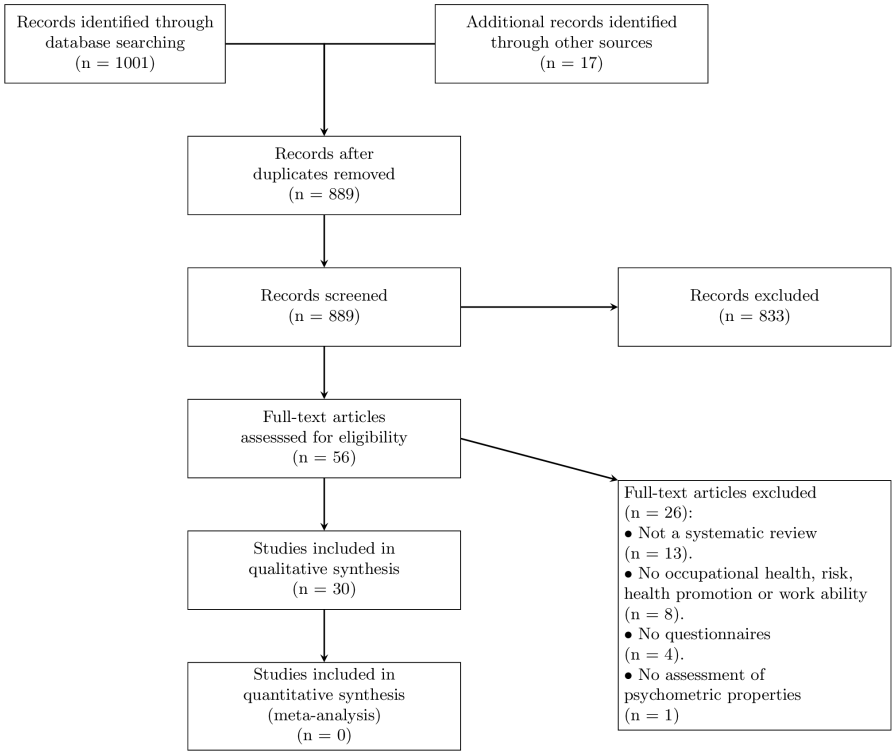
Two sensitivity analyses used different cut-offs for the indirect assessment: the 53% and 31% highest scoring individuals instead of the top 20%. These values are consistent with, respectively, the proportion of individuals in the top two categories of the necessity question (53%), and the proportion of individuals in the highest category of the necessity question (31%). A third sensitivity analysis used the employee's own assessment of whether they are able to cope with the mental burdens of their work, as opposed to the physicians' assessment, as it can be difficult for employees to disclose mental health issues to occupational physicians (J. S. Steel et al., 2021b).

2.2.3 Structural validity and internal consistency study

Study 2 was conducted in four Belgian hospitals between June and October 2019 to evaluate structural validity ("the degree to which the scores of a measurement instrument are an adequate reflection of the dimensionality of the construct to be measured" (De Vet et al., 2011)) and internal consistency ("the degree of interrelatedness among the items" (De Vet et al., 2011)). In the hospitals, which were recruited on voluntary basis by their occupational physician, employees who were eligible for yearly periodic health screenings were invited by email or letter to participate in the study.

As suggested by De Vet et al (De Vet et al., 2011), a confirmatory factor analysis (CFA) was performed to test whether the data fit a priori hypothesized factor structures based on the systematic review and confirmed by a Delphi panel. Model fit was assessed by the comparative fit index (CFI > 0.95), the root mean square error of approximation (RMSEA < 0.06) and the standardized root mean square residual (SRMR < 0.08) (De Vet et al., 2011). Cronbach's alpha was calculated as a measure of the internal consistency of the questionnaire. Acceptable Cronbach's alpha values are between 0.70 – 0.95.

Figure 2.2: PRISMA flow diagram



2.3 Results

2.3.1 Questionnaire and algorithm development

The systematic review led to the qualitative synthesis of 30 systematic reviews of questionnaires (PRISMA flow diagram in Figure 2.2). Within these reviews, 443 unique questionnaires were identified: 111 questionnaires related to physical exposure, 108 on mental health, 54 on physical health, 47 on the psychosocial work environment, 42 questionnaires on lifestyle variables, 40 related to productivity, 29 on work (in)ability, and 12 on work status. After extracting their psychometric properties, their topics and contents were subsequently categorized into 12 categories and 82 dimensions and presented to the nine members of the Delphi panel. The panel added 15 dimensions to a total of 107 (Appendix A.1). After the third round, 25 dimensions were seen as important

or very important by 89% of the panel.

The focus group indicated 10 of the 12 categories from the systematic review as important to include without prompts from the moderator (Appendix A.1). After prompting, the other categories were also deemed relevant. Four categories were seen as pivotal to include: physical health (especially musculoskeletal), mental health, social work environment, and workplace well-being and satisfaction.

Next, the dimensions that were deemed important were matched to a psychometrically validated survey identified in the systematic review. For 22 of the 25 dimensions, a questionnaire was thus identified in the systematic review (Table 2.2). For the remaining three dimensions – risk on mental or psychosocial burden; neoplasm; and endocrine, food, or metabolic diseases – no suitable questionnaire was available. After translating to Dutch where necessary, this first version of the survey was tested for clarity in a pilot study. Adjustments were made to minimize language barriers and complexity. This version of the questionnaire was then used in Study 1.

2.3.2 Criterion validation study

Study 1 had a sample size of 58 employees (two answers were incomplete) with various occupational backgrounds. Average time to completion was 15 minutes. Descriptive statistics of the sample are available in Appendix A.2, 31% of the consultations was in the highest necessity category in question 1, 53% was in the top two answer categories. In the baseline analysis, the ordinal scoring system (2) had the best criterion validity, with AUC values of 0.72 [ci=0.55-0.89] for the indirect criterion, and 0.72 [ci=0.58-0.86] for the direct criterion (Figure 2.3). Youden analysis (Table 2.1) showed one optimal cut-off point (at a score of 30.81) with a sensitivity of 0.64, specificity of 0.78, and a Youden index of 0.42.

The first two sensitivity analyses (Table 2.1) resulted in better values for the 53% criterion, but worse values for the 31% criterion. The third sensitivity analysis resulted in an AUC of 0.74 [ci=0.53-0.95], with better overall values. For the latter case, note that while the higher values are partially the logical consequence of the estimation, it could also indicate that the appraisal of mental health by the physicians is less consistent with the employee's view.

Figure 2.3: ROC curves for the ordinal scoring mechanism against the assessment by the occupational physician. A=How necessary is this PHS?, B=Sum of 4 questions: how well do you think the employee stands up against physical demands of his work? How well do you think the employee stands up against mental demands of his work? Workability (1-10), and workability by category. For panel A, the top two answers ("4" & "5 very necessary" = 1) are compared to the answers below ("0 not at all necessary" & "1" & "2" & "3" =0). In panel B, the 20% poorest scoring individuals on the questions are cases (=1).

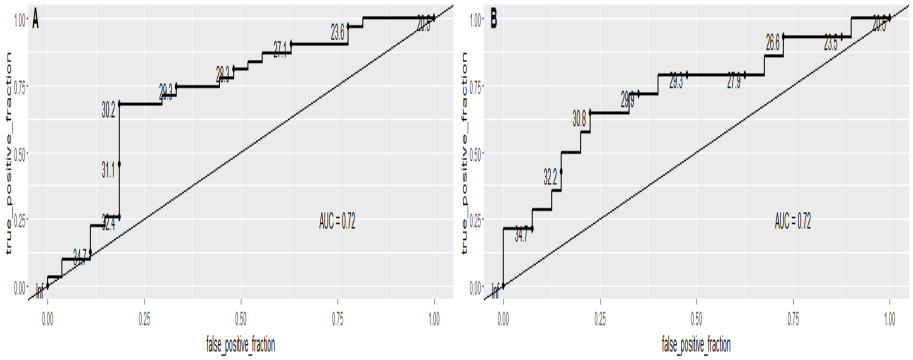


Table 2.1: Youden analyses for different case definitions by the physician: 20% worst scores, 31% worst scores, 53% worst cases, and a criterion with employee mental health assessment. AUC=Area Under ROC Curve, PPV=Positive Predictive Value, NPV=Negative Predictive Value, DLR=Diagnostic Likelihood Ratio, FP=False Positives, FN=False Negatives, CI=confidence interval

	Criterion 20% (baseline)	**Criterion 31% (two optimal cut-off points)**		**Criterion 53%**	**Criterion employee mental health**
AUC	0.72	0.69		0.75	0.74
(CI)	(0.55 - 0.89)	(0.53-0.84)		(0.62-0.88)	(0.53-0.95)
cutoff	30.81	30.19	30.75	30.19	30.81
Sensitivity	0.64	0.67	0.61	0.61	0.73
Specificity	0.78	0.69	0.75	0.83	0.77
PPV	0.5	0.52	0.55	0.83	0.44
NPV	0.86	0.81	0.79	0.61	0.92
DLR.Positive	2.86	2.18	2.44	3.52	3.13
DLR.Negative	0.46	0.48	0.52	0.47	0.36
FP	9	11	9	4	10
FN	5	6	7	12	3
Criterion value	0.42	0.36	0.36	0.44	0.49

2.3.3 Structural validity and internal consistency study

Study 2 had an eligible population of 3150 employees, with an effective sample size of 869 employees (776 fully completed the survey), meaning on average 28% responded (and 25% fully completed the survey). Descriptive statistics of the sample are shown in Appendix A.3.

Table 2.2: Questionnaire sources. Variables included in the screening tool (agreed upon by the Delphi panel and focus group), questionnaire source used per variable, and their validity and reliability. /=no suitable questionnaire identified; no validity or reliability tests available

Category	Dimension	Questionnaire source	Validity & Reliability
Job content	Risk On Mental Or Psychosocial Burden	/	/
Employment conditions	Work Time (Shift Or Nightwork)	custom question on night-, shift-, or weekendwork	/
Social work environment	Role Clarity Or Conflicts	WOCCQ (Hansez, 2008)	Construct validity (structural validity, hypothesis testing)
	Bullying	COPSOQ III (2)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
	Risk On Violence	HIS	/
Attitude towards job organisation	Job Satisfaction	COPSOQ III (Kristensen, Hannerz, Høgh, & Borg, 2005)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
	Happiness And Enthousiasm At Work	COPSOQ III (Kristensen et al., 2005)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
Productivity	Abseenteism	iPCQ (Bouwmans et al., 2015)	Reliability, content and face validity, cross-cultural validity, criterion validity (concurrent), construct validity (hypothesis testing)
Mental Health	Recovery Needs After Work	NFR (De Croon, Sluiter, & Frings-Dresen, 2006)	Internal consistency, reliability, Cross-cultural validity, construct validity (hypothesis testing), responsiveness
	General Mental Health And Well-Being	GHQ-12 (Goldberg, 1972)	Internal consistency, reliability, construct validity (structural, hypothesis testing)
	Stress	COPSOQ III (Kristensen et al., 2005)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
	Burnout	COPSOQ III (Kristensen et al., 2005)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
Physical Health	General Health	EQ-5D (EuroQol Research Foundation, 2019)	Reliability, content and face validity, construct validity (structural validity, hypothesis testing), responsiveness
	Neoplasma	/	

	Endocrine, Food, Or Metabolical Diseases	/	
	Sleep-Wake Diseases	COPSOQ III (Kristensen et al., 2005)	Internal consistency, content and face validity, criterion validity (concurrent), construct validity (structural, cross-cultural, hypothesis testing)
	Diseases Of The Respiratory System	Breathlessness scale (Fletcher, Elmes, Fairbairn, & Wood, 1959)	Criterion validity (concurrent), construct validity (hypothesis testing)
	Diseases Of The Musculoskeletal System/Connective Tissue	NMQ (Kuorinka et al., 1987)	Reliability, content and face validity, criterion validity (concurrent and predictive)
	Allergies Or Contactdermatitis	NOSQ (Susitaival et al., 2003)	Criterion validity, cross-cultural validity
	Pain	NMQ (Kuorinka et al., 1987)	Reliability, content and face validity, criterion validity (concurrent and predictive)
	Other Health Problems	custom question	/
Medication and healthcare use	Medication Use	HIS	/
	Medication Dependance	HIS	/
Health(care) behaviour and literacy	Lifestyle	AUDIT-C (Bush et al. 1998); HIS	AUDIT-C: reliability, internal consistency, criterion validity (concurrent), cross-cultural validity, construct validity (structural validity, hypothesis testing) HIS: test-retest reliability, concurrent validity
	Spontaneous Consultation	custom question	/

The models tested in CFA relate a number of latent variables (factors) to manifest variables (questions) by estimating the factor loadings. The latter represent the strength of the relationship between the latent variable and the manifest variables. A first model included all factors shown in Table 2.2, but the model showed high collinearity between variables and low frequency of some health problems. A 20-factor model included musculoskeletal problems and skin problems (both split out as separate factors by region), need for recovery, general mental health, job satisfaction, sleeping problems, burnout, stress, role conflicts, lifestyle (alcohol abuse and smoking), and absenteeism, with CFI=0.99, RMSEA=0.03, and SRMR=0.07. The final model used a second-order factor ('Follow-up'), removed skin problems and lifestyle, maintained only musculoskeletal functioning questions (complaints in the last 12 months and last seven days were removed), and demonstrated CFI=0.96, RMSEA=0.05, and SRMR=0.07.

The structure and estimates of the final model are depicted in Figure 2.4 and Table 2.3, showing the central second-order factor ('Follow-up'), its relation to the other factors, and the estimates of each factor with its manifest variables.

Given a few exceptions, all estimates were well above 0.50. The factor loadings of the second-order factor are highest with regard to the mental health and psychosocial factors.

The internal consistency of the final model was good, with a Cronbach's alpha of 0.94 for all questions taken together. All Cronbach's alpha coefficients of the subscales were higher than 0.75, indicating adequate internal consistency: Musculoskeletal=0.76, NFR=0.86, GHQ=0.88, Job Satisfaction=0.84, Sleep=0.86, Burnout=0.91, Stress=0.88, and Role Conflicts=0.78. The same goes for the unused scales (present in the 20-factor model, but not in the second-order model) of Skin problems at the wrist=0.89, and Skin problems at the hand=0.87.

2.4 Discussion

We developed an occupational health screening tool aimed at a broad employee population, containing questions on diverse and frequently occurring health problems. Although inferred from a limited sample size, the instrument has an acceptable but not excellent ability (Hosmer & Lemeshow, 2000) to identify employees in need of a follow-up by the occupational physician (AUC=0.72 [ci=0.58-0.86]), which is similar to other questionnaires in this field (Chen et al., 2020). The internal consistency of the algorithm was excellent ($\alpha = 0.94$). The 20-factor model of the questionnaire had the best fit, but the second-order model also met all criteria of adequate fit with the added benefit of being interpretable as a single score.

In line with our expectations, the process we followed resulted in a diverse instrument that captures frequent health problems. The algorithm emphasizes mental health, psychosocial risk, and musculoskeletal complaints, all of which are among the leading health problems, both in and outside of the healthcare sector (Godderis et al., 2015). Contrary to our expectations, several factors initially deemed important by the focus group and Delphi panel were not retained in the final second-order model. This need not be problematic, as some were too infrequent in the sample (e.g. respiratory problems), while others rely strongly on the employees' willingness to disclose information (J. S. Steel et al., 2021b), e.g. drug abuse. Both could be detected through other means, e.g. a face-to-face consultation with the physician, or job- and task specific questionnaires.

Previous occupational screening questionnaires focused on one domain (Gärtner, Nieuwenhuijsen, Van Dijk, & Sluiter, 2011; Hildebrandt, Bongers, Van Dijk, Kemper, & Dul, 2001), or on specific jobs and tasks (Boezeman et al., 2016; Boschman et al., 2016; Gärtner, Ketelaar, et al., 2011; Noben et al., 2014;

Ruitenburg et al., 2016; Van Schaaijk et al., 2017). In contrast, our instrument included diverse health problems, and the focus group and structural validation study took place in the hospital sector, which houses a broad population (e.g. technicians, administrative personnel, caretakers, cleaning staff), and is prone to a wide range of risks (physical, chemical, psychological, biological). Although this should be investigated more closely, this suggests that the instrument is also suitable outside the healthcare sector.

For clinicians the developed instrument can be used in this specific triage setting, but once implemented it will also be useful for epidemiological studies (Appendix A.3) and to inform occupational health practice. In addition, the algorithm can increase the efficiency of occupational health screenings, focusing resources where they are needed the most while still upholding universal monitoring and access to the occupational health services.

Although the results on criterion validity are encouraging, the sample size was limited, and more reliability and psychometric validity tests are needed before implementing the survey on a large scale, preferably following the COSMIN taxonomy (De Vet et al., 2011; Mokkink et al., 2010). Future research could include a second criterion analysis, a predictive validity study, or item scores can be weighted according to the importance of each subdomain in the survey (De Vet et al., 2011). To increase the instrument's relevance, it would be beneficial to add risks specific to an employee's occupation (Ruitenburg et al., 2016; J. K. Sluiter, 2017).

We conclude that the questionnaire-based algorithm has adequate face validity, criterion validity, structural validity, and internal consistency, and can be useful in occupational health screening and epidemiological studies.

Figure 2.4: Plot of the second-order factor model (standardized estimates, also available in Table 2.3.)

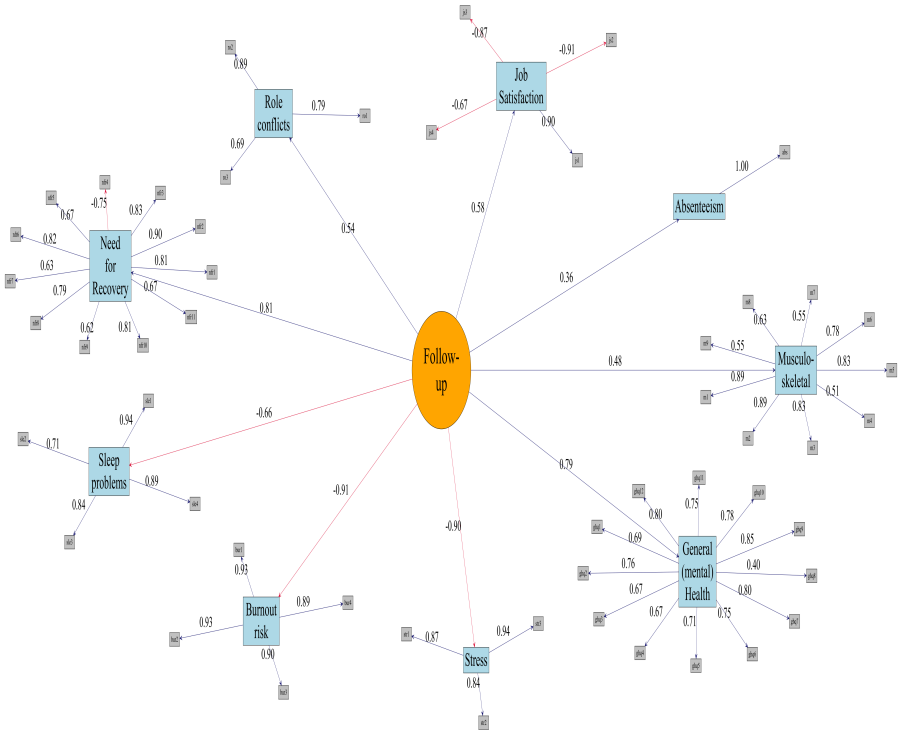


Table 2.3: CFA (standardized) estimates of the second-order factor model with robust standard errors. Estimates reflect factor loadings of each manifest variable on the latent variable. sleep, burnout and stress are reversed scales. All estimates are significant at 0.01 confidence

LATENT VARIABLE	MANIFEST	ESTIMATE	STANDARD ERROR	LATENT VARIABLE	MANIFEST	ESTIMATE	STANDARD ERROR
Musculoskeletal	neck	0.88	0.04	GHQ (cont.)	GHQ11	0.76	0.03
	upper back	0.88	0.05		GHQ12	0.8	0.02
	shoulder	0.84	0.04	Job Satisfaction	job satisfaction	0.9	0.02
	wrists-hands	0.52	0.08		enthusiasm1	-0.91	0.02
	lower back	0.83	0.04		enthusiasm2	-0.87	0.02
	hip thigh	0.79	0.06		enthusiasm3	-0.67	0.03
	elbow	0.57	0.09	Sleep	sleep1	0.94	0.01
	ankle-feet	0.63	0.08		sleep2	0.71	0.03
	knee	0.55	0.08		sleep3	0.84	0.01
Need For Recovery	nfr_1	0.84	0.03		sleep4	0.89	0.01
	nfr_2	0.89	0.02	Burnout	burnout1	0.93	0.01
	nfr_3	0.83	0.02		burnout2	0.93	0.01
	nfr_4	-0.75	0.04		burnout3	0.9	0.01
	nfr_5	0.67	0.04		burnout4	0.89	0.01
	nfr_6	0.82	0.03	Stress	stress1	0.87	0.01
	nfr_7	0.62	0.04		stress2	0.84	0.01
	nfr_8	0.79	0.03		stress3	0.94	0.01
	nfr_9	0.61	0.04	Role conflicts	role1	0.79	0.03
	nfr_10	0.81	0.03		role2	0.9	0.02
	nfr_11	0.67	0.04		role3	0.69	0.03
General (mental) Health	GHQ1	0.69	0.03	Absenteeism	abs.days	1	0
	GHQ2	0.77	0.02	Follow-up (2nd order)	Musculoskeletal	0.48	0.04
	GHQ3	0.64	0.03		Need for recovery	0.82	0.02
	GHQ4	0.67	0.03		General mental health	0.78	0.02
	GHQ5	0.72	0.03		Job satisfaction	0.58	0.03
	GHQ6	0.76	0.02		Sleep	-0.67	0.02
	GHQ7	0.8	0.02		Burnout	-0.92	0.01
	GHQ8	0.4	0.04		Stress	-0.91	0.01
	GHQ9	0.86	0.02		Role conflicts	0.55	0.03
	GHQ10	0.79	0.02		Absenteeism	0.23	0.03

KEY POINTS

What is already known about this subject

- The combination of a drop in occupational health personnel and limited (financial) resources makes it difficult to provide in-depth periodic health screening to all employees.
- A screening tool could help in this process, but few psychometrically valid and reliable surveys exist that can adequately detect and triage diverse employees who are at risk.
- Surveys that have been developed hitherto focus on specific occupational health topics (mental health, musculoskeletal disorders) or specific subpopulations (nurses, ambulance workers, or rail safety workers), and are not suitable for a broad employee population.

What this study adds

- A new generalized screening tool has been developed based upon literature, expert opinion, and the target population, aimed at a broad population and making use of frequent health problems.
- The face and content validity, criterion validity, structural validity, and internal consistency of the instrument is good and acceptable.
- The recurring health and risk data from the survey could prove useful for workers' follow-up and for epidemiological studies.

What impact this may have on practice or policy

- The survey could help organize periodic health screenings more efficiently, targeting mainly those that benefit from a resource-intensive follow-up.

References

All Party Parliamentary Group on Occupational Safety and Health. (2016). Occupational medical workforce crisis. the need for action to keep the uk workforce healthy. All Party Parliamentary Group on Occupational Safety and Health London.

Boezeman, E. J., Nieuwenhuijsen, K., & Sluiter, J. K. (2016). Predictive value and construct validity of the work functioning screener–healthcare (wfs-h). *J Occup Health*, 15-0056–OA. doi:10.1539/joh.15-0056-OA

Boschman, J., Hulshof, C., Frings-Dresen, M., & Sluiter, J. (2016). Improving fit to work assessments for rail safety workers by exploring work limitations. *Int Arch Occup Environ Health*, *89*(5), 803–811. doi:10.1007/s00420-016-1117-7

Bouwmans, C., Krol, M., Severens, H., Koopmanschap, M., Brouwer, W., & Hakkaart-van Roijen, L. (2015). The imta productivity cost questionnaire: A standardized instrument for measuring and valuing health-related productivity losses. *Value in health*, *18*(6), 753–758.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., & Bradley, K. A. (1998). The audit alcohol consumption questions (audit-c): An effective brief screening test for problem drinking. *Archives of internal medicine*, *158*(16), 1789–1795.

Chen, Y.-C., Guo, Y.-L. L., Lin, L.-C., Lee, Y.-J., Hu, P.-Y., Ho, J.-J., & Shiao, J. S.-C. (2020). Development of the nurses' occupational stressor scale. *Int J Environ Res Public Health*, *17*(2), 649. doi:10.3390/ijerph17020649

Cherry, N., Meyer, J., Holt, D., Chen, Y., & McDonald, J. (2000). Surveillance of work-related diseases by occupational physicians in the uk: Opra 1996–1999. *Occupational Medicine*, *50*(7), 496–503.

Colosio, C., Mandic-Rajcevic, S., Godderis, L., van der Laan, G., Hulshof, C., & van Dijk, F. (2017). Workers' health surveillance: Implementation of the directive 89/391/eec in europe. *Occupational Medicine*, *67*(7), 574–578.

De Croon, E. M., Sluiter, J. K., & Frings-Dresen, M. H. (2006). Psychometric properties of the need for recovery after work scale: Test-retest reliability and sensitivity to detect change. *Occupational and environmental medicine*, *63*(3), 202–206.

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

EU. (1989a). *Council directive 89/391/eec of 12 june 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work*. Official Journal of the European Communities.

EuroQol Research Foundation. (2019). *Eq-5d-5l user guide*.

Fletcher, C. M., Elmes, P. C., Fairbairn, A. S., & Wood, C. H. (1959). Significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *British medical journal*, *2*(5147), 257.

Gärtner, F. R., Ketelaar, S. M., Smeets, O., Bolier, L., Fischer, E., van Dijk, F. J., ... Sluiter, J. K. (2011). The mental vitality@ work study: Design of a randomized controlled trial on the effect of a workers' health surveillance mental module for nurses and allied health professionals. *BMC Public Health*, *11*(1), 1–13. doi:10.1186/1471-2458-11-290

Gärtner, F. R., Nieuwenhuijsen, K., Van Dijk, F. J., & Sluiter, J. K. (2011). Psychometric properties of the nurses work functioning questionnaire (nwfq). *PLoS One*, *6*(11), e26565. doi:10.1371/journal.pone.0026565

Godderis, L., Mylle, G., Coene, M., Verbeek, C., Viaene, B., Bulterys, S., & Schouteden, M. (2015). Data warehouse for detection of occupational diseases in ohs data. *Occup Med (Lond)*, *65*(8), 651–658.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness*. Swindon, Wiltshire, UK: Oxford U. Press.

Goldberg, D. (1978). *Manual of the general health questionnaire*. Windsor, England: NFER Publishing.

Hansez, I. (2008). The "working conditions and control questionnaire"(woccq): Towards a structural model of subjective stress. *European review of Applied psychology*, *58*(4), 253–262.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the delphi survey technique. *J Adv Nurs*, *32*(4), 1008–1015. doi:10.1046/j.1365-2648.2000.t01-1-01567.x

Hildebrandt, V., Bongers, P., Van Dijk, F., Kemper, H., & Dul, J. (2001). Dutch musculoskeletal questionnaire: Description and basic qualities. *Ergonomics*, *44*(12), 1038–1055. doi:10.1080/00140130110087437

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.

Kristensen, T. S., Hannerz, H., Høgh, A., & Borg, V. (2005). The copenhagen psychosocial questionnaire-a tool for the assessment and improvement of the psychosocial work environment. *Scand J Work Environ Health*, 438–449.

Kuorinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sorensen, F., Andersson, G., & Jorgensen, K. (1987). Standardised nordic questionnaires for the analysis of musculoskeletal symptoms. *Appl Ergon*, *18*. doi:10.1016/0003-6870(87)90010-x

Mairiaux, P., Chanteux, A., Schleich, E., & Delwick, B. (2010). *Financement des missions de prévention des services externes de prévention et de protection au travail*. SPF Emploi, travail et concertation sociale.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS medicine*, *6*(7), e1000097.

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... De Vet, H. C. (2010). The cosmin checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol*, *10*(1), 1–8. doi:10.1186/1471-2288-10-22

Noben, C., Smit, F., Nieuwenhuijsen, K., Ketelaar, S., Gärtner, F., Boon, B., ... Evers, S. (2014). Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud*, *51*(10), 1321–1331. doi:10.1016/j.ijnurstu.2014.01.017

Ruitenburg, M. M., Frings-Dresen, M. H., & Sluiter, J. K. (2016). How to define the content of a job-specific worker's health surveillance for hospital physicians? *Saf Health Work*, *7*(1), 18–31. doi:10.1016/j.shaw.2015.08.004

Sluiter, J. K. (2017). The new why when designing mandatory medical examinations. *Occup Med (Lond)*, *67*(5), 325–327. doi:10.1093/occmed/kqx044

Steel, J. S., Godderis, L., & Luyten, J. (2021a). Development and validation of a triage tool to follow-up workers' health. *Occup. Med. (Lond)*, *(under review)*.

Steel, J. S., Godderis, L., & Luyten, J. (2021b). Disclosure in online vs. face-to-face occupational health screenings: A cross-sectional study in belgian hospital employees. *International Journal of Environmental Research and Public Health*, *18*(4), 1460. doi:10.3390/ijerph18041460

Susitaival, P., Flyvholm, M.-A., Meding, B., Kanerva, L., Lindberg, M., Svensson, Å., & Olafsson, J. (2003). Nordic occupational skin questionnaire (nosq-2002): A new tool for surveying occupational skin diseases and exposure. *Contact dermatitis*, *49*(2), 70–76.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., ... Stewart-Brown, S. (2007). The warwick-edinburgh mental well-being scale (wemwbs): Development and uk validation. *Health and Quality of life Outcomes*, *5*(1), 63.

Torrance, I., & Heron, R. (2017b). Occupational health should be part of the nhs. Generic. British Medical Journal Publishing Group.

Valenty, M., Homère, J., Lemaitre, A., Plaine, J., Ruhlman, M., Cohidon, C., & Imbernon, E. (2015b). Surveillance programme for uncompensated work-related diseases in france. *Occup Med (Lond)*, *65*(8), 642–650. doi:10.1093/occmed/kqv104

Van Schaaijk, A., Boschman, J., Frings-Dresen, M., & Sluiter, J. (2017). Appraisal of work ability in relation to job-specific health requirements in ambulance workers. *Int J Environ Res Public Health*, *90*(1), 123–131. doi:10.1007/s00420-016-1181-z

Chapter 3

EFFECTIVENESS OF PERIODIC HEALTH SCREENING

This chapter will be submitted as:
Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals:

This part reports a randomized controlled trial of periodic health screening in the Belgian hospital sector, wherein the effectiveness of the electronic screening tool is evaluated vis-à-vis a face-to-face health screening in the short term. In the control arm of the experiment, employees all have a face-to-face periodic health examination with the occupational physician. In the intervention arm, employees complete an electronic health screening survey, after which the worst scoring 20% (i.e. the least healthy) will be referred to the occupational physician for a consultation.

Abstract

Objectives In many countries there is a legal obligation for employers to regularly screen their employees by an occupational physician. However this system is time-intensive and there may be more cost-effective alternatives. Our objective is to compare the effectiveness of periodic occupational health screening of hospital employees by a physician, with a system of electronic screening with targeted follow-up by an occupational physician.

Methods A randomised controlled trial was set up among personnel of four Belgian hospitals, with three measurement moments between June 2019 and December 2020, to compare differences in relevant outcomes: self-assessed health, healthcare use, productivity and intermediate outcomes over 19 months. Mixed effects models were used to assess differences in effectiveness. Superiority and non-inferiority post-hoc tests were used as a robustness check. The experiment coincided with the first two covid-19 waves during which hospital employees were exposed to an exceptional period of occupational stress.

Results In total, 1077 employees (34% of the target population) participated. Although we observed some immediate effects of the intervention (lower trust in the physician, absenteeism, and healthcare use), over time all these effects disappeared. After 19 months, including two waves of covid-19 hospitalizations, no significant differences were observed between employees screened through face-to-face contact and those screened electronically.

Conclusion Our study finds no indication that, in the short-term, substituting workforce screening by a physician with a quicker survey-based screening with targeted follow-up has different effects on the studied endpoints. However, as health and disease are often the result of a long-term process, more evidence is needed to determine long-term effects.

Keywords health screening; occupational stressors; mixed-effect models; multilevel modelling; randomised controlled experiment

3.1 Introduction

Work and health are intricately linked. On the one hand, evidence points out that work is an important source of value and meaning (Roe & Ester, 1999; Ros et al., 1999), and has positive effects on health (Désiron et al., 2015; Waddell & Burton, 2006). On the other hand, the workplace can be a hazardous environment when employees are confronted with physical, chemical, biological, or psychosocial hazards (Takala et al., 2014; World Health Organization, 1995).

Table 3.1: Implementation of EU Directive 89/391/EEC in the EU (and UK); n=28; based on (Colosio et al., 2017)

Question	Answer	n	Countries
Is health surveillance offered (as in Article 14)	Yes	28	All
For all workers?	All	15	BG, HR, CZ, EE, FI, DE, HU, LU, NL, PL, RO, SK, SI, ES, UK
	Groups of workers	12	AT, BE, CY, FR, GR, IE, IT, LV, LT
Obligatory for employers to offer?	Yes	27	All except DK
Obligatory for employees to attend?	Yes	21	AT, BE, HR, CY, CZ, EE, FI, FR, GR, HU, IE, IT, LV, LT, LU, MT, PL, RO, SK, SI, UK
Financing	Employer	22	BE, BG, HR, CY, CZ, EE, FI, FR, GR, HU, IE, IT, LU, NL, PL, PT, RO, SK, SI, ES, SE, UK
	Other	5	MT, LV, LT, AT, DE

AT, Austria; BE, Belgium; BG, Bulgaria; HR, Croatia; CY, Cyprus; CZ, Czech Republic; DK, Denmark; EE, Estonia; FI, Finland; FR, France; DE, Germany; GR, Greece; HU, Hungary; IE, Ireland; IT, Italy; LV, Latvia; LT, Lithuania; LU, Luxembourg; MT, Malta; NL, Netherlands; PL, Poland; PT, Portugal; RO, Romania; SK, Slovakia; SI, Slovenia; ES, Spain; SE, Sweden; UK, United Kingdom.

To minimise the impact of occupational risks on health and functioning, many countries make use of periodic occupational health screenings. In the US, this is regulated by the Occupational Safety and Health administration (OSHA), which requires medical screening for numerous hazardous substances (A. M. Herman & Jeffress, 2000). In the EU (and UK), these routine medical examinations are offered to nearly all workers on an annual basis (see Table 3.1) (Colosio et al., 2017; EU, 1989b). EU-OSHA's European Survey of Enterprises on New and Emerging Risks (ESENER), a survey among European workplaces, indicated that 65% of establishments (58% for micro and 89% for large enterprises) arrange regular medical examinations to monitor the health of employees (EU, 2017). These periodic health screenings can include measuring biometrics, and offer a clinical investigation by an occupational physician (Colosio et al., 2017; EU, 1989b). They focus on primary (e.g. vaccination), secondary (e.g. screening for diseases), and tertiary prevention (e.g. return to work), often take a job-specific risk analysis as a starting point, and can lead to both individual (employee) and collective (worksite) measures (International Labour Office, 1998; J. Sluiter et al., 2013b).

However, there are as of yet no strong research indications for their effectiveness. A systematic review of general health screening found 14 studies of sufficient quality, and did not find studies that indicated an effect on all-cause mortality, cardiovascular mortality, or cancer mortality. In contrast, it suggested they could lead to adverse effects such as overtreatment, misplaced reassurance, or

(in the case of false positives) unnecessary worries about health (Krogsbøll et al., 2013). Even if only high risk individuals are targeted, it remains an open question whether their screening is beneficial (MacAuley, 2012).

Another common challenge of such workers' health surveillance systems is that they are demanding in terms of working hours. In Belgium for instance, about 70% of the employees is regularly seen by the occupational physician, thus taking up a significant portion of both employees' and physicians' time (Lambrechts et al., 2015). Meanwhile, occupational physicians are proving more and more difficult to recruit, and their numbers are difficult to increase (All Party Parliamentary Group on Occupational Safety and Health, 2016; Mairiaux et al., 2010; Torrance & Heron, 2017a). Therefore, the question rises whether it is possible to allow for alternative means of screening workers, which are less labour-intensive without unduly compromising effectiveness.

A more targeted approach might be more promising. A possible alternative is therefore to use a screening survey that selectively refers employees to the occupational physician when it detects indications of functioning and/or (work-related) health problems. This could save occupational physicians' time, and it is a priori an attractive option for employers and employees. The former might save on screening costs, the latter might prefer a survey's practical ease (an online survey can be completed everywhere). Research also points out that some health problems are more easily disclosed in surveys than in face-to-face consultations (J. S. Steel et al., 2021b). Several countries already use surveys as surveillance systems to monitor occupational health, but not for the purpose of triage and screening (Centers for Disease Control and Prevention (CDC), 2018; Cherry et al., 2000; Valenty et al., 2015a). Belgium has taken legal steps in the direction of implementing novel triage instruments, and allows some employees to be followed up at two-yearly intervals with a medical questionnaire in between, which serves to identify employees in need of closer follow-up ("Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017", n.d.). A new instrument has therefore been developed especially to this end (J. S. Steel et al., 2021a).

In this study, we present the results of a randomised trial in four Belgian hospitals, that compares care-as-usual (periodic health screenings by the occupational physician) with employees who complete an electronic health survey with selective follow-up (the intervention group). The experiment coincided with the first two covid-19 waves. Since this affects both control and intervention groups equally, this allowed us to assess the effects of the intervention during a time that was, particularly in the hospital sector, characterised by heightened stress, work pressure, health complaints, and demand for health services.

3.2 Methods

3.2.1 Experimental design

Population

The study protocol is published on ClinicalTrials.gov (Identifier NCT04684316). In a population of 53 Flemish hospitals , we recruited four large hospitals that were willing to participate in this study. In these four hospitals, 3150 employees were eligible for periodic health screenings: personnel with safety functions, jobs with heightened vigilance, work that involves physical, biological or chemical agents or tasks that are an ergonomic or mental burden ("Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017", n.d.). Occupational groups that perform especially risky activities (frequent exposure to ionizing radiation, preparation of cytostatics, or exposure to carcinogens, mutagens, or reprotoxic substances) are excluded from the study population, as it is deemed that in these cases the electronic survey does not constitute adequate care that minimises health risks.

We estimated that at minimum 1700 employees (selected from four hospitals) must be recruited (with an assumed dropout of 50%, effectively participating 850), equally divided over the two groups. This is grounded on power analyses in Stata 2014 MP, estimating sample size for a two-sample means test assuming equal standard deviations in intervention and control group, power 0.9 and significance level of 0.05, based on three data sources (Hackl, Halla, Hummer, & Pruckner, 2015; Rongen, Robroek, van Lenthe, & Burdorf, 2013; Sørensen et al., 2015)

Design

Employees were allocated to one of two groups: the intervention group where employees received an electronic survey with selective follow-up by the occupational physician, and the control group where all employees underwent a periodic health screening by the occupational physician. Some employees (that had gone the longest without a consultation) already received a periodic health screening between January and June 2019, and were therefore assigned to the non-random control. In May 2019, the remaining participants were randomised between control and intervention group until both groups were of equal size (1575 employees). The remaining PHS consultations (random control group and employees of the intervention group that were referred by

the algorithm) occurred between July and December 2019. The quality of the randomisation is evaluated by comparing the descriptive statistics of control and intervention group, and of the samples against aggregate (population) data from the hospitals.

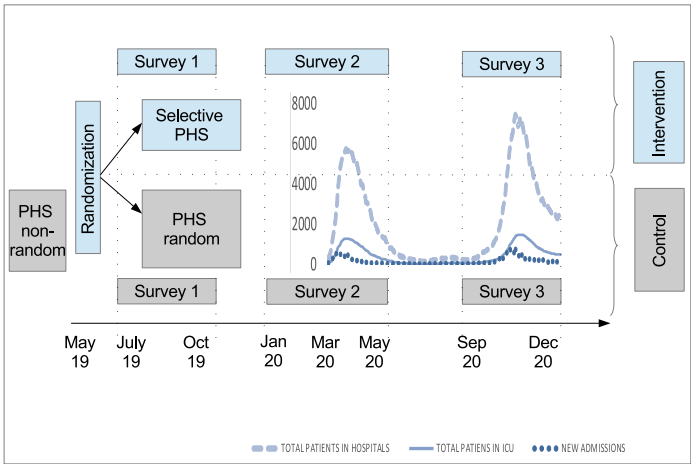


Figure 3.1: Study Design. PHS=periodic health screening. [source COVID-19 data: Dong, Du, and Gardner, 2020]

There are three measurement moments (each time with the same survey, see J. S. Steel et al., 2021a, and in both groups): between June and October 2019 (before COVID-19), between February 2020 and May 2020 (seeing the start of the first COVID wave in Belgium), and between September 2020 and December 2020 (during the second peak in the hospitals). If participants in the intervention group did not complete the first survey, they were referred to the occupational physician for a consultation due to ethical reasons. The design is graphically represented in Figure 3.1, together with the data for COVID-19 hospitalizations in Belgium, and how this relates to the timing of the measurements. Both COVID-19 waves were notoriously large in Belgium, amongst the highest per capita in the world at the time (Ritchie et al., 2020).

In care-as-usual, all employees at risk were invited by the occupational physician to attend a yearly screening. The details of these consultations are available on ClinicalTrials.gov (Identifier NCT04684316). In the intervention group, all employees at risk received a(n) (employee-unique) link to an online health screening questionnaire by email. Dependent upon their answers, 20% of the employees (i.e. the 20% of the employees that mostly needed contact with the

occupational physician, as described in J. S. Steel et al., 2021a) were referred to the occupational physician for a discussion of the results. The consultation then proceeded similar to care-as-usual.

A complication that many preventive interventions have to face, is that health ameliorations and decreases in health care use, absenteeism and presenteeism often only occur after several years. Given the relatively brief timing of our data collection (19 months), we tried to circumvent this complication by not only focusing on long-term outcomes (such as health, healthcare use, productivity), but also incorporating short-term outcomes (such as health literacy, and access behaviour to the occupational physician). The former are expected to improve in the long run, but their short-term behaviour is also relevant (e.g. a rise in healthcare use might point to overconsumption). In addition, because the second and third measurement largely coincide with peaks in COVID-19 hospitalizations, the COVID pandemic has likely increased the healthcare needs of hospital personnel during this period. We believe this is an unforeseen advantage of our study: whereas in normal situations it can take a long time before employees develop complaints, the external health shock could in principle accelerate this process, as employees were temporarily put under extreme strain.

Therefore, we analyse 19 outcome variables over a follow-up period of 19 months: EQ-5D visual analog scale (0-100), nordic musculoskeletal questionnaire (NMQ), stress (0-12), burnout (0-16), sleep problems (0-16), general mental health (GHQ, 0-12), need for recovery at work (NFR, 0-11), absenteeism (days absent), referrals by the occupational physician (0-5), spontaneous consultations with the occupational physician (0-5), work-related consultations with other providers (0-5), health literacy (HLS, 0-100), trust in physician (0-55), job satisfaction (0-16), role conflicts (0-12), turnover intention, use of prescribed medication (no-yes), use of non-prescribed medication (no-yes), and weighted presenteeism (multiplying days of reduced functioning by a functioning weight, as recommended in Bouwmans et al., 2015). For absenteeism in the last 4 weeks, the absence spells in our absenteeism measure above 28 days (4 weeks) were removed from the analyses (n=1684 instead of n=1714). We also assess one adverse effect of screening: worrying, both in intensity (how much do you worry about your health), and frequency (how often do you worry about your health). Finally, we included a weighted score of the two worry-variables.

3.2.2 Statistical analysis

Generalised Linear Mixed Effects models (GLMM) were used to assess the effect of the intervention upon the outcome measures. We discuss their nature in depth in Appendix A.2.1 In our analyses, fixed effects were used on the level of

the employee, and a random intercept and (time) slope effect was introduced to allow each employee to deviate from the (overall) fixed effect.

Since some of the measured outcomes are typically left- or right-skewed, we used logarithmic transformations for musculoskeletal complaints (NMQ), general mental health (GHQ), and need for recovery (NFR). The choice of the covariance structure (and random effects) for each outcome was based on the Akaike criterion.

It is often likely that count variables diverge from the normality assumption of the residuals because of frequent zero-values (e.g. zero days absent, zero consultations, etc.) and dispersed non-zero values (long tails). Therefore, we estimated these outcomes with binomial (B) and Poisson-type (e.g. generalised Poisson (GP) or negative binomial(NB)) mixed models. The fit of these models is tested assessed by several tests for simulated residuals (Hartig & Hartig, 2017). Finally, we make use of continuous ordinal (OC) regression (Manuguerra & Heller, 2010) to analyse the visual analog scale for general health status (EQ-5D), as several authors have argued that it should be analysed as an ordinal variable (Kersten, Küçükdeveci, and Tennant, 2012; Philip, 1990; Svensson, 2000; Wewers and Lowe, 1990 as cited in Manuguerra, Heller, and Ma, 2020).

If the analysis of the quality of the sample indicated significant differences, we controlled for those covariates in the regression analyses, and made use of post-hoc Tukey-Kramer tests (also called honest significant difference tests) as a robustness test of our estimates (Appendix A.2.1). Results were averaged over the levels of the covariates (hospital, gender, education), and a p-value adjustment was used by the Tukey method for comparing a family of 6 estimates (3 times * 2 groups).

In addition, non-inferiority and non-superiority tests of means were performed, with tests assuming a difference of delta (δ) affirms non-inferiority or non-superiority of the intervention mean versus the control mean. We thus verified whether the intervention was non-inferior (when higher outcomes are better) or non-superior (when higher outcomes are worse) in comparison to the control group (Mascha & Sessler, 2011). If the test was non-significant, non-inferiority or non-superiority cannot be concluded. Delta values were based on power calculations, clinical meaningful effects, and recommendations (Mascha & Sessler, 2011; Walker & Nowacki, 2011).

3.3 Results

3.3.1 Sample

In the four hospitals, 1077 unique employees participated in one or more of the three survey rounds and completely filled out the survey : 516 had been allocated to the intervention group (441 random, 75 non-random), and 561 to the control group (81 random, 480 non-random). This means the average overall participation rate was 34% (1077 in 3150). Appendix A.2.2 shows the participation over the three survey rounds, and descriptive statistics by wave. It also includes a column for the 208 respondents (6.6%) that completed all three survey waves, and presents health status information (a complete table can be found in the Appendix Table A.6).

The 1077 unique employees gave rise to 1782 response records over all three measurement moments: 906 responses were from the control group, 876 from the intervention group. Of the 876 responses in the intervention group, 126 were referred by the algorithm (the top 20% of the scores), 684 were not referred (524 because they were in the bottom 80% of the scores, 160 because they missed the first measurement), and 66 were not referred but still opted for a face-to-face consultation. The 160 employees in the intervention group that missed the first measurement were invited for a face-to-face consultation with the physician.

We assessed the quality of our sample in two ways. First, we evaluated the randomisation process, by verifying whether the control and intervention group differed substantially on background characteristics. The tables in Appendix Table A.2.3 show descriptive statistics by group for waves 1, 2, and 3, which demonstrate small differences for education in wave 1 (not wave 2 or 3). Second, we compared our sample with aggregated population data from hospitals B,C and D (there was no data available for hospital A), the results are in Appendix A.10. In light of these results, we controlled all regressions for age, gender, and education, but not occupation (since this is highly correlated to education), and made use of post-hoc tests.

3.3.2 Generalised Mixed Models for final and intermediate outcomes

The estimates from the linear mixed regressions are shown in tables 3.2 to 3.4, full models are available in tables A.11 to A.15, and show the average fixed effects for group and time, and the interaction effects. Although this is not the primary focus of this article, the time fixed effects demonstrate the impact of COVID-19 on personnel over time, for instance through decreased health,

increased musculoskeletal complaints, or increased stress and burnout risk in wave 2 and 3 compared to wave 1. Note how this makes the first measurement (pre-COVID) fundamentally different from the second and third measurements.

The results indicate no significant overall differences between intervention and control group for health status (EQ-5D vas), musculoskeletal problems, stress, burnout, sleep problems, general mental health, need for recovery, health literacy, job satisfaction, turnover intention, weighted presenteeism, worrying about health, and non-prescribed medication. The intervention group had, on average, lower scores for trust in the physician (-0.94 on a scale of 55), and had more role conflicts (0.30 on a scale of 12). The intervention group was also, on average, fewer days absent from work (IRR=0.30), had fewer work-related consultations (OR=0.69), had fewer referrals by the occupational physician (IRR=0.46), had fewer spontaneous consultations (IRR=0.35), and had more prescribed medication use (OR=1.53). However, these were mainly the result of differences in the first measurement.

The interactions at time 2 indicate an increase in the difference of intervention and control group for musculoskeletal problems, health literacy, worry intensity, the weighted worry-score, and (because turnover intention in the control group increased) a decrease in the difference of turnover intention, compared to the difference of intervention and control group in time 1.

On the final measurement moment, burnout risk is slightly reduced for the intervention group, general mental health is reduced, and need for recovery is reduced. The between group differences of absenteeism and work-related and spontaneous consultations decreased (the intervention group rose in comparison to the control group in time 1, making the difference smaller). The interactions thus indicate a short-term difference between groups at time 1 that diminished over time. We judge the absence of effects in time 2 and 3 as more important than the presence of group differences in time 1, as the occupational burdens to which participants were exposed were higher during COVID-19 (time 2 and 3).

3.3.3 Post-hoc tests

The tables 3.5 and 3.6 and figs. 3.2 and 3.3 summarise the post-hoc estimations. For the superiority post-hoc differences, only absenteeism, referrals, and spontaneous consultations are significantly different between groups (first three columns of Table 3.5). In the non-inferiority tests, the majority of the estimates indicate the intervention is at least as good as care-as-usual (EQ-5D, NMQ, stress, GHQ, NFR, HLS, trust, job satisfaction, absenteeism, healthcare use, and presenteeism). If the null hypothesis cannot be rejected, this does not indicate that the reverse (inferiority of the intervention group) is true.

Table 3.2: Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals

Dependent Var Range	EQ-5D vas (0-100)	Musculo (log) (1-10)	Stress (0-12)	Burnout (0-16)	Sleep Problems (0-16)	GHQ(log) (1-13)
Intervention group	-0.26	-0.06	0.02	0.13	0.36	0.07
(ref.=control)	(-2.01 – 1.49)	(-0.15 – 0.02)	(-0.32 – 0.36)	(-0.36 – 0.63)	(-0.14 – 0.86)	(-0.04 – 0.19)
Time 2 (ref.=Time1)	-1.72 *	-0.12 ***	0.46 ***	0.59 ***	0.07	0.08
	(-3.50 – 0.05)	(-0.20 – -0.04)	(0.16 – 0.77)	(0.15 – 1.03)	(-0.34 – 0.48)	(-0.04 – 0.19)
Time 3	-1.11	0.45 ***	0.48 ***	0.69 ***	0.53 **	0.17 ***
	(-2.80 – 0.57)	(0.36 – 0.54)	(0.19 – 0.77)	(0.25 – 1.13)	(0.08 – 0.98)	(0.05 – 0.28)
intervention:time2	-0.46	0.13 **	-0.29	-0.25	-0.27	-0.10
	(-2.94 – 2.01)	(0.02 – 0.24)	(-0.72 – 0.14)	(-0.87 – 0.36)	(-0.84 – 0.30)	(-0.26 – 0.06)
intervention:time3	0.00	0.01	-0.21	-0.57 *	-0.53	-0.15 *
	(-2.41 – 2.41)	(-0.12 – 0.14)	(-0.62 – 0.20)	(-1.20 – 0.06)	(-1.18 – 0.11)	(-0.32 – 0.01)
Random Effects						
σ^2	13.75	0.03	0.43	0.94	0.91	0.06
τ_{00}	144.97 ID	0.36 ID	5.73 ID	12.10 ID	12.30 ID	0.60 ID
τ_{11}	134.63 ID.time2	0.28 ID.time2	4.23 ID.time2	8.43 ID.time2	6.75 ID.time2	0.64 ID.time2
	173.98 ID.time3	0.50 ID.time3	4.66 ID.time3	11.21 ID.time3	11.83 ID.time3	0.90 ID.time3
ρ_{01}	-0.33	-0.45	-0.44	-0.43	-0.39	-0.5
	-0.51	-0.47	-0.49	-0.47	-0.46	-0.57
ICC	0.92	0.92	0.93	0.93	0.93	0.91
N	1041 ID	1042 ID	1042 ID	1042 ID	1042 ID	1042 ID
Observations	1733	1736	1736	1736	1736	1736
Marginal R2 / Conditional R2	0.029 / 0.923	0.132 / 0.932	0.048 / 0.932	0.046 / 0.931	0.025 / 0.934	0.037 / 0.916

Note: *p<0.1; **p<0.05; ***p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept ; τ_{11} = variance of random slope; ρ_{01} = correlation between random intercept and slope

Dependent Var Range	NFR(log) (1-12)	Role conflicts (0-12)	Health literacy (0-100)	Trust in physician (0-55)	Job Satisfaction (0-20)	Turnover intention (0-4)
Intervention group	0.00	0.30 *	-0.97	-0.94 **	-0.07	-0.02
(ref.=control)	(-0.09 – 0.09)	(-0.04 – 0.64)	(-2.43 – 0.49)	(-1.78 – -0.10)	(-0.38 – 0.25)	(-0.17 – 0.13)
Time 2 (ref.=Time1)	0.04	0.16	-1.31 *	-0.32	-0.09	0.19 ***
	(-0.03 – 0.12)	(-0.13 – 0.46)	(-2.62 – -0.00)	(-0.99 – 0.36)	(-0.35 – 0.17)	(0.05 – 0.32)
Time 3	0.08 **	0.13	1.22 *	-0.04	-0.06	-0.15 **
	(0.00 – 0.15)	(-0.17 – 0.43)	(-0.19 – 2.63)	(-0.79 – 0.72)	(-0.33 – 0.21)	(-0.28 – -0.02)
intervention:time2	-0.08	-0.27	1.55 *	0.32	0.08	-0.18 *
	(-0.18 – 0.03)	(-0.68 – 0.14)	(-0.27 – 3.38)	(-0.61 – 1.25)	(-0.28 – 0.44)	(-0.37 – 0.00)
intervention:time3	-0.09 *	-0.22	-0.36	0.76	0.05	0.15
	(-0.20 – 0.02)	(-0.65 – 0.22)	(-2.38 – 1.65)	(-0.32 – 1.85)	(-0.33 – 0.43)	(-0.04 – 0.34)
Random Effects						
σ^2	0.03	0.44	8.57	2.39	0.36	0.09
τ_{00}	0.37 ID	5.70 ID	99.95 ID	33.41 ID	5.13 ID	1.11 ID
τ_{11}	0.23 ID.time2	3.60 ID.time2	72.66 ID.time2	15.86 ID.time2	2.48 ID.time2	0.70 ID.time2
	0.31 ID.time3	5.22 ID.time3	113.80 ID.time3	31.01 ID.time3	3.86 ID.time3	0.95 ID.time3
ρ_{01}	-0.45	-0.43	-0.49	-0.36	-0.28	-0.29
	-0.44	-0.46	-0.45	-0.45	-0.46	-0.45
ICC	0.93	0.93	0.92	0.93	0.93	0.93
N	1042 ID	1042 ID	1017 ID	1001 ID	1040 ID	1034 ID
Observations	1736	1736	1688	1642	1734	1723
Marginal R2 / Conditional R2	0.023 / 0.930	0.058 / 0.932	0.027 / 0.926	0.094 / 0.940	0.024 / 0.936	0.098 / 0.938

Note: *p<0.1; **p<0.05; ***p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept ; τ_{11} = variance of random slope; ρ_{01} = correlation between random intercept and slope

Table 3.3: Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals.

Dependent var.	OC EQ5D vas	GP Weighted Presenteeism	GP absenteeism	GP absenteeism last 4 weeks	B work-related consultations	NB referrals	NB spontaneous consultations
Range	(0-100)	(0-140)	(0-526)	(0-28)	(no-yes)	(0-5)	(0-5)
Estimates	OC	IRR	IRR	IRR	OR	IRR	IRR
Intervention group (ref.=control)	0.16 (0.15– 0.17)	1.15 (0.95 – 1.39)	0.59 ** (0.37 – 0.95)	0.30 *** (0.14 – 0.63)	0.69 ** (0.47 – 1.00)	0.46 *** (0.28 – 0.76)	0.35 *** (0.23 – 0.55)
Time 2 (ref.=Time1)	0.3 (0.29– 0.31)	0.98 (0.77 – 1.25)	1.45 (0.91 – 2.29)	1.36 (0.74 – 2.51)	0.80 (0.52 – 1.23)	0.58 * (0.33 – 1.00)	0.54 *** (0.34 – 0.86)
Time 3	0.26 (0.25– 0.26)	1.07 (0.87 – 1.31)	0.83 (0.52 – 1.33)	1.38 (0.74 – 2.59)	0.62 ** (0.43 – 0.92)	0.48 *** (0.29 – 0.80)	0.59 *** (0.40 – 0.87)
intervention:time2	-0.05 (-0.06– -0.04)	0.97 (0.69 – 1.35)	1.38 (0.68 – 2.79)	4.11 *** (1.69 – 10.01)	1.20 (0.66 – 2.18)	1.16 (0.46 – 2.88)	2.51 ** (1.20 – 5.27)
intervention:time3	-0.06 (-0.07– -0.05)	0.88 (0.66 – 1.18)	1.91 * (0.94 – 3.85)	2.96 ** (1.11 – 7.89)	1.73 * (1.00 – 2.99)	1.78 (0.80 – 3.97)	2.47 *** (1.29 – 4.75)
Zero-Inflated Model							
Intercept		see Appendix		7.19 ***		0.24	
Dispersion Model							
Overdispersion parameter		see Appendix	521	7.19		3.5	1.11
Random Effects							
σ^2	3.144				3.29		
τ_{00}	1.765 ID				1.37 ID		
ICC					0.29		
N					1042 ID		
Observations	1037	1638	1714	1684	1735	1734	1736
R2 conditional / R2 marginal		/ 0.026	/ 0.410	/ 0.344	0.028 / 0.313	/ 0.077	/ 0.084

Note: *p<0.1; **p<0.05; *** p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept
OC = ordinal continuous; GP = Generalised Poisson; B=Binomial; NB=Negative Binomial; IRR = Incidence Rate Ratios; OR = Odds Ratios

Table 3.4: Regressions mixed methods: group, time, and interaction effect, controlled for age, gender, education, and hospital with 95% confidence intervals.

Dependent var.	B prescribed medication	B non-prescribed medication	Worry intensity	Worry frequency	Worry weighted score
Range	(no-yes)	(no-yes)	(1-5)	(1-5)	(0-1.55)
Estimates	*Odds Ratios*	*Odds Ratios*	*linear*	*linear*	*linear*
Intervention group	1.53 *	1.23	-0.06	-0.03	-0.02
(ref.=control)	(0.93 – 2.50)	(0.83 – 1.82)	(-0.17 – 0.06)	(-0.18 – 0.13)	(-0.07 – 0.03)
Time 2	1.30	1.51 *	-0.14 **	0.01	-0.03
(ref.=Time1)	(0.78 – 2.17)	(0.97 – 2.34)	(-0.24 – -0.03)	(-0.12 – 0.15)	(-0.08 – 0.01)
Time 3	0.88	1.32	-0.06	0.24 ***	0.03
	(0.56 – 1.40)	(0.89 – 1.96)	(-0.16 – 0.04)	(0.11 – 0.38)	(-0.01 – 0.08)
intervention:time2	0.56	0.97	0.20 ***	0.10	0.07 **
	(0.28 – 1.16)	(0.53 – 1.79)	(0.05 – 0.35)	(-0.08 – 0.29)	(0.01 – 0.13)
intervention:time3	0.95	1.27	0.03	-0.07	-0.00
	(0.49 – 1.83)	(0.73 – 2.21)	(-0.11 – 0.18)	(-0.27 – 0.12)	(-0.07 – 0.06)
Random Effects					
σ^2	3.29	3.29	0.05	0.09	0.01
τ_{00}	4.30 ID	1.44 ID	0.62 ID	1.15 ID	0.14 ID
τ_{11}			0.46 ID.time2	0.72 ID.time2	0.08 ID.time2
			0.61 ID.time3	1.07 ID.time3	0.12 ID.time3
ρ_{01}			-0.42	-0.39	-0.36
			-0.57	-0.48	-0.5
ICC	0.57	0.31	0.92	0.93	0.93
N	1042 ID	1042 ID	1042 ID	1042 ID	1042 ID
Observations	1734	1734	1736	1735	1735
R2 conditional / R2 marginal	0.073 / 0.598	0.046 / 0.337	0.037 / 0.922	0.030 / 0.930	0.031 / 0.931

Note: *p<0.1; **p<0.05; ***p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept
B=Binomial; OR = Odds Ratios

Table 3.5: Posthoc estimations with contrasts of within (one group over time) and between (same time different groups) group with 95% confidence intervals

contrasts	c1-i1	c2-i2	c3-i3	c1-c2	c2-c3	i1-i2	i2-i3
EQ-5D vas	0.261	0.723	0.26	1.724	-0.609	2.186	-1.072
	(-2.297 – 2.82)	(-2.907 – 4.354)	(-2.801 – 3.322)	(-0.874 – 4.321)	(-3.48 – 2.262)	(-0.345 – 4.717)	(-3.875 – 1.731)
NMQ	0.065	-0.066	0.056	0.115 *	-0.566	-0.016	-0.443 ***
	(-0.063 – 0.192)	(-0.226 – 0.093)	(-0.11 – 0.222)	(-0.002 – 0.233)	(-0.705 – -0.426)	(-0.131 – 0.099)	(-0.581 – -0.306)
Stress	-0.02	0.27	0.19	-0.462 **	-0.02	-0.172	-0.1
	(-0.517 – 0.476)	(-0.344 – 0.884)	(-0.351 – 0.732)	(-0.912 – -0.013)	(-0.486 – 0.447)	(-0.611 – 0.268)	(-0.553 – 0.353)
Burnout	-0.134	0.121	0.436	-0.586 *	-0.102	-0.332	0.213
	(-0.858 – 0.591)	(-0.774 – 1.016)	(-0.397 – 1.268)	(-1.232 – 0.059)	(-0.803 – 0.599)	(-0.964 – 0.299)	(-0.471 – 0.897)
Sleep Problems	-0.361	-0.092	0.173	-0.07	-0.455	0.199	-0.191
	(-1.088 – 0.366)	(-0.964 – 0.78)	(-0.681 – 1.026)	(-0.669 – 0.529)	(-1.151 – 0.24)	(-0.387 – 0.785)	(-0.873 – 0.491)
GHQ	-0.075	0.028	0.076	-0.078	-0.091	0.024	-0.042
	(-0.241 – 0.092)	(-0.192 – 0.248)	(-0.124 – 0.277)	(-0.248 – 0.092)	(-0.275 – 0.093)	(-0.143 – 0.191)	(-0.223 – 0.138)
NFR	-0.001	0.077	0.091	-0.041	-0.036	0.036	-0.022
	(-0.126 – 0.125)	(-0.074 – 0.227)	(-0.054 – 0.236)	(-0.15 – 0.067)	(-0.159 – 0.087)	(-0.07 – 0.141)	(-0.143 – 0.099)
HLS	0.97	-0.582	1.333	1.314	-2.532 **	-0.239	-0.617
	(-1.167 – 3.107)	(-3.152 – 1.987)	(-1.321 – 3.987)	(-0.605 – 3.232)	(-4.769 – -0.295)	(-2.109 – 1.631)	(-2.825 – 1.592)
Trust	0.936	0.617	0.172	0.318	-0.282	-0.001	-0.727
	(-0.293 – 2.166)	(-0.84 – 2.075)	(-1.265 – 1.609)	(-0.668 – 1.305)	(-1.469 – 0.906)	(-0.95 – 0.949)	(-1.9 – 0.446)
Job Satisfaction	0.065	-0.015	0.013	0.086	-0.023	0.005	0.005
	(-0.399 – 0.53)	(-0.603 – 0.573)	(-0.507 – 0.532)	(-0.294 – 0.465)	(-0.48 – 0.435)	(-0.364 – 0.375)	(-0.444 – 0.453)
Role conflicts	-0.299	-0.032	-0.082	-0.162	0.033	0.105	-0.018
	(-0.796 – 0.197)	(-0.637 – 0.572)	(-0.658 – 0.493)	(-0.593 – 0.268)	(-0.459 – 0.524)	(-0.316 – 0.525)	(-0.499 – 0.464)
Turnover intention	0.023	0.206	-0.126	-0.187 *	0.335 ***	-0.003	0.003
	(-0.196 – 0.242)	(-0.083 – 0.496)	(-0.378 – 0.126)	(-0.382 – 0.008)	(0.115 – 0.555)	(-0.194 – 0.187)	(-0.213 – 0.218)
Absenteeism GP	0.527	0.206	-0.117	-0.37	0.556	-0.691	0.232
	(-0.164 – 1.219)	(-0.551 – 0.964)	(-0.874 – 0.64)	(-1.038 – 0.299)	(-0.194 – 1.306)	(-1.469 – 0.088)	(-0.532 – 0.997)
Absenteeism last 4 weeks GP	1.207 **	-0.206	0.122	-0.309	-0.016	-1.723 ***	0.312
	(0.13 – 2.283)	(-0.895 – 0.482)	(-0.79 – 1.034)	(-1.198 – 0.58)	(-0.846 – 0.814)	(-2.759 – -0.686)	(-0.451 – 1.075)
Presenteeism weighted GP	-0.138	-0.103	-0.011	0.019	-0.086	0.053	0.007
	(-0.416 – 0.14)	(-0.505 – 0.298)	(-0.337 – 0.315)	(-0.338 – 0.375)	(-0.455 – 0.284)	(-0.284 – 0.39)	(-0.359 – 0.372)
HC Work-related B	0.375	0.193	-0.172	0.222	0.248	0.04	-0.117

	(-0.17 – 0.92)	(-0.554 – 0.94)	(-0.812 – 0.468)	(-0.397 – 0.842)	(-0.418 – 0.914)	(-0.573 – 0.653)	(-0.774 – 0.54)
HC Referrals NB	0.772 **	0.626	0.195	0.547	0.181	0.401	-0.25
	(0.057 – 1.488)	(-0.535 – 1.787)	(-0.738 – 1.127)	(-0.256 – 1.35)	(-0.752 – 1.115)	(-0.661 – 1.462)	(-1.402 – 0.902)
HC Spontaneous NB	1.038 ***	0.117	0.132	0.62 *	-0.086	-0.3	-0.071
	(0.411 – 1.664)	(-0.763 – 0.998)	(-0.586 – 0.849)	(-0.065 – 1.305)	(-0.856 – 0.685)	(-1.135 – 0.535)	(-0.906 – 0.764)
Prescribed Medication B	-0.424	0.147	-0.369	-0.261	0.386	0.31	-0.13
	(-1.141 – 0.294)	(-0.8 – 1.095)	(-1.193 – 0.456)	(-1.006 – 0.485)	(-0.414 – 1.186)	(-0.423 – 1.044)	(-0.914 – 0.654)
Non-Prescribed Medication B	-0.206	-0.18	-0.443	-0.41	0.131	-0.384	-0.132
	(-0.774 – 0.361)	(-0.934 – 0.575)	(-1.088 – 0.202)	(-1.05 – 0.229)	(-0.541 – 0.803)	(-1.004 – 0.237)	(-0.788 – 0.525)

Note: c=control; i=intervention; 1,2,3=time1,2,3; GP = Generalised Poisson; B=Binomial; NB=Negative Binomial

Table 3.6: Posthoc estimations with contrasts of within (one group over time) and between (same time different groups). Non-inferiority tests

contrasts	delta	direction	i1-c1	i2-c2	i3-c3	i2-i1	i3-i2	c2-c1	c3-c2
EQ-5D VAS	5	noninferior	-0.26 **	-0.72 **	-0.26 **	-2.19 **	1.072 **	-1.72 **	0.609 **
			0	0.006	0.0001	0.012	0	0.003	0
NMQ	log(0.45)	nonsuperior	-0.07 **	0.066 **	-0.06 **	0.016 **	0.443 **	-0.12 **	0.566 **
			0	0	0	0	0	0	0.00002
Stress	0.6	nonsuperior	0.02 **	-0.27 **	-0.19 **	0.172 *	0.1 **	0.462	0.02 **
			0.007	0.0004	0.0003	0.041	0.013	0.959	0.003
Burnout	0.8	nonsuperior	0.134	-0.121 *	-0.44 **	0.332	-0.21 **	0.586	0.102 *
			0.064	0.025	0.0002	0.23	0.0002	0.941	0.034
Sleep Problems	0.8	nonsuperior	0.361	0.092	-0.17 **	-0.2 **	0.191	0.07 **	0.455
			0.478	0.144	0.009	0.00001	0.079	0.004	0.707
GHQ	log(0.6)	nonsuperior	0.075 **	-0.03 **	-0.08 **	-0.02 **	0.042 **	0.078 **	0.091 **
			0	0	0	0	0	0	0
NFR	log(0.6)	nonsuperior	0.001 **	-0.08 **	-0.09 **	-0.04 **	0.022 **	0.041 **	0.036 **
			0	0	0	0	0	0	0
HLS	5	noninferior	-0.97 **	0.582 **	-1.33 **	0.239 **	0.617 **	-1.31 **	2.532 **
			0	0.00001	0	0	0	0	0.013
Trust	2.75	noninferior	-0.94 **	-0.62 **	-0.17 **	0.001 **	0.727 **	-0.32 **	0.282 **
			0.0002	0.0002	0	0	0	0	0
Job Satisfaction	1	noninferior	-0.07 **	0.015 **	-0.01 **	-0.01 **	-0.01 **	-0.09 **	0.023 **

			0	0.00001	0	0	0	0	0
Role Conflicts	0.6	nonsuperior	0.299	0.032 *	0.082	-0.11 **	0.018 **	0.162 *	-0.03 **
			0.474	0.054	0.075	0.00002	0.004	0.028	0.002
Turnover Intention	0.2	nonsuperior	-0.023 *	-0.21 **	0.126	0.003 *	-0.003 *	0.187	-0.34 **
			0.028	0.0005	0.965	0.025	0.053	1	0
Absenteeism GP	1	nonsuperior	-0.53 **	-0.21 **	0.117 **	0.691	-0.23 **	0.37 *	-0.56 **
			0	0.00004	0.007	0.873	0.00003	0.053	0
Absenteeism last 4 weeks GP	1	nonsuperior	-1.21 **	0.206 **	-0.12 **	1.723	-0.31 **	0.309	0.016 **
			0	0.008	0.003	1	0.00001	0.183	0.006
Presenteeism weighted GP	0.5	nonsuperior	0.138 **	0.103 *	0.011 **	-0.05 **	-0.01 **	-0.02 **	0.086 **
			0.002	0.036	0.0001	0.00002	0.001	0.0003	0.011
HC Work-related B	1	nonsuperior	-0.38 **	-0.19 **	0.172 **	-0.04 **	0.117 **	-0.22 **	-0.25 **
			0	0.00004	0.002	0.00001	0.001	0	0
HC Referrals NB	1	nonsuperior	-0.77 **	-0.63 **	-0.2 **	-0.4 **	0.25	-0.55 **	-0.18 **
			0	0.001	0.002	0.001	0.383	0	0.002
HC Spontaneous NB	1	nonsuperior	-1.04 **	-0.12 **	-0.13 **	0.3	0.071 **	-0.62 **	0.086 **
			0	0.002	0.0001	0.12	0.011	0	0.005
Prescribed Medication B	1	nonsuperior	0.424	-0.15 **	0.369	-0.31 **	0.13 **	0.261 *	-0.39 **
			0.154	0.004	0.197	0	0.012	0.035	0.00001
Non-Prescribed Medication B	1	nonsuperior	0.206 **	0.18 *	0.443	0.384 *	0.132 **	0.41	-0.13 **
			0.001	0.015	0.099	0.034	0.001	0.063	0.00001

Note: c=control; i=intervention; 1,2,3=time1,2,3; GP = Generalised Poisson; B=Binomial; NB=Negative Binomial

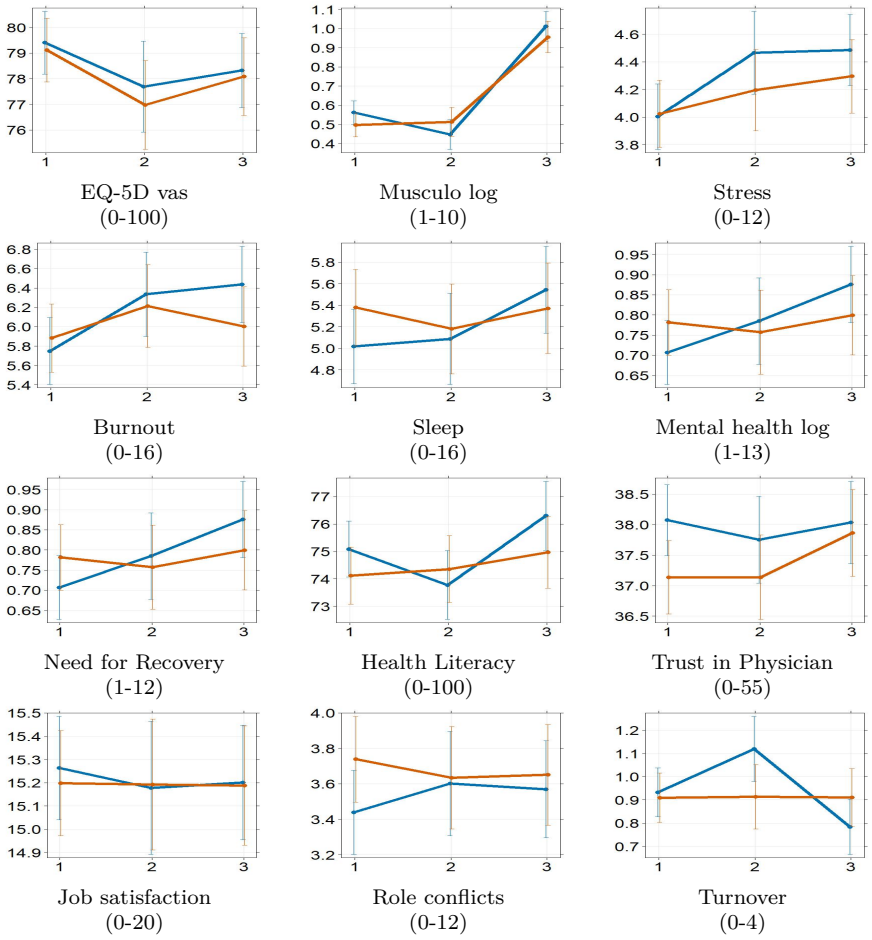


Figure 3.2: Predicted effect plots of mixed models with 95% confidence intervals, blue=control group, orange=intervention group.
x-axis=measurement moment (1,2,3)

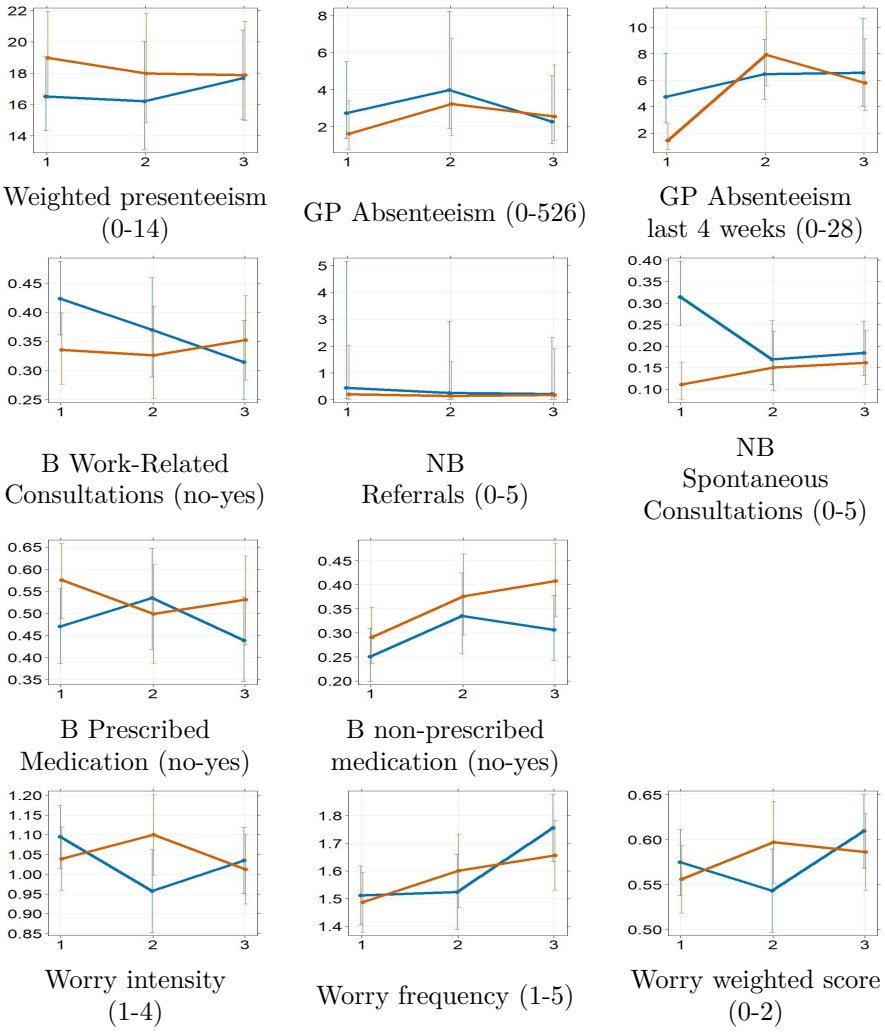


Figure 3.3: Predicted effect plots of mixed models with 95% confidence intervals, blue=control group, orange=intervention group. x-axis=measurement moment (1,2,3); GP = Generalised Poisson; B=Binomial; NB=Negative Binomial

3.4 Discussion

Non-inferiority tests indicate the intervention group performs at least as good as the control group on the majority of the outcomes. The regressions and (superiority) post-hoc analyses indicate that there were some short-term effects of the intervention: trust in the physician, absenteeism and healthcare use (referrals, spontaneous consultations, and work-related consultations) were lower in the intervention group. However, these differences predominantly occurred in wave 1 and were no longer present in the third measurement moment when participants had been exposed to exceptional circumstances.

In a way, the absence of large differences between the control and the intervention could have been expected for the long-term (e.g. health) outcomes: it often takes a long time before the beneficial effects of screening become apparent, and our study duration (19 months) might have been too short to pick this up. Although the COVID-19 pandemic had a clear impact on the data, for instance visible in the increased stress, burnout risk, and musculoskeletal functioning problems over time, it did not increase the differences between the face-to-face screenings of the control group compared to the targeted screenings. While outcomes changed over time, differences between groups remained largely absent.

The differences we did find (absenteeism, trust, and healthcare use) were mainly driven by differences in the first measurement moment. Their direction seems plausible: after the control group had a face-to-face screening, it seems natural that their trust in their physician is increased. Likewise, if health or functioning problems are discussed in the screening, this might lead to additional consultations (hence increased healthcare use) or increased absenteeism (e.g. to consult specialists or by taking up advice by the physician to stay home). The rise in healthcare use is consistent with other literature, where a higher healthcare use is one of the short-term consequences of face-to-face screening, but the rise in absenteeism is not (Hackl et al., 2015). The fact that these effects are no longer present in the second and third wave might be explained by several factors. COVID-19 might have urged personnel (from control and intervention group alike) to refrain from taking absence leave because of the increased pressure on the hospitals, or from visiting physicians because of contact restrictions. In addition, it is also possible that the effects from face-to-face contact were only temporary: an immediate increase in healthcare use, absenteeism, and trust that faded away by the time of the second and third measurement. Third, the non-perfect randomisation might have played a role, although the robustness checks of the quality of our sample (both vis-à-vis the population and based on background differences between groups) did not indicate large problems in this respect.

Some limitations and potential biases should also be pointed out. Although we verified the quality of our sample, the randomisation process was only partial and participation could have been affected by selection bias. Similarly, our results are dependent on the four selected hospitals, although care was taken to select large hospitals from different regions. The results are also limited to the short time-frame of the study (19 months), and to the use of self-assessed outcomes. Because self-assessed outcomes were used, in contrast to diagnostic information from occupational physicians, this study focused on the effectiveness of the intervention, not on its ability to detect health problems (e.g. burnouts). The design and performance of the survey and algorithm was however taken up in a previous article (J. S. Steel et al., 2021a). While COVID-19 might have increased the burden on personnel and gives an indication of long-term effects, the true long-term effects should be measured in further research. It is still possible that, as in other screening studies, the benefits of face-to-face screening only become apparent in the long run. Nevertheless, the short-term impact of screening is far from irrelevant: turnover is traditionally high in the hospital sector (causing short employment periods), intermediate outcomes give an indication of long-term effects, and short-term changes can strongly affect cost-effectiveness (e.g. a higher healthcare use in the short term without long-term improvements can point to overuse of care). Finally, it is possible that the intervention (leaving out a face-to-face screening with the physician) had little effect because the employee and physician already had a long-standing bond, built up from past screenings and workplace visits. As shown in Appendix A.2.2 about half the sample knew the occupational physician for five years or longer. A robustness analysis where this relation was added as a fixed covariate effect did not show substantial differences.

For policy makers and practitioners, we want to emphasise that we did not compare screening by the physician with no screening, but with more targeted screening. This means we cannot make claims on the effectiveness of screening by the physician, but can only compare it to the intervention: screening after an additional risk selection. We refer to other studies, that did assess the effect of screening versus not screening, e.g. general screenings (Krogsbøll et al., 2013; Wynn, 2013), vision screening of older drivers (Desapriya et al., 2014), tuberculosis screening of healthcare workers (Zwerling et al., 2012), tomography for lung cancer (Black et al., 2006), pre-employment investigations (Schaafsma, Mahmud, Reneman, Fassier, & Jungbauer, 2016), or biometric screening (Arena et al., 2014; J. Baker, Mitchell, Lawson, & Pell, 2013; Behling et al., 2013; Burden & Rouse, 2010; Lawson, Fenwick, Pell, & Pell, 2010; Robbins et al., 2013; Selvarajah et al., 2013; Soler et al., 2010; Task Force on Community Preventive Services, 2005). In addition, we wish to highlight that occupational health services and periodic health screenings can have other purposes, as they can be used for primary, secondary, and tertiary prevention.

Nevertheless, for those populations where physical consultations are not strictly necessary, a lower frequency of face-to-face health screening might prove at least as effective if combined with a more targeted approach to ensure those who really need it are still offered adequate care.

References

All Party Parliamentary Group on Occupational Safety and Health. (2016). Occupational medical workforce crisis. the need for action to keep the uk workforce healthy. All Party Parliamentary Group on Occupational Safety and Health London.

Arena, R., Arnett, D. K., Terry, P. E., Li, S., Isaac, F., Mosca, L., ... Whitsel, L. P. (2014). The role of worksite health screening: A policy statement from the american heart association. *Circulation*, *130*(8), 719–34. doi:10.1161/CIR.0000000000000079

Baker, J., Mitchell, R., Lawson, K., & Pell, J. (2013). Ethnic differences in the cost-effectiveness of targeted and mass screening for high cardiovascular risk in the uk: Cross-sectional study. *Heart*, *99*(23), 1766–71. doi:10.1136/heartjnl-2013-304625

Behling, C., Kelly, R., Ruppert, J., Whitsel, L., Hudson, M., Terry, P., ... Loeppke, R. (2013). Biometric health screening for employers consensus statement of the health enhancement research organization, american college of occupational and environmental medicine, and care continuum alliance. *J. Occup. Environ. Med.*, *55*(10), 1244–1251.

Black, C., Bagust, A., Boland, A., Walker, S., McLeod, C., De Verteuil, R., ... Waugh, N. (2006). The clinical effectiveness and cost-effectiveness of computed tomography screening for lung cancer: Systematic reviews. *Health Technol Assess*, *10*(3), iii–iv, ix–x, 1–90.

Bouwmans, C., Krol, M., Severens, H., Koopmanschap, M., Brouwer, W., & Hakkaart-van Roijen, L. (2015). The imta productivity cost questionnaire: A standardized instrument for measuring and valuing health-related productivity losses. *Value in health*, *18*(6), 753–758.

Burden, A. C., & Rouse, A. (2010). Cardiovascular prevention. go for targeted screening. *BMJ*, *340*, c2866. doi:10.1136/bmj.c2866

Centers for Disease Control and Prevention (CDC). (2018). Workplace health in america 2017. Atlanta, GA. Centers for Disease Control, Prevention, U.S. Department of Health, and Human Services.

Cherry, N., Meyer, J., Holt, D., Chen, Y., & McDonald, J. (2000). Surveillance of work-related diseases by occupational physicians in the uk: Opra 1996–1999. *Occupational Medicine*, *50*(7), 496–503.

Colosio, C., Mandic-Rajcevic, S., Godderis, L., van der Laan, G., Hulshof, C., & van Dijk, F. (2017). Workers' health surveillance: Implementation of the directive 89/391/eec in europe. *Occupational Medicine*, *67*(7), 574–578.

Desapriya, E., Harjee, R., Brubacher, J., Chan, H., Hewapathirane, D. S., Subzwari, S., & Pike, I. (2014). Vision screening of older drivers for preventing road traffic injuries and fatalities. *Cochrane Database Syst Rev*, (2), CD006252. doi:10.1002/14651858.CD006252.pub4

Désiron, H. A., Donceel, P., Godderis, L., Van Hoof, E., & de Rijk, A. (2015). What is the value of occupational therapy in return to work for breast cancer patients? a qualitative inquiry among experts. *European journal of cancer care*, *24*(2), 267–280.

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533–534. doi:10.1016/S1473-3099(20)30120-1

EU. (1989b). Council directive 89/391/eec of 12 june 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work. *Official Journal of the European Communities*, *183*, 1–8.

EU. (2017). Commission staff working document, ex-post evaluation of the european union occupational safety and health directives (refit evaluation). *Official Journal of the European Communities*, *SWD 10*, 21.

Hackl, F., Halla, M., Hummer, M., & Pruckner, G. J. (2015). The effectiveness of health screening. *Health Econ*, *24*(8), 913–35. doi:10.1002/hec.3072

Hartig, F., & Hartig, M. F. (2017). Package 'dharma'.

Herman, A. M., & Jeffress, C. N. (2000). *Screening and surveillance: A guide to osha standards*. Department of labor Washington DC occupational safety and health administration.

International Labour Office. (1998). *Technical and ethical guidelines for workers' health surveillance (osh no. 72)*. International Labour Office (ILO).

Kersten, P., Küçükdeveci, A. A., & Tennant, A. (2012). The use of the visual analogue scale (vas) in rehabilitation outcomes. *Journal of rehabilitation medicine*, *44*(7), 609–610.

Krogsbøll, L. T., Jørgensen, K., & Gøtzsche, P. C. (2013). General health checks in adults for reducing morbidity and mortality from disease. *JAMA*, *309*(23), 2489–2490. doi:10.1001/jama.2013.5039

Lambrechts, M.-C., Ketterer, F., Symons, L., Mairiaux, P., Peremans, L., Remmen, R., ... Godderis, L. (2015). The approach taken to substance abuse by occupational physicians: A qualitative study on influencing factors. *Journal of occupational and environmental medicine*, *57*(11), 1228–1235.

Lawson, K. D., Fenwick, E. A., Pell, A. C., & Pell, J. P. (2010). Comparison of mass and targeted screening strategies for cardiovascular risk: Simulation

of the effectiveness, cost-effectiveness and coverage using a cross-sectional survey of 3921 people. *Heart*, *96*(3), 208–212.

MacAuley, D. (2012). The value of conducting periodic health checks. British Medical Journal Publishing Group.

Mairiaux, P., Chanteux, A., Schleich, E., & Delwick, B. (2010). *Financement des missions de prévention des services externes de prévention et de protection au travail*. SPF Emploi, travail et concertation sociale.

Manuguerra, M., & Heller, G. Z. (2010). Ordinal regression models for continuous scales. *The International Journal of Biostatistics*, *6*(1).

Manuguerra, M., Heller, G. Z., & Ma, J. (2020). Continuous ordinal regression for analysis of visual analogue scales: The r package ordinalcont. *Journal of Statistical Software*, *96*(1), 1–25.

Mascha, E. J., & Sessler, D. I. (2011). Equivalence and noninferiority testing in regression models and repeated-measures designs. *Anesthesia & Analgesia*, *112*(3), 678–687.

Philip, B. K. (1990). Parametric statistics for evaluation of the visual analog scale. *Anesthesia & Analgesia*, *71*(6), 710.

Ritchie, H., Ortiz-Ospina, E., Beltekian, D., Mathieu, E., Hasell, J., Macdonald, B., ... Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. https://ourworldindata.org/coronavirus.

Robbins, C. L., Keyserling, T. C., Pitts, S. B., Morrow, J., Majette, N., Sisneros, J. A., ... Dietz, P. M. (2013). Screening low-income women of reproductive age for cardiovascular disease risk factors. *J Womens Health (Larchmt)*, *22*(4), 314–21. doi:10.1089/jwh.2012.4149

Roe, R. A., & Ester, P. (1999). Values and work: Empirical findings and theoretical perspective. *Applied psychology*, *48*(1), 1–21.

Rongen, A., Robroek, S. J., van Lenthe, F. J., & Burdorf, A. (2013). Workplace health promotion: A meta-analysis of effectiveness. *American journal of preventive medicine*, *44*(4), 406–415.

Ros, M., Schwartz, S. H., & Surkiss, S. (1999). Basic individual values, work values, and the meaning of work. *Applied psychology*, *48*(1), 49–71.

Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017. (n.d.).

Schaafsma, F. G., Mahmud, N., Reneman, M. F., Fassier, J. B., & Jungbauer, F. H. (2016). Pre-employment examinations for preventing injury, disease and sick leave in workers. *Cochrane Database Syst Rev*, (1), CD008881. doi:10.1002/14651858.CD008881.pub2

Selvarajah, S., Haniff, J., Kaur, G., Hiong, T. G., Bujang, A., Cheong, K. C., & Bots, M. L. (2013). Identification of effective screening strategies for cardiovascular disease prevention in a developing country:

Using cardiovascular risk-estimation and risk-reduction tools for policy recommendations. *BMC Cardiovasc Disord*, *13*(1), 10.

Sluiter, J., Weel, A., & Hulshof, C. (2013b). Leidraad preventief medisch onderzoek van werkenden. *NVAB, Utrecht*.

Soler, R. E., Leeks, K. D., Razi, S., Hopkins, D. P., Griffith, M., Aten, A., ... Task Force on Community Preventive, S. (2010). A systematic review of selected interventions for worksite health promotion. the assessment of health risks with feedback. *Am J Prev Med*, *38*(2 Suppl), S237–62. doi:10.1016/j.amepre.2009.10.030

Sørensen, K., Pelikan, J. M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., ... Uiters, E. (2015). Health literacy in europe: Comparative results of the european health literacy survey (hls-eu). *European journal of public health*, *25*(6), 1053–1058.

Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals.

Steel, J. S., Godderis, L., & Luyten, J. (2021a). Development and validation of a triage tool to follow-up workers' health. *Occup. Med. (Lond)*, *(under review)*.

Steel, J. S., Godderis, L., & Luyten, J. (2021b). Disclosure in online vs. face-to-face occupational health screenings: A cross-sectional study in belgian hospital employees. *International Journal of Environmental Research and Public Health*, *18*(4), 1460. doi:10.3390/ijerph18041460

Svensson, E. (2000). Comparison of the quality of assessments using continuous and discrete ordinal rating scales. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *42*(4), 417–434.

Takala, J., Hämäläinen, P., Saarela, K. L., Yun, L. Y., Manickam, K., Jin, T. W., ... Lim, S., et al. (2014). Global estimates of the burden of injury and illness at work in 2012. *Journal of occupational and environmental hygiene*, *11*(5), 326–337.

Task Force on Community Preventive Services. (2005). *The guide to community preventive services: What works to promote health?* Oxford University Press.

Torrance, I., & Heron, R. (2017a). Occupational health should be part of the nhs. British Medical Journal Publishing Group.

Valenty, M., Homère, J., Lemaitre, A., Plaine, J., Ruhlman, M., Cohidon, C., & Imbernon, E. (2015a). Surveillance programme for uncompensated work-related diseases in france. *Occupational Medicine*, *65*(8), 642–650.

Waddell, G., & Burton, A. K. (2006). Is work good for your health and well-being?

Walker, E., & Nowacki, A. S. (2011). Understanding equivalence and noninferiority testing. *Journal of general internal medicine*, *26*(2), 192–196.

Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in nursing & health*, *13*(4), 227–236.

World Health Organization. (1995). *Global strategy on occupational health for all: The way to health at work, recommendation of the second meeting of the who collaborating centres in occupational health, 11-14 october 1994, beijing, china*. World Health Organization.

Wynn, P. (2013). Periodic health checks in the workplace—is it time to change the prescription? Oxford University Press UK.

Zwerling, A., van den Hof, S., Scholten, J., Cobelens, F., Menzies, D., & Pai, M. (2012). Interferon-gamma release assays for tuberculosis screening of healthcare workers: A systematic review. *Thorax*, *67*(1), 62–70. doi:10.1136/thx.2010.143180

Chapter 4

METHODOLOGICAL CHALLENGES OF ECONOMIC EVALUATIONS IN OCCUPATIONAL HEALTH AND SAFETY

This chapter was previously published as:
Steel, J., Godderis, L., & Luyten, J. (2018a). Methodological challenges in the economic evaluation of occupational health and safety programmes. *International journal of environmental research and public health*, *15*(11), 2606. doi:10.3390/ijerph15112606

As a secondary objective, the thesis explores several key methodological topics at the crossroads of occupational health and safety, its efficient organization and economic evaluation. It describes the methodological pitfalls and challenges of economic evaluations in OHS, with a special focus on the estimation of health-related productivity.

4.1 Methodological challenges in economic evaluation of occupational health and safety interventions

Abstract

An emerging issue in occupational health and safety (OHS) is that interventions increasingly have to demonstrate that they offer sufficient value for money. To this end, the last decennia have seen more and more economic evaluation methods being employed in this field. However, several recent publications have indicated that many of the published studies suffer from important shortcomings. This paper aims to highlight difficulties in assessing the value of OHS by use of current economic evaluation methods. First, a summary framework presents an overview of the costs and benefits relevant for OHS interventions. Next, three elements from this framework are selected that are at the same time crucial to OHS value, but also challenging to measure and monetise: Effects on worker productivity, 'intangible' benefits, such as reputation effects, and the influence of the broader legal–fiscal context in which an intervention takes place. The following sections then discuss the following research questions for each of these elements: Why is it difficult to exclude these factors from OHS economic evaluations? Why do they pose a challenge to the quality of economic evaluations in OHS? How can they be included, and what are the known advantages and disadvantages of the methods to measure these factors? Future work should investigate (and standardise) better methods to include these elements.

Keywords: occupational health and safety; economic evaluation; methodology; workplace; productivity

4.1.1 Introduction

Matching the continuous emergence of new treatments and technological changes with limited resources is a crucial challenge in the healthcare domain (Chandra & Skinner, 2012). This delicate exercise makes it necessary for decision-makers to assess opportunities critically, in order to ensure that limited resources are invested in interventions with effects that merit their costs, without foregoing better alternatives. Economic evaluation can help to make these difficult decisions by comparing the costs and effects of different interventions. The last decades, publications related to economic evaluation have been strongly increasing in number.

An analogous increase of economic evaluation is noticeable for occupational health and safety (OHS) programmes—interventions that focus on improving the health, well-being, and safety of employees at the workplace (e.g., hazard control, return to work, screening, or health promotion at work). This is logical, as—although important improvements have been made—OHS faces similar challenges: Limited resources need to be reconciled with a considerable burden of OHS injuries and diseases (Heuvel et al., 2017), an aging working population (Burton & World Health Organization, 2010), and upcoming health hazards, such as nanomaterials or e-waste (Lucchini & London, 2014). Resources are perhaps even more pressurised in OHS, given the competitive enterprise environment in which OHS is often situated. The need for economic evaluations in OHS is therefore substantial (Burdorf, 2007).

However, several publications have highlighted that many of these economic evaluation studies are lacking sufficient scientific quality (Luyten et al., 2017). To some extent, this is because the measurement and attribution of the effectiveness of OHS interventions is inherently more challenging than it is in, e.g., the domain of pharmaceutical evaluation. For instance, it is more difficult to implement randomised controlled trials (e.g., because of legal requirements) (Tompa, Verbeek, van Tulder, & de Boer, 2010; van der Molen, Stocks, & Frings-Dresen, 2016), a wide range of health and non-health outcomes need to be measured (Edwards, Charles, & Lloyd-Williams, 2013; Payne, McAllister, & Davies, 2013), the stability of measurement units over time can be problematic (e.g., because firms can have "high rates of closing, merging, moving, downsizing or restructuring" (O. Olsen et al., 2008)), there is often a long latency period before (positive or negative) effects can be observed, and OHS possesses multiple features of complex interventions and systems (Craig et al., 2008; Shiell, Hawe, & Gold, 2008). In addition, some researchers have observed that many studies diverge from the 'reference case' (e.g., as described by Drummond (Drummond et al., 2015)) in how economic evaluation should be executed and reported: Underreporting of the employed methodology and context, not including all relevant costs and consequences, incorrectly valuing them, insufficiently accounting for uncertainty or justifying assumptions (Hamberg-van Reenen et al., 2012; Niven, 2002; Tompa et al., 2006; Uegaki et al., 2010).

In this article, we want to extend upon these areas for improvement by arguing that the current methods of economic evaluation are at times ill-devised to evaluate OHS programmes. More precisely, calculating the economic benefits of OHS interventions typically requires including several factors that are notoriously hard to measure and monetise. Most of these methodological shortcomings are well known, but in the context of OHS they become essential, as they concern factors that are (in contrast to other domains) of key value in OHS programmes.

4.1.2 Methods

As a first step, the authors present an overview of the costs and benefits associated with OHS investment. This is based upon frameworks of economic evaluation (Drummond et al., 2015), of economic evaluations in OHS (Burton & World Health Organization, 2010; Mossink & de Greef, 2002; Tompa, Culyer, & Dolinschi, 2008), of case studies identified by previous reviews (J. Steel, Godderis, & Luyten, 2018b; J. Steel, Luyten, & Godderis, 2018), and upon descriptions of the decision-making process in implementing OHS (van Dongen et al., 2013). The result of this exercise is presented in Section 4.1.3. The second step selects factors from the framework that contribute to a lack of quality in OHS economic evaluations. We focused upon factors that are an important source of value in an OHS setting (but are typically overlooked or less important in other healthcare domains) and are challenging to measure or monetise. Other challenges that are not OHS-specific (e.g., how to incorporate health effects, the use of discounting for costs and effects, how to measure or value grief and suffering of relatives), or OHS specific factors that are less challenging (e.g., avoided damage to factory equipment) are beyond the scope of this article.

Three such factors were identified: The effects of OHS on health-related productivity, intangible benefits of OHS, and accounting for the fiscal–legal context. Sections section 4.1.3 investigate these factors in turn by answering three research questions: Why is it difficult to exclude these factors from OHS economic evaluations, in contrast to economic evaluation in other domains? Why do these factors pose a challenge to the quality of economic evaluations in OHS? How can these factors be included, and what are the known advantages and disadvantages of the methods to measure these factors?

4.1.3 Three Challenges of Economic Evaluation in Occupational Health and Safety

Figure 4.1 provides a graphical overview of OHS costs and benefits. It does not aim to provide a comprehensive or exhaustive overview, nor to describe what should (or should not) be included, but rather serves to highlight the differences of OHS with other economic evaluation domains. The left panel illustrates the direct and indirect costs of OHS interventions from the perspective of employees, employers, and society. The right panel uses those same perspectives, and categorises effects of OHS programmes into productivity benefits, health benefits, health care savings, intangible benefits, and administrative and legal cost savings. Note that the societal perspective takes into account the relevant costs and benefits for all stakeholders (employees, employers, taxpayers, etc.), but without double counting. For instance, productivity effects for the employee

(income loss) and the employer (output loss) are closely related and should not be included twice. Since many OHS programmes rely on group-based decisions (e.g., by representatives of unions, employees, employers, occupational health specialists), it is often important to include multiple perspectives in the analysis.

What comes to the fore in comparison to other healthcare domains is the addition of the employer perspective and the particular costs and effects that accompany it. In evaluations in other settings, a patient or healthcare payer perspective is customarily used, which puts the largest weight in the decision-making process upon health effects and healthcare costs. In contrast, the employer perspective is more commonly adopted in OHS (J. Steel, Godderis, & Luyten, 2018b; Tompa et al., 2009). This is logical, as managers play a pivotal role in the funding and implementation of OHS interventions (Colosio et al., 2017). In turn, this reduces the weight of direct health effects (compared to other settings), and makes other (groups of) factors that are typically overlooked in other healthcare domains more important. The sections below discuss the challenging nature of three such factors: Worker productivity, intangible benefits (e.g., firm reputation and safety), and effects linked to the legal–fiscal context (e.g., employee indemnity claims, workers' compensation expenses (Ison, 1998), or financing OHS through subsidies or tax reductions).

Effects of OHS on Worker Productivity

Productivity loss is often equated with the fact that healthier individuals will show less absenteeism and presenteeism (reduced performance at work). However, it might be more precise to see it as lower efficiency in production, which means less output is obtained from a given set of inputs (Syverson, 2011). In economic evaluations in other healthcare domains, the inclusion of productivity losses (sometimes referred to as an indirect cost (Anders et al., 2013)) remains controversial, with some guidelines advocating their use (e.g., Sweden and the Netherlands), while others advise to exclude it (e.g., the UK, Belgium, and New Zealand) (Krol, Brouwer, & Rutten, 2013). Amongst other reasons, it is put forward that productivity losses are not a natural element of healthcare evaluations, and that their in- or exclusion only serves to boost cost-effectiveness ratios if so desired (Zhang, Bansback, & Anis, 2011). In contrast, the central role of the employer as decision-maker makes it difficult to follow this argument in OHS. Notwithstanding their controversy in other healthcare domains, productivity losses are one of the core reasons why employers choose to implement OHS interventions beyond required legal standards. A recent review indicated several evaluations even take productivity up as the sole benefit in the analysis (J. Steel, Godderis, & Luyten, 2018b). As such, it would be hard

to defend the systematic exclusion of productivity in OHS, and it makes the precision of its estimation of great importance.

However, the measurement and monetisation of productivity is complex (J. Steel, Godderis, & Luyten, 2018b). First, at least seven different valuation methods are available (output-based methods, human capital approach, friction cost approach, multiplier approach, productivity in natural units, US panel approach, and making use of workers' compensation expenses), and there is as of yet no consensus on which one should be used in which circumstances (J. Steel, Godderis, & Luyten, 2018b). Second, it is not straightforward to obtain qualitative absenteeism and presenteeism data. Objective sickness days data are often incomplete (e.g., human resource data are not always accurate, insurance data often focus upon compensated absences), while subjective (self-reported) absenteeism (and presenteeism) is limited due to recall bias. Furthermore, the magnitude of time-loss estimates depends on the chosen survey, and not all surveys have been validated (Krol & Brouwer, 2014; Mattke, Balakrishnan, Bergamo, & Newberry, 2007). Third, to monetise absenteeism and presenteeism, an adequate price weight (e.g., the employee's wage) is needed. Depending on the guideline followed, the composition (are secondary benefits included?) and type (job, individual, or industry wage?) of the price weight differs, and requires additional data. Finally, if one wants to ensure that this monetary estimate reflects real loss in productivity, adjustments are needed. For instance, a (national) friction period must be estimated (only available in the Netherlands and the UK (Kigozi, Jowett, Lewis, Barton, & Coast, 2017a; M. A. Koopmanschap, Rutten, van Ineveld, & van Roijen, 1995)), or job-dependent multipliers should be used to indicate the absence of perfect substitutes, the effects of team production, and the penalty of not reaching an output target (S. Nicholson et al., 2006; Pauly et al., 2002).

Since there is as of yet no consensus on how to measure productivity, we recommend on the one hand to follow (national) guidelines (e.g., Oostenbrink et al. (Oostenbrink, Koopmanschap, & Rutten, 2002) for the Netherlands), as this contributes to standardisation. On the other hand, it has been observed that the appropriateness of (valuation) methods may depend on the intervention, outcome, perspective, time horizon, and study design of the evaluation (Lensberg, Drummond, Danchenko, Despiegel, & Francois, 2013; J. Steel, Godderis, & Luyten, 2018b). We therefore suggest to include productivity separately in economic evaluations, and to perform analyses using several methods. Given the strong differences between the two (Lensberg et al., 2013), one could argue that, at minimum, the human capital and friction cost approach should be applied, but sensitivity analyses with the multiplier approach and including

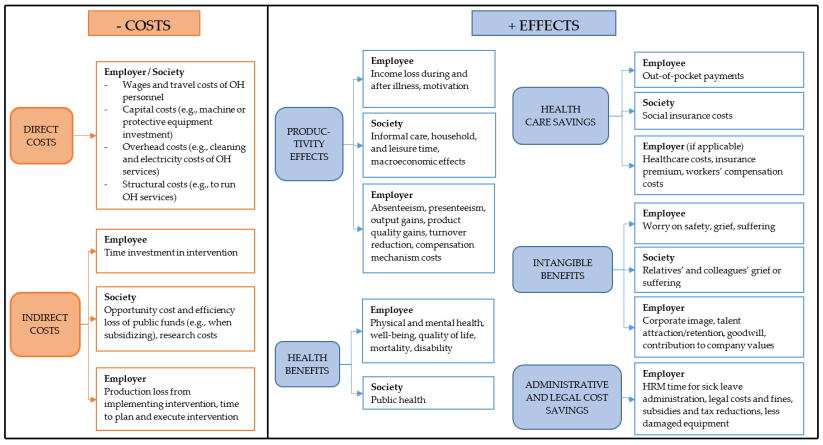


Figure 4.1: Costs and effects of occupational health and safety interventions, by cost and effect category and by perspective. [Blue = intervention benefits or effects; Orange = intervention costs, OH = occupational health, HRM = Human Resource Management].

costs of compensation mechanisms can be added. Since clear instructions on how to perform the different methods are available (e.g., the Dutch Costing Manual (Tan, Bouwmans, Rutten, & Hakkaart-van Roijen, 2012) for the friction cost approach), this should be a feasible opportunity. For instance, Van Wier et al. (van Wier et al., 2013) use both the human capital and the friction cost approach from a societal and employer perspective. Regarding productivity measurement, using objective data on sick leave that are separable from other absences (e.g., holidays) is often the most precise. If such data are absent, several validated and reliable instruments are available (e.g., PRODISQ, HLQ, WHO-HPQ, iPCQ) for self-reported absenteeism (Lofland, Pizzi, & Frick, 2004; J. Steel, Godderis, & Luyten, 2018b), although the recommended recall periods of these instruments should be respected. Presenteeism is more controversial, and is best measured by self-report (unless production data are available, e.g., number of files handled per hour in a phone centre), e.g., using the iPCQ, or HLQ (Kigozi et al., 2017a; J. Steel, Godderis, & Luyten, 2018b). Since estimates of presenteeism can differ across instruments, it seems best to follow the recommendations of guidelines (e.g., the Netherlands advises the use of iPCQ).

Intangible Benefits of OHS

In economic evaluations, the term "intangible costs" (or "intangible benefits") is used to indicate costs and consequences that are difficult to measure and monetise (Drummond et al., 2015), although there is not always a clear consensus on what this encompasses. Some examples in a clinical setting are effects on the social context (e.g., aggressiveness caused by substance abuse (Jeanrenaud & Pellegrini, 2007)), pain and grief of family and friends, or effects upon quality of life (e.g., pain, joy, physical limitations) (Anders et al., 2013; Jeanrenaud & Pellegrini, 2007). Because of the methodological difficulties posed with quantifying or valuing these effects, most studies do not include them. As Figure 4.1 indicates, OHS programmes lead to several intangible costs and benefits and these are a key part of the value offered by OHS programmes. Some of these occur in other healthcare domains—quality of life effects, pain and suffering—but OHS also has some unique intangible consequences.

First, investments in OHS can affect a firm's labour pool. A good reputation on the labour market can contribute to attracting more and more talented workers, and can help to retain key employees. Accumulating the right combination of human capital could in turn help firms to obtain a competitive advantage (Brymer, Molloy, & Gilbert, 2014; Kucherov & Zavyalova, 2012). Given the current workforce demographics (e.g., retirement of baby boomers), and especially in sectors or occupations with high turnover rates (e.g., teachers (Troman & Woods, 2000) or nurses (Buchan & Campbell, 2013)), this becomes an important factor to consider.

Second, investments in OHS can have an effect upon costs associated with a bad reputation. 'Sweat shops' (Ponsonby, 2017), unequal treatment of migrant workers (Ponsonby, 2017), textile factory accidents (Chowdhury & Tanim, 2016), or excessive overwork (McCurry, 2017) can catch public attention and affect a firm's image, and thus possibly hit customer brand loyalty or sales. A survey among large UK companies (the yearly Captains of Industry Survey) also indicated this as a concern among managers: 79% of the senior directors saw health and safety as having a tangible impact upon corporate reputation (Smallman & John, 2001), while later versions of the survey (e.g., in 2012) noted reputation or company brand as a key factor to stand out as a company (Clarke, 2013).

Third, many investments in OHS will affect the risks and safety of the workplace. Given its different definitions (Aven, 2010) and disciplinary contexts (Aven & Zio, 2014), the challenging nature of risk manifests itself on several levels in occupational health and safety. From the viewpoint of risk analysis, challenges rise in the appropriate assessment and management of hazards in the workplace

to minimise risks on employees' health and safety. In essence, the focus of these exercises lies upon mapping out the likelihood of hazardous events and the severity of the health burden they cause. However, questions are rising on whether such probability-based approaches are too narrow for risk assessments. Instead, some maintain the view that probabilities cannot reflect all the knowledge and information on uncertainty that is needed (Aven, 2010; Aven & Zio, 2014). For instance, Paté-Cornell (Paté-Cornell, 2012) is of the opinion that "perfect storms"—rare conjunctions of factors with known probabilities—require systemic methods that pay attention to dependencies among factors, and "black swans"—rare and difficult to anticipate events—call for proper attention to (less tangible) signals and precursors. In addition, Aven (Aven & Zio, 2014) suggests that qualitative methods are needed to ensure all aspects (e.g., social and political aspects, value judgements such as the (pre)cautionary principle) are taken up in decision-making. In this sense, both can be seen in the wider trend towards risk-informed decision making—seeing risk analysis as a decision support that aims to inform decision-makers—rather than risk-based decision making—(quantitative) risk analysis as the sole basis for decision making (Apostolakis, 2004).

From the viewpoint of decision modelling (Briggs, Sculpher, & Claxton, 2006), a challenge of safety in OHS lies in making adequate decisions (e.g., to implement a programme or not) given a scarcity of high-quality evidence on OHS interventions' cost-effectiveness. This scarcity multiplies the risk of making a wrong decision when choosing whether to implement an intervention, and thus the associated costs of making the wrong decisions (the opportunity cost of foregone funds). In situations of high stake investments with high uncertainty, it can therefore be useful to not only compare the costs and benefits of implementing or not implementing a programme, but to also compare it with postponing the decision until more evidence is available (a value of information approach (Briggs et al., 2006)).

Finally, in the field of economic evaluation, there is the challenge of how risk and safety should be taken up in the analysis. A common counterargument against including effects on safety is that it can lead to double counting, as increases in safety overlap with the health improvements that come along with them (e.g., fewer accidents or fractures). However, in OHS, safety seems to have benefits beyond their eventual translation into health improvement. A first observation is that employees worry when there is a lack of safety. Conversely, knowing that everything has been done to ensure a safe work environment might reduce these concerns, giving a clear additional benefit for employees. Next, insights from behavioural economic and psychological literature indicate individuals have other attitudes towards risk and safety than towards health. For example, they value high-impact incidents that occur infrequently (e.g.,

a terror attack or large accident) as disproportionally more burdensome than small incidents that occur frequently (e.g., getting influenza) (Slovic, 1987), even though the aggregate impact of the latter can be bigger (in terms of both total health impact and associated costs). If economic evaluations do not take these attitudes into account and look only at the expected health effects of an OHS intervention, return-on-investment estimates might not result in the right prioritisation of programmes (i.e., in line with employees' preferences).

In contrast to clinical settings, a study by Miller and colleagues (P. Miller & Haslam, 2009) indicates several of these intangible benefits are influential in managers' investment decisions in OHS: Interviews point out their choices rely on a combination of financial, ethical, social, and legal factors, as well as intuitive arguments on people management or corporate reputation. An economic evaluation aiming to aid employers in their decisions could therefore incorporate these effects where they are deemed important. At the moment, a first step to include these effects (upon the labour pool, firm reputation and sales, less tangible signs of risks, employee worry, and attitudes towards risk) is to highlight their existence to decision-makers when they are likely to change because of the intervention. However, this leaves the weighting of these factors in the decision entirely in their hands. Second, as indicated by Drummond et al. (Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005), "intangible" can be a misleading term, as these effects can still be measured and included in analyses. For instance, quality of life effects can be assessed through surveys (e.g., through the SF-36 health survey (Ware Jr & Sherbourne, 1992)), and other intangible costs are—in principle—measurable through willingness-to-pay estimates, indicating how much money people are willing to sacrifice for avoiding these intangible burdens. A more challenging possibility is thus to estimate intangible effects directly (e.g., surveys measuring employees' or managers' perceptions of labour pool effects or willingness-to-pay for risk reductions) or indirectly (e.g., through their effect upon sales or productivity). Finally, instruments exist that try to take intangible OHS effects into account, such as sustainability reports (a company reporting publicly on its economic, environmental, and/or social impacts) (Sancroft, 2017) (although these have received criticism (Evangelinos et al., 2018)), instruments that measure employee motivation (Gagné et al., 2010), job satisfaction (Piko, 2006), firm reputation (Cravens, Oliver, & Ramamoorti, 2003), or perceived safety climate (Barling & Hutchinson, 2000). While these instruments often originate from the OHS community, they have not been regularly considered as outcomes in economic evaluations of OHS. In general, more methodological work is needed to ensure these issues can be taken up consistently.

Accounting for the Legal and Fiscal Context

OHS is, more than other healthcare domains, set in a complex social, legal, and fiscal context that strongly diverges between countries. Dekker et al. have pointed out the far-going bureaucratisation of safety in many countries, including "increases in rules, paperwork, costs, time drain, safety people involved, and compliance expectations that are insensitive to the demands of front-line activities" (Dekker, 2014). While acknowledging the safety gains, rationalisation, order, and efficiency that safety administration has brought, they claim the yield of further bureaucratisation is declining or plateauing in many industries (Dekker, 2014). Regardless of their effectiveness, extensive safety regulations complicate economic evaluations of occupational health and safety programs as they entail a significant administrative burden and time investment for the employer, who may therefore be less inclined to devote time to other programs. In addition, safety regulations differ across industries and countries, posing challenges to the generalisability of research results.

On the cost side, OHS programmes are often subsidised by the government, or tax reductions can be applicable to purchases related to OHS. A report by the European Agency for Safety and Health at Work (EU-OSHA) indicates most European countries take measures to incentivise OHS investment (Esler et al., 2010). While tax incentives were rather rare, most EU countries make use of insurance-related incentives (e.g., a bonus–malus system for insurance premiums based on a companies' accident rates), and nearly every EU country used funding schemes for OHS (e.g., subsidies or grants for purchasing materials and tools, or to implement OHS management systems) (Esler et al., 2010). For instance, in the German butchery sector, a combination of positive premium variations (when investing in occupational safety) and funding schemes for safety and health, was introduced (Esler et al., 2010).

On the effects side, an important part of the benefits of OHS is avoiding costly occupational injury, disease, or death. Without countermeasures, the burden of not investing in OHS falls largely on the employee and their families, as they would bear the health and income consequences (e.g., loss of wage or rise in healthcare costs). For this reason, several systems have been devised that help to relieve this burden upon employees, by shifting (part of the) costs back to the employer or to society. These systems range from worker compensation systems (e.g., USA or Canada), sick pay, taking OHS compensation up in the general social insurance system, accident compensation, (employer organised) disability insurance, to employers' liability (and the possibility to legally claim compensation) (Ison, 1998). The exact implementation of these systems also differs. For instance, while the Netherlands requires up to two years of sick pay to be paid by the employer, this is only up to 30 days in Belgium, with social

insurance taking up a much larger portion of the payments. Each system, thus, has very different consequences on who (employee, employer, society) bears what portion of the burden of occupational disease and injury. Some systems will result in little reimbursements to employees (e.g., due to underdeveloped OHS legislation), others will emphasise the responsibility of the employer and ask them to compensate the majority of the costs to the employee, while again other systems will rely more strongly upon society's solidarity by shifting OHS costs to the general social security system.

In turn, these differences can strongly affect the economic benefits from investing in OHS interventions, even when exactly the same intervention is carried out (J. Steel, Godderis, & Luyten, 2018b). An economic evaluation of an OHS programme, aiming to indicate the expected return to a stakeholder relative to its required costs, should therefore carefully assess and report what OHS compensation systems are in place to correctly determine the consequences of OHS investments and disentangle the burden and benefits of each stakeholder. The relevance of this context is obvious from an employer's perspective as, in this case, fiscal–legal effects are the most tangible. However, ideally, economic evaluation involves taking a societal perspective, as this ensures that the effects upon the healthcare system as a whole (and thus taxpayers) are taken into account. This can be complicated, as particular money streams (e.g., subsidies or fines) must be seen as transfers rather than costs. Precisely performing and reporting this exercise becomes even more important when one realises economic evaluations in one setting are often used to inform decisions in other settings, since many choices (rightfully) rely on prior research rather than re-evaluating the same programmes.

4.1.4 Discussion

Economic evaluation can be a valuable tool for decision-makers in OHS facing resource allocation tasks, as it provides detailed information on the costs and outcomes of interventions. When carefully executed, it provides useful estimates of the return-on-investment of competing programmes. However, several publications have outlined that published studies are often insufficiently reliable and that better economic evaluation studies are needed. This article aimed to point out that this is a difficult task, as the available methods of economic evaluation are often ill equipped for the particular OHS context. It did so by exploring the challenges in correctly measuring three effects that are typically not encountered (or deemed less relevant) in other healthcare domains: Productivity losses, intangible effects, and effects of the legal and fiscal context. As such, it contributes to previous work (Tompa et al., 2006; Tompa et al., 2010) that offered broader methodological guidelines in OHS, and it answers

to the call for more research on economic evaluation in occupational health (Lalloo et al., 2018). The three elements discussed above remain difficult to address. They are less relevant in other contexts, which makes them less of a methodological priority. However, OHS does not always fit into the standardised approaches for other disciplines, as these elements cannot be as readily excluded because of their importance in the decision-making process (our first question).

Instead, they remain challenging factors that need to be included (our second question): Estimating productivity loss knows many practical difficulties (how to adequately measure absenteeism and presenteeism) and methodological disagreements (what should be included and how should it be monetised). Intangible benefits are by definition difficult to grasp and monetise within an economic evaluation. Nevertheless, they indirectly lead to tangible consequences: Effects upon the labour pool and firm reputation can lead to changes in firm profit, picking up less tangible signs of risks can avoid costly accidents, and reducing an employee's worries could benefit work satisfaction and production. Finally, the diversity of the legal and fiscal context and the bureaucratic burden that often accompanies it complicates research and the generalisability of its results. Each of these elements, thus, poses a barrier for economic evaluations of OHS to attain a higher quality.

While we referred to recent methodological work that tries to resolve these matters, there is as of yet no clearly dominating methodology (our third question), and all of these elements would benefit from more fundamental research. Productivity needs methodological work on estimating friction periods for other countries, correct estimation of presenteeism (recent attempts attempted to link it to health (Jones, 2018; Krol, Stolk, & Brouwer, 2014)), and the implementation of job-dependent multipliers. For now, it seems best to follow guidelines as much as possible, but to use sensitivity analyses to assess the effects of different valuation methods and corrections. For intangible effects, awareness (when relevant) is the starting point, but methodological work is needed to see how these can be consistently included in the decision-making process. Effects of the legal and fiscal context are in less need of methodological work, but the quality of economic evaluations and models would benefit from more thorough investigation and reporting of contextual factors by researchers performing economic evaluations.

The challenges of economic evaluation in OHS do not restrict themselves to the three elements discussed above. As indicated in the introduction, the measurement and attribution of effectiveness in OHS is particularly challenging, and often requires creative methodological solutions. Next, the loss of a colleague, friend, or family member can have substantial psychological consequences for the social environment of an employee, but it is unclear whether and how this should be taken into account. Normative standpoints also often remain implicit

in analyses. For instance, the inclusion of productivity costs in economic evaluations can lead to a wider implementation of programmes that target productive individuals. By contrast, from an equity point of view, many populations that are less productive (e.g., retired employees, children) require more attention. However, it is as of yet not evident how to control for these effects and, more broadly, whether equity aspects should be considered altogether. Finally, barriers exist between researchers and business. For example, managers are not always well acquainted with economic evaluation methods (e.g., the difference between cost-effectiveness, cost–utility, or cost–benefit analyses), and have more affinity with simple summary measures (return-on-investment, payback period, benefit–cost ratio) that focus upon monetary consequences, while their use has substantial drawbacks (Ambler & Roberts, 2008; Rosen & Gayer, 2010; Steen & Welch, 2011). Since these topics, as well as the three elements discussed above, build upon multiple domains—(health) economics, occupational health and safety, risk analysis, psychology—multidisciplinary research is key to work towards solutions.

4.1.5 Conclusions

In summary, general guidelines on economic evaluation remain largely applicable to the field of OHS, but researchers should pay attention to the specific challenges listed in this and other articles (Tompa et al., 2006; Tompa et al., 2010). To ensure economic evaluations in OHS adequately inform decision-makers (especially employers), future methodological work should continue research on productivity, intangibles, and legal/fiscal effects, and work towards standardised methods of including them in economic evaluations, either directly or indirectly. We hope this article can help to generate more research in this direction.

4.2 Productivity estimation in economic evaluations of occupational health and safety interventions: a systematic review

This chapter was previously published as:
Steel, J., Godderis, L., & Luyten, J. (2018b). Productivity estimation in economic evaluations of occupational health and safety interventions: A systematic review. *Scand J Work Environ Health*, *44*(5). doi:10.5271/sjweh.3715

Abstract

Objectives Occupational health and safety (OHS) interventions' effect on worker productivity is an essential, but complex element of the value of these programs. The trustworthiness of economic evaluation studies, aiming to provide guidance to decision-makers in the field of OHS, depends at least partly on how accurately productivity changes are measured. We aim to review the methods used to estimate productivity changes in recently published economic evaluations of OHS interventions.

Methods We performed systematic searches of economic evaluations of OHS programs published between 2007 and 2017 and reviewed these studies' methods to quantify the programs' impact on worker productivity.

Results Of the 90 identified studies, 44 used a human capital approach, 17 a friction cost approach, 13 stated productivity in natural units (eg, a cost-per-absence-day-avoided), 7 made use of compensation expenses, 4 used output-based methods, 4 an "ad hoc" approach, and 1 study did not state its method. Different approaches were combined in 19 studies. Within these methods, we observed a wide diversity in their precise implementation, especially regarding the measurement and valuation of absenteeism and presenteeism.

Conclusions Productivity is a key element of the economic attractiveness of investing in OHS. Economic evaluation studies of OHS would benefit from more methodological standardization in their approach to quantifying productivity change. Future research should better account for the methodological uncertainty that occurs in estimating it in order to demonstrate the impact that particular choices and approaches to productivity estimation can have on cost-effectiveness results.

Key terms absenteeism; friction cost; human capital; indirect cost; methodology; measurement; occupational health; presenteeism; valuation; work.

4.2.1 Introduction

Productivity loss can be broadly defined as reduced efficiency in production: less output is obtained from a given set of inputs (Syverson, 2011). In occupational health and safety (OHS), avoiding these productivity losses is an indispensable element of the value of interventions: a healthy workforce reduces the need for costly (internal or external) replacement workers, and avoids output not being produced (eg, when no replacement is found). Whereas in the economic evaluation of healthcare, productivity effects are often excluded for ethical or pragmatic reasons (Drummond & McGuire, 2001; Drummond et al., 2015; Krol & Brouwer, 2014; J. A. Olsen & Richardson, 1999; Zhang et al., 2011), in the context of OHS, they are often a central point of focus (Luyten et al., 2017). This is logical, as in this domain, investments beyond the strict legal minimum requirements are often made at the discretion of managers and shareholders, for whom productivity gains are a major element of OHS' economic attractiveness.

However, defining and quantifying productivity change is a complex endeavour, and several approaches exist to estimate it: output-based methods, productivity in natural units, the human capital approach (HCA), the friction cost approach (FCA), the US panel approach (USP), the multiplier approach (MA), and workers' compensation expenses (CE). The rest of the introduction discusses these methods and their challenges in turn, and Table 4.1 indicates additional advantages and disadvantages. Note that aspects of lost-time measurement stated for CEA LT are also applicable to the HCA and FCA.

A first approach measures output change directly. However, production data are not always available or costly to obtain, while (more available) revenue data might reflect market circumstances or other effects on profits (Syverson, 2011). In addition, defining and measuring output is not always evident (see Table 4.1), and even with an output measure, it is difficult to link this to individual employees (for instance when work is performed in teams or projects) (Hassink & van den Berg, 2017). Instead, the most common approaches to estimate productivity make use of data on labor inputs (Zhang et al., 2011). In this way, they measure the effect (of occupational health improvements) upon the volume (number of hours worked) and quality (human capital) of labor. A second approach to include productivity in OHS is thus to measure "lost-work-time" and state the conclusion of the study as a cost per natural unit (eg, a cost per absent day avoided). This requires measuring absenteeism – health problems leading to reduced time at work – and (possibly) presenteeism – health problems leading to reduced performance while working (Zhang et al., 2011), by means of objective data or surveys. For the latter, a wide variety of standardized instruments is available, which differ in length, scope, method, and can be disease-specific or general (Mattke et al., 2007). A detailed comparison

Table 4.1: Definition, advantages and disadvantages of productivity estimation methods. [QALY=quality-adjusted life year]

Method	Definition	Advantages	Disadvantages
Output-based methods	Estimate productivity costs by analyzing output or revenue data	Automatically adjusts for different effects upon productivity, such as compensation mechanisms, multipliers, etc; allows estimation of multifactor productivity (relating a measure of output to multiple inputs such as capital or labor) instead of focusing on a single factor (ie, labor) (De Loecker & Goldberg, 2014; Organisation for Economic Co-operation and Development, 2001; Syverson, 2011)	Production data are not always available or costly to obtain; it can be difficult to define and measure output, eg, in services or knowledge industries, if one firm or employee produces multiple outputs, or when output differs in quality; difficult to link productivity changes to individual employees (eg, in case of team work)
Cost-effectiveness analysis expressing productivity in natural units CEA LT	Measure lost-time and state the conclusion of the study as a cost per natural unit	Intuitive; easily implemented; employee surveys are valid and affordable given an adequate recall period (Mattke, Balakrishnan, Bergamo, & Newberry, 2007) [eg, 3 months for absenteeism, one week for presenteeism (Zhang, Bansback, & Anis, 2011)]	Lack of high-quality (uncontaminated) objective data on absenteeism and presenteeism; magnitude of time-loss estimates depends on the chosen survey and not all surveys have been validated (Krol & Brouwer, 2014; Mattke, Balakrishnan, Bergamo, & Newberry, 2007); some surveys do not grasp all relevant dimensions, or include too much (such as quality of life, stress, job satisfaction, etc.) (Zhang, Sun, Woodcock, & Anis, 2015)
Human capital approach (HCA)	Monetizes lost-time with a price weight, such as individual or average wages	Wide registration of the required data; computational ease; accordance with economic theory (Lensberg, Drummond, Danchenko, Despiegel, & Francois, 2013; Zhang, Bansback, & Anis, 2011); broader perspective	Assumes no replacements (or compensation mechanisms) take place in the economy
Friction cost approach (FCA)	Revision of the HCA that focuses on the time required to restore the lost productivity due to the absent worker (Lensberg, Drummond, Danchenko, Despiegel, & Francois, 2013)	Takes replacements (and costs of compensation mechanisms) into account; adjusts for the elasticity of labor to output	Theoretical inconsistencies (Johannesson & Karlsson, 1997; M. A. Koopmanschap, Rutten, van Ineveld, & van Roijen, 1997); more narrow view (Lensberg, Drummond, Danchenko, Despiegel, & Francois, 2013); stronger data requirements (eg, the duration of absence spells)
US panel approach (USP)	Productivity is part of quality of life estimates in health state utilities (eg, QALY)	Does not require separate estimation of productivity	Quality of life surveys can ask their respondents not to take income or productivity effects of health states into account (J. A. Olsen & Richardson, 1999); not often used in OHS or other healthcare economic evaluations
Multiplier approach (MA)	Apply job-dependent multipliers to wages	Accounts for the impact of the absence of perfect substitutes, team production, and output penalties on absenteeism cost	Needs a substantial library of multipliers; requires more research (Burdorf, 2007; Krol & Brouwer, 2014), especially regarding presenteeism; multipliers are assumed to be constant across firms, while jobs can be very different across enterprises (Mattke, Balakrishnan, Bergamo, & Newberry, 2007).
Workers' compensation expenses (CE)	Include CE directly in the analysis or calculate CE per employee and use it as the price weight for valuing lost-time	Wide registration of the required data; direct cost to the employer	Not necessarily related to the loss of productivity capacity (Mossink & de Greef, 2002; Tompa, Dolinschi, de Oliveira, & Irvin, 2009); can give rise to contaminated absence data (eg, compensated days only)

of survey instruments is beyond the scope of this review, but can be found elsewhere (Kigozi et al., 2017a; Krol & Brouwer, 2014; Lofland et al., 2004; Mattke et al., 2007; Prasad, Wahlqvist, Shikiar, & Shih, 2004; Uegaki, de Bruijne, van der Beek, van Mechelen, & van Tulder, 2011; Zhang et al., 2011). Table 4.1 indicates some additional measurement challenges.

Apart from these two broad approaches to quantifying productivity change, three concrete estimation methods are suggested in economic evaluation guidelines (Drummond et al., 2015). In HCA, the monetary value of productivity losses is approximated by multiplying lost working hours or days with a relevant value of production per time unit – a "price weight" – such as individual or average wages (Uegaki et al., 2011). This is in accordance with the assumption (from economic theory) that the marginal revenue product of labor is proportionate to the wage (Kigozi et al., 2017a; Zhang et al., 2011). However, the HCA also (3) implicitly assumes that changes in inputs (absenteeism and presenteeism) have a direct linear relationship with changes in output (lost production) (Krol, Brouwer, Severens, Kaper, & Evers, 2012), whereas compensation mechanisms can mitigate some production losses: less important tasks can be postponed and made up for on return, internal labor slack/reserves could be used, and temporary or permanent replacements can be trained and hired (M. A. Koopmanschap et al., 1995; Krol et al., 2012). As such, some criticize the HCA because it takes a long(er) time horizon, where productivity costs can become quite substantial (eg, when they accumulate over several years) and may include lost productivity due to disability or early retirement (Lensberg et al., 2013). FCA instead proposes to limit long-term production losses to the "friction period" [the time needed to replace a sick worker (M. A. Koopmanschap et al., 1995)] and to adjust shorter absence spells for the "elasticity of labor to output" – the less than proportional change in output per unit of labor – to reflect compensation mechanisms (Kigozi, Jowett, Lewis, Barton, & Coast, 2016; M. A. Koopmanschap et al., 1995). This friction period will likely depend on the location, industry, firm, and category of worker (Drummond et al., 2015). Finally, the costs of compensation mechanisms (eg, costs of hiring workers) should ideally be included (Brouwer & Koopmanschap, 2005; Kigozi et al., 2016). A third approach, USP, argues that productivity should not be measured separately since it is part of quality of life estimates in health state utilities such as quality-adjusted life years (QALY) (Drummond et al., 2015; Gold, Siegel, Russell, Weinstein, & editors, 1996; Uegaki et al., 2011; Zhang et al., 2011). However, up to now, neither QALY nor the USP approach have been frequently used in OHS (Tompa et al., 2009).

A theoretical model by Pauly et al gives rise to a fourth method: the multiplier approach (MA) (S. Nicholson et al., 2006; Pauly et al., 2002). It poses a critique on salary conversion methods, such as the HCA and the FCA, by stating that

the cost of absenteeism can be "higher than the wage when perfect substitutes are not available to replace absent workers, and there is team production or a penalty associated with not meeting an output target [for a given time interval]" (Pauly et al., 2002). Job-dependent wage multipliers have been suggested – and to a certain extent estimated – to correct for this underestimation (Krol et al., 2012; S. Nicholson et al., 2006; Pauly et al., 2002). However, while the empirical impact of wage multipliers has been demonstrated (Kigozi et al., 2017a; Krol et al., 2012; Strömberg, Aboagye, Hagberg, Bergstrom, & Lohela-Karlsson, 2017), little research has been done on how to include them, especially in the case of presenteeism (Burdorf, 2007; Pauly et al., 2002).

Some studies make use of compensation expenses data (Tompa et al., 2009). The regulations for workers' compensation differ strongly region by region (Burton & World Health Organization, 2010; Mossink & de Greef, 2002) and the expenses are often a mix of replacement wages, disability benefits, insurance expenses and healthcare costs. They can be included directly in the analysis (eg, by using the total sum of compensation expenses) or can be calculated per employee and used as the price weight for valuing lost-time (ie, a variant on the HCA). While compensation data can be easy to collect, the size of these payments is not necessarily related to the loss of productivity capacity (Mossink & de Greef, 2002; Tompa et al., 2009) and can give rise to contaminated absence data, since some datasets will only contain "compensated days" instead of all sick leave.

While all approaches estimate the marginal effect of the intervention upon productivity, they can also be seen as fulfilling distinct purposes, and the appropriate method may depend on the intervention, outcomes, perspective, time horizon, study design, etc (Lensberg et al., 2013). It might even be interesting to use several methods (eg, the HCA and FCA) (Drummond et al., 2015; Krol & Brouwer, 2014) or to combine them. For instance, the FCA can include additional measures to control for compensation mechanisms (correcting not only for long-term replacements, but also for colleagues' work), or can use job-dependent multipliers.

Earlier studies have systematically reviewed how the economic evaluation literature of OHS, up to 2007, estimates productivity change. These authors observed significant diversity within different applications of the same methods, thereby hindering comparability of results and decision-making and ample margins for improvement (Kigozi et al., 2016; Tompa et al., 2009; Uegaki et al., 2011). In this article, we aim to extend on these studies by reviewing economic evaluations of OHS interventions that have emerged in the decade after the last review (2007–2017).

4.2.2 Methods

The study selection process and outcome is set out in Figure 4.2. We conducted systematic searches in Embase, PubMed, NHS EED, DARE and Cochrane Library from 2007 up to 2017. Search terms included controlled vocabulary (MeSH and Emtree) and free-text terms, and consisted of three sets: related to economic evaluation, workplace settings or occupational interventions, and productivity. Additional articles were identified from reviews (Carroll, Rick, Pilgrim, Cameron, & Hillage, 2010; Elsler et al., 2010; Hamberg-van Reenen et al., 2012; Neuhaus et al., 2014; Osilla et al., 2012; Pelletier, 2009, 2011; Shrestha et al., 2016; Torbeyns, Bailey, Bos, & Meeusen, 2014; Uegaki et al., 2011; van Dongen et al., 2011; Verbeek, Pulliainen, & Kankaanpaa, 2009), reference screening, and the authors' own database.

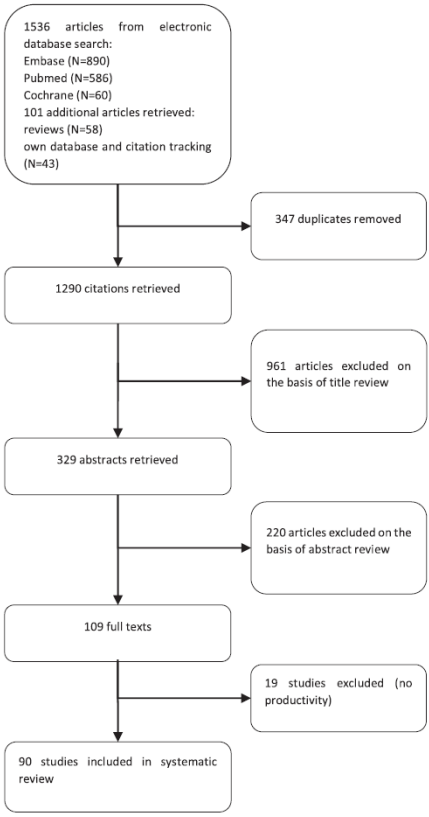


Figure 4.2: Study selection process and outcome

Studies were included based on four criteria: the analysis (i) was a full or partial economic evaluation; (ii) included OHS interventions targeted at an employed (or return-to-work) population; (iii) included productivity in paid work; and (iv) was written in English, French, or Dutch. Studies were included if they concerned an intervention on health or safety at work, regardless of whether these interventions considered primary, secondary, or tertiary prevention.

After study selection, a data extraction form was created based on previous literature (Drummond et al., 2015; Evers, Goossens, de Vet, van Tulder, & Ament, 2005; Kigozi et al., 2016, 2017a; Uegaki et al., 2011), to appraise the studies included in the review. A first set of measures describes general study characteristics: study method, type of economic analysis (following

Drummond et al., 2015), perspective, country of the study, sector, study design, and intervention details. Studies were then assessed on how productivity was estimated (the methodological approach), whether they included absenteeism and presenteeism, and how this was measured (objectively or through surveys, what instrument and its characteristics, and data sources). The method to measure presenteeism was also indicated: perceived impairment (asking employees how much their illness hindered them in performing tasks), comparative productivity (versus other colleagues or usual performance, often using Likert scales), and directly asking employees to estimate unproductive time at work Kigozi et al., 2016; Mattke et al., 2007. When applicable, the valuation/monetization methods of lost time were analyzed, as well as their characteristics: price weights, friction period length used and source for this friction period estimate, and the use of the elasticity of labor to output. Finally, studies were assessed whether they controlled lost output for compensation mechanisms, used multipliers, or included additional costs (such as the costs of compensation mechanisms).

4.2.3 Results

Study background

Study characteristics of the 90 included studies are represented in Table 4.2. Of these, 88 studies were written in English, with the remaining in French (Dutheil et al., 2008), and Dutch (Hak, Knol, Wilschut, & Postma, 2010). Studies originated in Europe (56%), North America (38%), or other continents (7%), were performed in diverse sectors and industries, and studied a wide range of OHS intervention types. A substantial part (39%) of these economic evaluations was based on data from randomized controlled trials, and (18%) used uncontrolled before–after studies, of which some employed statistical analyses to control for confounders [eg,(J. Brown et al., 2015)], while others simply attributed all changes in outcomes to the intervention [eg, (Seeley, Marklin, Usher, & Yager, 2008)]. Decision-analytic models were used in 14 studies (16%); other designs (28%) are detailed in the table. Cost-effectiveness analyses (CEA) were conducted in 71 studies (79%), 19 were cost-utility analyses (CUA) (21%), using QALY as outcome measure. Most of these CUA also included a CEA. Regarding primary perspective, 37 studies (41%) adopted a societal perspective, taking into account costs and effects that were not strictly relevant to the firm. Costs and effects relevant for the employer were the focus of 24 studies (27%), while 3 (3%) took a healthcare provider or insurance perspective. The other 26 studies (29%) did not explicitly mention a perspective, however, for all but 3 (Schene, Koeter, Kikkert, Swinkels, & McCrone, 2007; Smedley, Harris, Cox,

Ntani, & Coggon, 2013; Thijs et al., 2012), it could be inferred based on the context. Analyses were repeated from an additional perspective in 23 studies (26%); 3 studies (3%) used three perspectives. Among the 90 selected studies, which all included some form of productivity estimation, 47 (52%) also included the intervention's effect on health outcomes and 60 (67%) the effect on healthcare costs or worker's compensation. In 20 studies (22%), neither health outcomes nor healthcare costs were included, making the overall impact of productivity estimation the major element of importance to the study, while 43 studies (48%) did not include an economic evaluation guideline – such as (Cleemput, Van Wilder, Vrijens, Huybrechts, & Ramaekers, 2008; Drummond et al., 2015; Gold et al., 1996; M. Koopmanschap et al., 2005; M. A. Koopmanschap & Rutten, 1996; M. A. Koopmanschap et al., 1995; Oostenbrink et al., 2002; Pauly et al., 2002; Tan et al., 2012) in their methods.

Productivity estimation method

As summarized in Table 4.2, 44 studies (49%) used the HCA, 17 (19%) the FCA, 13 studies (14%) stated productivity in natural units (CEA LT), 7 studies (8%) made use of compensation expenses (CE), 4 studies (4%) estimated productivity changes by analyzing output or revenue data, and 1 study (1%) did not state its method. The remaining 4 studies (4%) all used ad hoc approaches that cannot be reduced to one of the above. For instance, Ektor-Andersen, Ingvarsson, Kullendorff, and Orbaek, 2008 valued lost time by using social security expenditures. Another study (Alamgir et al., 2008) made use of direct-to-indirect cost multipliers: direct costs of an intervention are gathered, after which a multiplier is applied to approximate the indirect productivity costs (or cost savings). However, note that the term "indirect costs" can also indicate additional effects (eg, reductions in informal care, home production, etc). No studies made use of the USP approach. In 18 studies, a second method was applied: most used either a HCA or compensation expenses. Less common were the addition of a (Pauly-Nicholson) multiplier approach, productivity in natural units, and an output-based method. Alamgir et al., 2008 combined three methods: a direct-to-indirect cost multiplier, stating productivity in natural units, and compensation expenses.

Measuring time loss

Only 5 (Dale et al., 2014; Dowd et al., 2010; Evanoff & Kymes, 2010; Lahiri, Latif, & Punnett, 2013; Ødegaard & Roos, 2014) of the 90 studies did not include a measure of lost-time (absenteeism or presenteeism). This is because, in addition to studies with a primary or secondary method that used lost-time

Table 4.2: Characteristics of the included studies (N=90) and the interventions they studied. An extended table is available in an online annex.

Category	Variable	Number of studies
	Europe	50
Continent	North-America	34
	Other continents	6
Perspective	Employer	Stated: 24 Inferred: 19
	Societal	Stated: 37 Inferred: 2
	Health care provider or insurance	Stated: 3 Inferred: 2
	Not inferable	3
Sector / industry	Patient population	21
	Multiple sectors	15
	Health care	15
	Manufacturing & manual labor	8
	Services and administration	7
	Transport	4
	Construction	4
	Education	2
	Military/police	2
	Not stated or not applicable	12
Intervention type	Multiple intervention types	36
	Health promotion	12
	Vaccination	10
	Ergonomics	9
	Return to work	9
	Mental health	6
	Disease prevention & management	4
	Risk assessment	2
	Environment enhancement (eg, air and noise control)	2
Economic evaluation type	CEA	71
	CUA	19
	RCT	35
	BA	16
	DAM	14
	(matched) CBA	7
Study design	Quasi-experiment	6
	PCS	5
	RCS	3
	BA with DAM	2
	HCT	1
	HCA	44
	FCA	17
	CEA LT	13
Productivity method	Compensation expenses	7
	Output based	4
	Other	4
	Not stated	1

BA=(uncontrolled) before–after study; CBA=controlled before–after study;
CE=compensation expenses; CEA=cost-effectiveness analysis;
CEA LT=cost-effectiveness analysis with lost-time in natural units
CUA=cost-utility analysis; DAM=decision-analytic modelling; FCA=friction-cost approach;
HCA=human-capital approach; HCT=historical controlled trial;
glspcs=prospective cohort study; RCS=retrospective cohort study;
RCT=(cluster) randomized controlled trial

(productivity in natural units, HCA, or FCA), several of the ad hoc and output-based methods measured absenteeism or presenteeism. For instance, At'kov, Azarov, Zhukov, Nicoloyannis, and Durand, 2011 measured absenteeism and presenteeism and valued it with an estimate of the contribution of the employee to firm profits (the marginal revenue of labor), and Ektor-Andersen et al., 2008 valued lost time using social security expenditures.

Absenteeism As indicated in Table 4.3, a measure for absenteeism (sick leave, disability leave or other health-related time away from work) was included in 83 studies (92%), 3 of which combined two measurement methods (Adepoju et al., 2014; Dutheil et al., 2008; Evans-Lacko et al., 2016). Of these 83 studies (or 86 measurements), 26 (31%) made use of firm data, 26 (31%) of employee surveys, 10 (12%) used information from previous studies, 9 (11%) insurance data, 3 (4%) simulated or assumed absenteeism, 3 (4%) used data from the healthcare provider, 3 (4%) from national databases, and 1 (1%) each used data from a private research company, compensation data, and an employer survey. Among the 27 absenteeism surveys, a majority (44%) were self-constructed questionnaires, 19% did not state which instrument was used, 11% used the PROductivity and DISease Questionnaire (PRODISQ) although one adapted it, 7% used the Short-Form Health and Labour Questionnaire (SF-HLQ), 7% the net-cost model questionnaire, and 4% each used the TNO survey for cancer and work, the WHO-Health and Productivity Questionnaire (WHO HPQ), and a cost diary. The recall periods of these instruments ranged from 24 hours (Şen et al., 2012) up to 6 months (Noben, Evers, et al., 2015).

The FCA was applied in 17 studies. This requires, in addition to observing days of absenteeism, an estimation of a friction period: the time it requires to replace an employee. A period of 154 days was used in 7 studies (41%), 4 (24%) used 23 weeks (roughly 161 days), 4 studies (24%) did not specify, 1 study (6%) used 92.68 days, and 1 (6%) 10 weeks (roughly 70 days). When a source was stated, it referred to the Koopmanschap et al seminal study (M. A. Koopmanschap et al., 1995) and/or the Dutch Costing Manual (Oostenbrink et al., 2002; Tan et al., 2012), or the Chartered Institute of Personnel and Development (Chartered Institute of Personnel and Development, 2007). However, one of these studies (Goorden et al., 2014) stated the source (M. A. Koopmanschap et al., 1995) but not the friction period length. Finally, Noben, Hoefsmit, et al., 2015 estimated their own friction period using data on job vacancies.

Table 4.3: Inclusion and measurement details of absenteeism, presenteeism, and friction periods by the studies included in this article (N=90).

Category	Variable description	Number of studies
Absenteeism	Measure for sick leave, disability leave, time away from work	83
Source absenteeism	Firm data	26
	Employee survey	26
	Previous studies	10
	Insurance data	9
	Simulated / assumptions	3
	Healthcare provider	3
	National (governmental) databases	3
	Source not stated	3
	Private data gatherer / research company	1
	Compensation data	1
	Employer survey	1
Survey/instrument absenteeism	Non-standardized / custom made survey / study-specific questionnaire	12
	Survey / instrument not stated	5
	PRODISQ	3
	SF-HLQ	2
	Net-cost model questionnaire	2
	TNO survey cancer and work	1
	(sections of) WHO-HPQ	1
	Cost diary	1
Friction period length	154 days	7
	23 weeks	4
	92.68 days	1
	10 weeks	1
Friction period source	Dutch costing manual and/ or Koopmanschap	11
	Estimated by the author	1
	The Chartered Institute of Personnel and Development	1
Presenteeism	Measure for reduced productivity at work/while working	31
Source presenteeism	Employee survey	18
	Simulated / assumptions	7
	Previous studies	6
	Firm data	1
	Employer survey	1
Survey/instrument presenteeism	WHO-HPQ	10
	Non-standardized / custom made survey / study-specific questionnaire	2
	PRODISQ	2
	TiC-P (or SF-HLQ in TiC-P)	2
	QQ	1
	Net-cost model questionnaire	1
	Survey / instrument not stated	1
Presenteeism summary approach (of instruments)	Comparative approach	10
	Not stated	4
	Direct approach	3
	Perceived impairment approach	2
Time unit of absenteeism / presenteeism measurement	Days (unspecified)	49

Hours	27
Other	7
Weeks	3
Time unit not stated	2

PRODISQ=PROductivity and DISease Questionnaire;
SF-HLQ=Short-Form Health and Labour Questionnaire;
TiC-P=Trimbos/ iMTA questionnaire for Costs associated with Psychiatric Illness;
QQ=Quantity and Quality questionnaire;
WHO HPQ=WHO-Health and Productivity Questionnaire

Presenteeism This was not taken into account in 59 studies (66%). Nonetheless, research indicates that both absenteeism and presenteeism are important factors: presenteeism can play a larger role than absenteeism in some diseases and for jobs that are physically and cognitively demanding, whereas absenteeism plays a larger role for jobs with high physical job demands (Jinnett, Schwatka, Tenney, Brockbank, & Newman, 2017; Schultz, Chen, & Edington, 2009; Zhang et al., 2011). A measure for reduced performance at work was included in 31 studies (34%), 2 studies used two measures (Barbosa et al., 2015; Tompa, Dolinschi, & Natale, 2013). Employee surveys were used in 18 studies (58%), 7 studies (23%) simulated presenteeism or assumed that productivity under presenteeism was decreased by a certain percentage based on the literature [eg, a 50% (Tompa, Dolinschi, Alamgir, Sarnocinska-Hart, & Guzman, 2016) or a 25% (Vermeulen et al., 2013) decrease in productivity], 6 studies (19%) used information from previous studies, and 1 study (3%) each used firm data or an employee survey. Of the 19 surveys, 10 studies (53%) used the WHO HPQ, of which Osilla (Osilla et al., 2010) used only part, 2 studies (11%) used PRODISQ, and 1 study (5%) each used the Trimbos/ iMTA questionnaire for Costs associated with Psychiatric Illness (TiC-P), the SF-HLQ in TiC-P, the Quantity and Quality (QQ) questionnaire, and the net-cost model questionnaire (in the employer survey). The remainder used self-constructed surveys (11%) or did not state which survey was used (5%). The range of recall periods for presenteeism instruments was similar to those of absenteeism above.

Regarding the method to estimate presenteeism within these surveys, 10 studies (53%) (those using the WHO HPQ) used a combination of perceived impairment – by measuring self-assessed Likert scale reduced work performance due to sickness – and the comparative approach – by comparing global presenteeism estimates with those of a colleague (Kigozi et al., 2016). No approach was stated in 4 studies (21%), while 3 (16%) used a direct approach, asking for "extra hours needed to catch up on tasks unable to complete in normal working hours" (Arends, Bultmann, van Rhenen, Groen, & van der Klink, 2013; Goorden et al., 2014) or the number of days with decreased efficiency at work (Şen et al., 2012). The perceived change approach was used in 2 studies (11%) that asked employees to give their productivity a 0–1 inefficiency score (Noben, Evers, et al., 2015; Noben et al., 2014).

Data quality Despite their frequent inclusion (83 of 90 studies), several studies reported having difficulties in obtaining good quality absence data (Barbosa et al., 2015; Bernaards, Bosmans, Hildebrandt, van Tulder, & Heymans, 2011; J. Brown et al., 2015; Dewa & Hoch, 2014a; Driessen et al., 2012; Dutheil et al., 2008; Karrholm, Ekholm, Ekholm, Bergroth, & Ekholm, 2008; Nyman, Abraham, Jeffery, & Barleen, 2012; Palumbo, Sikorski, & Liberty, 2013; Pedersen & Arendt, 2014; Roelofs et al., 2010; Smedley et al., 2013; Spekle et al., 2010; Van Buynder et al., 2015; Wendt, Tsai, Bhojani, & Cameron, 2010). Often, data was confounded (eg, by other cause sick leave or holiday leave), or included only part of the relevant absences (eg, only compensated leave). In addition, the time unit of absenteeism and presenteeism measurements influenced the precision of the estimate. Of the 85 studies measuring lost-time, 49 (58%) used unspecified days (no mention of differentiation between partial and full sick leave days), 27 (32%) used hours, 3 (4%) used weeks, and 2 (2%) did not state a time unit. Other approaches were applied in 7 studies (8%): explicitly count part-time absenteeism as full-time absent days (Bernaards et al., 2011), use only full days of absence (Driessen et al., 2012), use both gross (total number of calendar days – partial or full – absent) and net (partial absent days) sick leave (Hlobil et al., 2007; Rebergen, Bruinvels, van Tulder, van der Beek, & van Mechelen, 2009; Seeley et al., 2008); and use short-term absences of 4–9 days (Dutheil et al., 2008), or absence spells >4 weeks (Smedley et al., 2013). A second time unit was used in 4 studies, thereby combining measurements in hours and days (Goorden et al., 2014; Schene et al., 2007; Tompa et al., 2016; van Eijsden, Gerhards, de Bie, & Severens, 2009).

Valuing time loss

A price weight to monetize lost-time measurements was used in 69 studies (77%). The other 21 studies (23%) (Alamgir et al., 2008; Braun, Bambra, Booth, Adetayo, & Milne, 2015; Chan, Shie, Lee, & Lin, 2008; Dale et al., 2014; Dallat, Hunter, Tully, Cairns, & Kee, 2013; Dewa, Hoch, Carmen, Guscott, & Anderson, 2009; Dowd et al., 2010; Evanoff & Kymes, 2010; Hoch & Dewa, 2014; Lahiri et al., 2013; Mukhopadhyay & Wendel, 2013; Nyman et al., 2012; Ødegaard & Roos, 2014; Phillips et al., 2012; Şen et al., 2012; Smedley et al., 2013; Smith, Raymund, Nowalk, Roberts, & Zimmerman, 2010; Taimela et al., 2008; Tallant et al., 2014; Thijs et al., 2012; von Thiele Schwarz & Hasson, 2012)) did not make use of (separate) price weights for monetization because they stated productivity outcomes in natural units, made direct use of compensation expenses, or used an ad hoc approach or output-based method.

Composition Table 4.4 indicates that in about half of the 69 studies that used price weights (45%), these were composed of wages and secondary benefits. A third (32%) used wages only, 4 studies (6%) used compensation expenses (without wage) as the price weight, while 8 (12%) used alternative price weights. One of these used the labor cost and individual operating income per employee (At'kov et al., 2011) and another the replacement cost of a worker (Tompa et al., 2013). Others adjusted wages only for taxes (Norrefalk, Ekholm, Linder, Borg, & Ekholm, 2008) or used a multiplier for non-wage benefits (Osilla et al., 2010). Finally, 2 studies (Dewa & Hoch, 2014a, 2014b) used disability benefits divided by wage, while Ektor-Andersen et al., 2008 used social security expenditures and Karrholm et al., 2008 applied income qualifying for sickness benefit, adjusted for payroll taxes.

Type Regarding the price weight type, 19 studies (28%) used job averages and 18 (26%) applied national averages adjusted for age and gender (recommended by the Dutch costing manual). A generic national or regional price weight was used in 13 studies (19%), 9 (13%) used a worker-specific weight, 3 (4%) the industry or company average, and 1 (1%) a job average with age differentiation. Variations – such as using the mean annual salary per staff member in the NHSL (J. Brown et al., 2015), the study group average (Karrholm et al., 2008; Rebergen et al., 2009), or the national average for blue- and white-collar workers (Norrefalk et al., 2008) – were applied in 4 other studies (6%), while 3 used a second type of price weight: the national age and gender-specific average (Rebergen et al., 2009; van Wier et al., 2013) and a generic national weight (Groeneveld et al., 2011).

Source Data on the price weight was most often taken from previous studies (38%), company administration (17%), national databases (13%), employee self-reports (12%) or insurance data (7%).

Elasticity FCA recommends adjusting calculations for the "elasticity of labor", ie, the less than proportional change in output per unit of labor. In total, 15 studies incorporated an elasticity factor; 14 used a value of 0.8 and 1 used several values: 0.3, 0.55, 0.8. Interestingly, 4 of the 14 studies did not follow the FCA but did include an elasticity factor (Bernaards et al., 2011; Groeneveld et al., 2011; McEachan et al., 2011; Meijster et al., 2011). The source for these elasticity values was Koopmanschap et al (M. A. Koopmanschap et al., 1995) and/or the Dutch Costing Manual (Oostenbrink et al., 2002; Tan et al., 2012), except for Bernaards et al., 2011, who used the Netherlands economic institute as a source.

Table 4.4: Main results on the composition, type and source of price weights, and inclusion of an elasticity of labor to output, for studies that included a valuation of lost time (N=69).

Category	Variable description	Number of studies
Price weight composition	Wage including secondary benefits	31
	Wage	22
	Other	8
	Compensation expenses, no wage	4
Price weight type	Job average	19
	National average, age and gender specific	18
	Generic national or regional price weight	13
	Worker-specific	9
	Other	4
	Industry or company average	3
	Job average with age differentiation	1
Price weight source	Literature / previous studies	26
	Company administration	12
	National databases	9
	Self-reported	8
	Insurance data (private or public)	5
Elasticity of labor to output	Elasticity of 0.8	14
	Elasticity of 0.3, 0.55, 0.8, and 1	1
Elasticity source	Dutch costing manual and/ or Koopmanschap	14
	Netherlands Economic Institute	1

Additional costs and adjustments

Productivity costs of sick employees can also involve additional costs related to compensation mechanisms (eg, costs of replacement workers) or employee turnover. One author included overtime, staffing agency costs, and wages for temporary employees (von Thiele Schwarz & Hasson, 2012), while 4 studies included hiring costs (Seeley et al., 2008; Smith et al., 2010; Squires, Rick, Carroll, & Hillage, 2012; Tompa et al., 2013), such as equating the cost of lost time to the cost of a replacement worker (Seeley et al., 2008; Tompa et al., 2013), including the cost of advertising and recruiting the replacement (Squires et al., 2012), or adding temporary and permanent replacement workers' costs (Smith et al., 2010). The latter was also the only study to include the training costs of replacement workers. Employee turnover rates were included in 6 analyses (Barbosa et al., 2015; Evanoff & Kymes, 2010; Guimaraes, Ribeiro, & Renner, 2012; Lahiri et al., 2013; Meenan et al., 2010; Noben, Hoefsmit, et al., 2015) with the view that improving the working environment or avoiding injuries, disabilities and diseases causes less employees to leave the firm (Mossink & de Greef, 2002). For instance, Barbosa et al., 2015 approximated "voluntary termination costs" through workers' compensation and Lahiri et al., 2013 included the cost of

employee turnover rates by multiplying the change in personnel (by a headcount) by estimated costs of turnover by the company (expressed as a percentage of the job-category wage).

Finally, some studies separately adjusted the value of lost output for a compensation mechanism (apart from their productivity estimation method) or followed the Pauly-Nicholson approach (S. Nicholson et al., 2006) to reflect the fact that absenteeism costs might be higher than the wage. The latter was done in 2 studies (Meenan et al., 2010; Osilla et al., 2010) by including job-multipliers, while 7 (8%) included a separate measure for compensation mechanisms to reflect that these can reduce output losses. Evanoff and Kymes, 2010 adjusted productivity losses to (all) replacement workers, At'kov et al., 2011 to internal replacements, and Palumbo et al., 2013 took only 75% of absences into account to reflect the fact that some production losses are compensated.

Uncertainty and discounting

Because of the large uncertainties involved, it is necessary to subject productivity estimations to a sensitivity analysis to see how results change – for instance by relaxing assumptions (eg, on the percentage used for presenteeism), by using alternative data sources, by including and excluding presenteeism, etc. As indicated in Table 4.5, this was done in 61 studies (68%), while 21 (23%) stated to have discounted productivity estimates, and 53 (59%) mentioned inflation adjustments. Discounting and inflation adjustment is advised if the study length is more than a year.

Table 4.5: Main results on sensitivity analyses, discounting, and inflation of productivity loss by the included studies (N=90).

Category	Variable description	Number of studies
Sensitivity analysis	A sensitivity analysis was performed on the productivity losses	61
Discounting	Productivity costs were discounted	21
Inflation	Productivity costs were inflated	53

4.2.4 Discussion

In the past, some have advocated the view that quick and simple decision-making tools should be sufficient to convince employers of the benefits of occupational health and safety interventions (Burton & World Health Organization, 2010). Whereas this may be practical and responsive to the immediate information

needs of managers in a particular context, from a broader perspective there are good reasons to adhere to stronger scientific standards. First, even within a specific context, managers can implement many OHS programs and need to set priorities. If productivity estimations are insufficiently precise, they can lead to a sub-optimal allocation of available OHS resources, and thus to forgone benefits from other and "better" OHS programs. Second, economic evaluation of OHS programs is not only needed to convince local managers and decision makers, but it is also important to provide an evidence-base for discussions between governments, employers and trade unions. It can help to set out policy priorities within OHS, or facilitate comparisons with broader forms of healthcare and public health (eg, maximizing population health on a cost-per-QALY basis). When cost and effects for different parties are reported separately, it can even help to divide the burdens and benefits of OHS programs, and align incentives for OHS investments by means of subsidies or taxation (Asche & Aven, 2004; Cagno, Micheli, Masi, & Jacinto, 2013; Kankaanpaa, Suhonen, & Valtonen, 2008).

Previous reviews of economic evaluation studies of OHS laid bare substantial methodological shortcomings in their approach to quantifying productivity change: a misplaced reliance on workers' compensation expenses, a large group of studies that did not include presenteeism (H. E. Brown, Gilson, Burton, & Brown, 2011; Chapman, 2012; J. D. Jensen, 2011; Kigozi et al., 2017a; Trogdon, Finkelstein, Hylands, Dellea, & Kamal-Bahl, 2008; van Dongen et al., 2011, 2012; Verbeek et al., 2009), weak(er) study designs, and strong assumptions without controlling for these (Tompa et al., 2009). The authors urged for more precision and standardization in future studies. This systematic review assessed the economic evaluations of OHS published in the preceding decade. In analyzing 90 studies, the overall results sketched a largely similar picture: a large heterogeneity in productivity measurement and valuation/monetization methods, big differences in the precision of the implementation of each method, limited comparability across studies, non-transparent reporting of methods used, exclusion of relevant cost-dimensions related to absenteeism and presenteeism, and limited sensitivity analysis and accounting for the uncertainty embodied in the presented estimates. The approach used seems to be based more on data availability and practical considerations than on scientific reflection of what constitutes the most valid and reliable method in a certain context. For instance, both compensation mechanisms and wage multipliers have been empirically demonstrated to matter (Kigozi et al., 2017a; Krol et al., 2012; Strömberg et al., 2017), but were rarely included. More particularly, improvements are possible through increasing data quality or precision (eg, a lower reliance on compensation or undifferentiated sick leave data confounded with holiday leave), adding essential productivity components (such as presenteeism or replacement and turnover costs), reducing strong assumptions (eg, about presenteeism) or

testing their impact, shortening the recall periods used in surveys, and more adequate adjustment for inflation or discounting. In addition, studies using the HCA could additionally calculate productivity costs through the FCA, since – especially in the case of absences longer than the friction period – the HCA might reflect only the "potential" costs, and overestimate the "actual" losses (Lensberg et al., 2013). To make this possible, estimates for friction periods should be made in other countries since up to now they have only been calculated for the Netherlands and the UK (Kigozi et al., 2017a; M. A. Koopmanschap et al., 1995). Not all these quality improvements of productivity estimation require additional time, data or resource investments: detailed reporting, adjustments to inflation or discounting, relaxing assumptions in sensitivity analyses, making use of job-specific wage multipliers, or using (a range of) elasticity values as proposed by Koopmanschap et al can all greatly increase the quality of cost-effectiveness research in OHS, and could easily be added (eg, in sensitivity analyses next to the main analysis).

All this adds evidence to the observation of previous authors that there is, as of yet, no consensus on the specifics of the measurement and valuation/monetization method of productivity (Brouwer & Koopmanschap, 2005; Furlan et al., 2012; M. Koopmanschap et al., 2005; Krol & Brouwer, 2014; Lofland et al., 2004; S. Nicholson et al., 2006; Pauly et al., 2002; Uegaki et al., 2011; Zhang et al., 2011). Guidelines should be developed, keeping in mind cross-countries differences, for instance in worker's legislation and compensation regulation (Uegaki et al., 2011). Many Dutch studies seem already committed to increased standardization and have taken advantage of the Dutch Costing Manual (Kigozi et al., 2016; Oostenbrink et al., 2002; Tan et al., 2012). Given the close relations between OHS and the firm, this field might prove an ideal setting to advance the broader field of economic evaluation. First, the closer ties with companies in OHS may render objective measures for productivity changes more feasible. Even if only a few employers can be convinced to engage in a study, this can give many opportunities to validate measurement instruments objectively, compare productivity methods against an output-based standard, or further advance research of compensation mechanisms and multipliers. Several authors have noted that the general quality and relevance of economic evaluation in the OHS field falls below the standard in other domains of healthcare (Luyten et al., 2017; P. Nicholson, 2017; Tompa et al., 2009). The OHS community should not fall behind in incorporating state-of-the-art methods into economic evaluations and can even take a lead position experimenting with methods to assess how productivity should be measured.

References

Action, E. J., Heintz, E., Gerber-Grote, A., Ghabri, S., Hamers, F. F., Rupel, V. P., ... Davidson, T., et al. (2016). Is there a european view on health economic evaluations? results from a synopsis of methodological guidelines used in the eunethta partner countries. *Pharmacoeconomics*, *34*(1), 59–76.

Adepoju, O. E., Bolin, J. N., Ohsfeldt, R. L., Phillips, C. D., Zhao, H., Ory, M. G., & Forjuoh, S. N. (2014). Can chronic disease management programs for patients with type 2 diabetes reduce productivity-related indirect costs of the disease? evidence from a randomized controlled trial. *Popul Health Manag*, *17*(2), 112–120.

Alamgir, H., Yu, S., Fast, C., Hennessy, S., Kidd, C., & Yassi, A. (2008). Efficiency of overhead ceiling lifts in reducing musculoskeletal injury among carers working in long-term care institutions. *Injury*, *39*(5), 570–7. doi:10.1016/j.injury.2007.11.420

Ambler, T., & Roberts, J. H. (2008). Assessing marketing performance: Don't settle for a silver metric. *J Market Manag*, *24*(7-8), 733–750. doi:10.1362/026725708X345498

Anders, B., Ommen, O., Pfaff, H., Lüngen, M., Lefering, R., Thüm, S., & Janssen, C. (2013). Direct, indirect, and intangible costs after severe trauma up to occupational reintegration–an empirical analysis of 113 seriously injured patients. *GMS Psycho-Social-Medicine*, *10*.

Apostolakis, G. E. (2004). How useful is quantitative risk assessment? *Risk Anal*, *24*(3), 515–520. doi:10.1111/j.0272-4332.2004.00455.x

Arends, I., Bultmann, U., van Rhenen, W., Groen, H., & van der Klink, J. J. (2013). Economic evaluation of a problem solving intervention to prevent recurrent sickness absence in workers with common mental disorders. *PLoS One*, *8*(8), e71937. doi:10.1371/journal.pone.0071937

Asche, F., & Aven, T. (2004). On the economic value of safety. *Risk Decis Pol*, *9*(3), 253–267.

At'kov, O. Y., Azarov, A. V., Zhukov, D. A., Nicoloyannis, N., & Durand, L. (2011). Influenza vaccination in healthy working adults in russia: Observational study of effectiveness and return on investment for the employer. *Appl Health Econ Health Policy*, *9*(2), 89–99. doi:10.2165/11538680-000000000-00000

Aven, T. (2010). On how to define, understand and describe risk. *Reliab Eng Syst Saf*, *95*(6), 623–631. doi:10.1016/j.ress.2010.01.011

Aven, T., & Zio, E. (2014). Foundational issues in risk assessment and risk management. *Risk Anal*, *34*(7), 1164–1172. doi:10.1111/risa.12132

Barbosa, C., Bray, J. W., Dowd, W. N., Mills, M. J., Moen, P., Wipfli, B., ... Kelly, E. L. (2015). Return on investment of a work-family intervention:

Evidence from the work, family, and health network. *J Occup Environ Med*, *57*(9), 943–51. doi:10.1097/JOM.0000000000000520

Barling, J., & Hutchinson, I. (2000). Commitment vs. control-based safety practices, safety reputation, and perceived safety climate. *Can J Adm Sci*, *17*(1), 76–84. doi:10.1111/j.1936-4490.2000.tb00208.x

Bernaards, C. M., Bosmans, J. E., Hildebrandt, V. H., van Tulder, M. W., & Heymans, M. W. (2011). The cost-effectiveness of a lifestyle physical activity intervention in addition to a work style intervention on recovery from neck and upper limb symptoms and pain reduction in computer workers. *Occup Environ Med*, *68*(4), 265–72. doi:10.1136/oem.2008.045450

Braun, T., Bambra, C., Booth, M., Adetayo, K., & Milne, E. (2015). Better health at work? an evaluation of the effects and cost-benefits of a structured workplace health improvement programme in reducing sickness absence. *J Public Health-Uk*, *37*(1), 138–142.

Briggs, A., Sculpher, M., & Claxton, K. (2006). *Decision modelling for health economic evaluation*. Oxford: Oxford Univ Pr.

Brouwer, W. B., & Koopmanschap, M. A. (2005). The friction-cost method : Replacement for nothing and leisure for free? *PharmacoEconomics*, *23*(2), 105–11. doi:10.2165/00019053-200523020-00002

Brown, H. E., Gilson, N. D., Burton, N. W., & Brown, W. J. (2011). Does physical activity impact on presenteeism and other indicators of workplace well-being? *Sports Med*, *41*(3), 249–62. doi:10.2165/11539180-000000000-00000

Brown, J., Mackay, D., Demou, E., Craig, J., Frank, J., & Macdonald, E. B. (2015). The easy (early access to support for you) sickness absence service: A four-year evaluation of the impact on absenteeism. *Scand J Work Environ Health*, *41*(2), 204–15. doi:10.5271/sjweh.3480

Brymer, R. A., Molloy, J. C., & Gilbert, B. A. (2014). Human capital pipelines: Competitive implications of repeated interorganizational hiring. *J Manage*, *40*(2), 483–508. doi:10.1177/0149206313516797

Buchan, J., & Campbell, J. (2013). Challenges posed by the global crisis in the health workforce. *BMJ*, *347:f6201.* doi:10.1136/bmj.f6201

Burdorf, A. (2007). Economic evaluation in occupational health—its goals, challenges, and opportunities. *Scand J Work Environ Health*, 161–164. doi:10.5271/sjweh.1127

Burton, J., & World Health Organization. (2010). *Who healthy workplace framework and model: Background and supporting literature and practices* (Report No. Report No.: 9241500247). World Health Organization.

Cagno, E., Micheli, G. J. L., Masi, D., & Jacinto, C. (2013). Economic evaluation of osh and its way to smes: A constructive review. *Safety Sci*, *53*, 134–152. doi:10.1016/j.ssci.2012.08.016

Carroll, C., Rick, J., Pilgrim, H., Cameron, J., & Hillage, J. (2010). Workplace involvement improves return to work rates among employees with back pain on long-term sick leave: A systematic review of the effectiveness and cost-effectiveness of interventions. *Disabil Rehabil*, *32*(8), 607–21. doi:10.3109/09638280903186301

Chan, A. L., Shie, H. J., Lee, Y. J., & Lin, S. J. (2008). The evaluation of free influenza vaccination in health care workers in a medical center in taiwan. *Pharm World Sci*, *30*(1), 39–43. doi:10.1007/s11096-007-9137-8

Chandra, A., & Skinner, J. (2012). Technology growth and expenditure growth in health care. *J Econ Lit*, *50*(3), 645–80. doi:10.1257/jel.50.3.645

Chapman, L. S. (2012). Meta-evaluation of worksite health promotion economic return studies: 2012 update. *Am J Health Promot*, *26*(4), TAHP1–TAHP12. doi:10.4278/ajhp.26.4.tahp

Chartered Institute of Personnel and Development. (2007). *Recruitment, retention and turnover: Annual survey report 2007*. CIPD.

Chowdhury, M. F., & Tanim, T. R. (2016). Industrial accidents in bangladesh apparel manufacturing sector: An analysis of the two most deadliest accidents in history. *Asian J Soc Sci Manag Stud*, *3*(2), 115–126. doi:10.20448/journal.500/2016.3.2/500.2.115.126

Clarke, V. (2013). Honesty and integrity: Is this the edge businesses need in a stagnant economy? Web Page.

Cleemput, I., Van Wilder, P., Vrijens, F., Huybrechts, M., & Ramaekers, D. (2008). *Guidelines for pharmacoeconomic evaluations in belgium. health technology assessment (hta)* (Report No. KCE Reports 78C (D/2008/10.273/27)). Federal Health Care Knowledge Centre (KCE).

Colosio, C., Mandic-Rajcevic, S., Godderis, L., van der Laan, G., Hulshof, C., & van Dijk, F. (2017). Workers' health surveillance: Implementation of the directive 89/391/eec in europe. *Occupational Medicine*, *67*(7), 574–578.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., & Medical Research Council, G. (2008). Developing and evaluating complex interventions: The new medical research council guidance. *BMJ*, *337*, a1655. doi:10.1136/bmj.a1655

Cravens, K., Oliver, E. G., & Ramamoorti, S. (2003). The reputation index: Measuring and managing corporate reputation. *Eur Manag J*, *21*(2), 201–212.

Dale, A. M., Gardner, B. T., Zeringue, A., Werner, R., Franzblau, A., & Evanoff, B. (2014). The effectiveness of post-offer pre-placement nerve conduction screening for carpal tunnel syndrome. *J Occup Environ Med*, *56*(8), 840–7. doi:10.1097/JOM.0000000000000185

Dallat, M. A., Hunter, R. F., Tully, M. A., Cairns, K. J., & Kee, F. (2013). A lesson in business: Cost-effectiveness analysis of a novel financial incentive

intervention for increasing physical activity in the workplace. *BMC Public Health*, *13*, 953. doi:10.1186/1471-2458-13-953

De Loecker, J., & Goldberg, P. K. (2014). Firm performance in a global market. *Annu Rev Econ*, *6*(1), 201–227.

Dekker, S. W. (2014). The bureaucratization of safety. *Safety Sci*, *70*, 348–357. doi:10.1016/j.ssci.2014.07.015

Dewa, C. S., & Hoch, J. S. (2014a). Estimating the net benefit of a specialized return-to-work program for workers on short-term disability related to a mental disorder: An example exploring investment in collaborative care. *J Occup Environ Med*, *56*(6), 628–31. doi:10.1097/JOM.0000000000000157

Dewa, C. S., & Hoch, J. S. (2014b). When could a stigma program to address mental illness in the workplace break even? *Can J Psychiatry*, *59*(10), S34–S39.

Dewa, C. S., Hoch, J. S., Carmen, G., Guscott, R., & Anderson, C. (2009). Cost, effectiveness, and cost-effectiveness of a collaborative mental health care program for people receiving short-term disability benefits for psychiatric disorders. *Can J Psychiatry*, *54*(6), 379–88. doi:10.1177/070674370905400605

Dowd, B., McGrail, M., Lohman, W. H., Sheasby, B., O'Connor, H., Calasanz, M., ... Parente, S. (2010). The economic impact of a disability prevention program. *J Occup Environ Med*, *52*(1), 15–21. doi:10.1097/JOM.0b013e3181c5c358

Driessen, M., Bosmans, J., Proper, K., Anema, J., Bongers, P., & van der Beek, A. (2012). The economic evaluation of a participatory ergonomics programme to prevent low back and neck pain. *Work*, *41 Suppl 1*, 2315–20. doi:10.3233/WOR-2012-0458-2315

Drummond, M. F., & McGuire, A. (2001). *Economic evaluation in health care: Merging theory with practice*. Oxford: Oxford University Press.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Univ Pr.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). Methods for the economic evaluation of health care programmes. Book. Oxford: Oxford Univ Pr.

Dutheil, F., Delaire, P., Boudet, G., Rouffiac, K., Djeriri, K., Souweine, B., & Chamoux, A. (2008). [cost/effectiveness comparison of the vaccine campaign and reduction of sick leave, after vaccination against influenza among the clermont-ferrand university hospital staff]. *Med Maladies Infect*, *38*(11), 567–573. doi:10.1016/j.medmal.2008.09.019

Edwards, R. T., Charles, J. M., & Lloyd-Williams, H. (2013). Public health economics: A systematic review of guidance for the economic evaluation

of public health interventions and discussion of key methodological issues. *BMC Publ Health*, *13*(1), 1001. doi:10.1186/1471-2458-13-1001

Ektor-Andersen, J., Ingvarsson, E., Kullendorff, M., & Orbaek, P. (2008). High cost-benefit of early team-based biomedical and cognitive-behaviour intervention for long-term pain-related sickness absence. *J Rehabil Med*, *40*(1), 1–8. doi:10.2340/16501977-0127

Elsler, D., Treutlein, D., Rydlewska, I., Frusteri, L., Kruger, H., Veerman, T., ... Taylor, T. N. (2010). A review of case studies evaluating economic incentives to promote occupational safety and health. *Scand J Work Environ Health*, *36*(4), 289–98.

Esler, D., Eeckelaert, L., Knight, A., Treutlein, D., Pecillo, M., Elo-Schäfer, J., ... Dontas, S. (2010). *Economic incentives to improve occupational safety and health: A review from the european perspective*.

Evangelinos, K., Fotiadis, S., Skouloudis, A., Khan, N., Konstandakopoulou, F., Nikolaou, I., & Lundy, S. (2018). Occupational health and safety disclosures in sustainability reports: An overview of trends among corporate leaders. *Corp Soc Resp Env Ma.* doi:10.1002/csr.1512

Evanoff, B., & Kymes, S. (2010). Modeling the cost-benefit of nerve conduction studies in pre-employment screening for carpal tunnel syndrome. *Scand J Work Environ Health*, *36*(4), 299–304.

Evans-Lacko, S., Koeser, L., Knapp, M., Longhitano, C., Zohar, J., & Kuhn, K. (2016). Evaluating the economic impact of screening and treatment for depression in the workplace. *Eur Neuropsychopharmacol*, *26*(6), 1004–13. doi:10.1016/j.euroneuro.2016.03.005

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. *Int J Technol Assess Health Care*, *21*(2), 240–5.

Fleurbaey, M., & Abi-Rafeh, R. (2016). The use of distributional weights in benefit–cost analysis: Insights from welfare economics. *Review of Environmental Economics and Policy*, *10*(2), 286–307.

Furlan, A. D., Gnam, W. H., Carnide, N., Irvin, E., Amick, 3., B. C., DeRango, K., ... Bultmann, U. (2012). Systematic review of intervention practices for depression in the workplace. *J Occup Rehabil*, *22*(3), 312–21. doi:10.1007/s10926-011-9340-2

Gagné, M., Forest, J., Gilbert, M.-H., Aubé, C., Morin, E., & Malorni, A. (2010). The motivation at work scale: Validation evidence in two languages. *Educ Psychol Meas*, *70*(4), 628–646. doi:10.1177/0013164409355698

Gold, M., Siegel, J., Russell, L., Weinstein, M., & editors. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford Universitary Press.

Goorden, M., Vlasveld, M. C., Anema, J. R., van Mechelen, W., Beekman, A. T., Hoedeman, R., … Hakkaart-van Roijen, L. (2014). Cost-utility analysis of a collaborative care intervention for major depressive disorder in an occupational healthcare setting. *J Occup Rehabil*, *24*(3), 555–62. doi:10.1007/s10926-013-9483-4

Groeneveld, I. F., van Wier, M. F., Proper, K. I., Bosmans, J. E., van Mechelen, W., & van der Beek, A. J. (2011). Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med*, *53*(6), 610–617. doi:10.1097/JOM.0b013e31821b9c24

Guimaraes, L. B., Ribeiro, J. L., & Renner, J. S. (2012). Cost-benefit analysis of a socio-technical intervention in a brazilian footwear company. *Appl Ergon*, *43*(5), 948–57. doi:10.1016/j.apergo.2012.01.003

Hak, E., Knol, L. M., Wilschut, J. C., & Postma, M. J. (2010). [influenza vaccination of hospital healthcare staff from the perspective of the employer: A positive balance]. *Ned Tijdschr Geneeskd*, *154*, A1188.

Hamberg-van Reenen, H. H., Proper, K. I., & van den Berg, M. (2012). Worksite mental health interventions: A systematic review of economic evaluations. *Occup Environ Med*, *69*(11), 837–45. doi:10.1136/oemed-2012-100668

Hanly, P., Ortega, M. O., Pearce, A., Soerjomataram, I., & Sharp, L. (2020). Advances in the methodological approach to friction period estimation: A european perspective. *Social Science & Medicine*, *264*, 113289.

Hassink, W., & van den Berg, B. (2017). Methods and instruments for the estimation of production changes in economic evaluations. In D. Razzouk (Ed.), *Mental health economics* (Chap. Methods and Instruments for the Estimation of Production Changes in Economic Evaluations, pp. 431–440). Springer, Cham.

Heuvel, S., Zwaan, L., Dam, L. v., Oude Hengel, K., Eekhout, I., van Emmerik, M., … Wilhelm, C. (2017). *Estimating the costs of work-related accidents and ill-health: An analysis of european data sources* (Report No. 9292409972). European Agency for Safety and Health at Work (EU-OSHA).

Hlobil, H., Uegaki, K., Staal, J. B., de Bruyne, M. C., Smid, T., & van Mechelen, W. (2007). Substantial sick-leave costs savings due to a graded activity intervention for workers with non-specific sub-acute low back pain. *Eur Spine J*, *16*(7), 919–924. doi:10.1007/s00586-006-0283-9

Hoch, J. S., & Dewa, C. S. (2014). Advantages of the net benefit regression framework for economic evaluations of interventions in the workplace: A case study of the cost-effectiveness of a collaborative mental health care program for people receiving short-term disability benefits for psychiatric disorders. *J Occup Environ Med*, *56*(4), 441–5. doi:10.1097/JOM.0000000000000130

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., ... Loder, E. (2013). Consolidated health economic evaluation reporting standards (cheers) statement. *The European Journal of Health Economics*, *14*(3), 367–372.

Ison, T. G. (1998). Worker's compensation systems. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th edition). Geneva: International Labour Organization.

Jeanrenaud, C., & Pellegrini, S. (2007). Valuing intangible costs of alcohol dependence: A contingent valuation study. *Rev Econ Polit*, *117*(5), 813–825. doi:10.3917/redp.175.0813

Jensen, J. D. (2011). Can worksite nutritional interventions improve productivity and firm profitability? a literature review. *Perspect Public Health*, *131*(4), 184–92. doi:10.1177/1757913911408263

Jinnett, K., Schwatka, N., Tenney, L., Brockbank, C. V., & Newman, L. S. (2017). Chronic conditions, workplace safety, and job demands contribute to absenteeism and job performance. *Health Aff (Millwood)*, *36*(2), 237–244. doi:10.1377/hlthaff.2016.1151

Johannesson, M., & Karlsson, G. (1997). The friction cost method: A comment. *J Health Econ*, *16*(2), 249–55, discussion 257–9. doi:http://dx.doi.org/10.1016/S0167-6296(97)00006-4

Jones, C. (2018). *The economics of presenteeism in the context of rheumatoid arthritis, ankylosing spondylitis and psoriatic arthritis* (Thesis).

Kankaanpaa, E., Suhonen, A., & Valtonen, H. (2008). Promoting prevention with economic arguments - the case of finnish occupational health services. *BMC Public Health*, *8*, 130. doi:10.1186/1471-2458-8-130

Karrholm, J., Ekholm, K., Ekholm, J., Bergroth, A., & Ekholm, K. S. (2008). Systematic co-operation between employer, occupational health service and social insurance office: A 6-year follow-up of vocational rehabilitation for people on sick-leave, including economic benefits. *J Rehabil Med*, *40*(8), 628–36. doi:10.2340/16501977-0233

Kigozi, J., Jowett, S., Lewis, M., Barton, P., & Coast, J. (2016). Estimating productivity costs using the friction cost approach in practice: A systematic review. *Eur J Health Econ*, *17*(1), 31–44. doi:10.1007/s10198-014-0652-y

Kigozi, J., Jowett, S., Lewis, M., Barton, P., & Coast, J. (2017a). The estimation and inclusion of presenteeism costs in applied economic evaluation: A systematic review. *Value Health*, *20*(3), 496–506. doi:10.1016/j.jval.2016.12.006

Koopmanschap, M., Burdorf, A., Jacob, K., Meerding, W. J., Brouwer, W., & Severens, H. (2005). Measuring productivity changes in economic evaluation: Setting the research agenda. *PharmacoEconomics*, *23*(1), 47–54. doi:10.2165/00019053-200523010-00004

Koopmanschap, M. A., & Rutten, F. F. (1996). A practical guide for calculating indirect costs of disease. *PharmacoEconomics*, *10*(5), 460–6.

Koopmanschap, M. A., Rutten, F. F., van Ineveld, B. M., & van Roijen, L. (1995). The friction cost method for measuring indirect costs of disease. *J Health Econ*, *14*(2), 171–89. doi:10.1016/0167-6296(94)00044-5

Koopmanschap, M. A., Rutten, F. F. H., van Ineveld, B. M., & van Roijen, L. (1997). Reply to johanneson's and karlsson's comment. *J Health Econ*, *16*(2), 257–259. doi:10.1016/s0167-6296(96)00524-3

Krol, M., & Brouwer, W. (2014). How to estimate productivity costs in economic evaluations. *Pharmacoeconomics*, *32*(4), 335–44. doi:10.1007/s40273-014-0132-3

Krol, M., Brouwer, W. B., Severens, J. L., Kaper, J., & Evers, S. M. (2012). Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects. *Soc Sci Med*, *75*(11), 1981–8. doi:10.1016/j.socscimed.2012.07.012

Krol, M., Stolk, E., & Brouwer, W. (2014). Predicting productivity based on eq-5d: An explorative study. *Eur J Health Econ*, *15*(5), 465–75. doi:10.1007/s10198-013-0487-y

Krol, M., Brouwer, W., & Rutten, F. (2013). Productivity costs in economic evaluations: Past, present, future. *PharmacoEconomics*, *31*(7), 537–549. doi:10.1007/s40273-013-0056-3

Kucherov, D., & Zavyalova, E. (2012). Hrd practices and talent management in the companies with the employer brand. *European Journal of Training and Development*, *36*(1), 86–104. doi:10.1108/03090591211192647

Lahiri, S., Latif, S., & Punnett, L. (2013). An economic analysis of a safe resident handling program in nursing homes. *Am J Ind Med*, *56*(4), 469–78. doi:10.1002/ajim.22139

Lalloo, D., Demou, E., Smedley, J., Madan, I., Asanati, K., & Macdonald, E. B. (2018). Current research priorities for uk occupational physicians and occupational health researchers: A modified delphi study. *Occup Environ Med*, oemed-2018–105114. doi:10.1136/oemed-2018-105114

Lensberg, B. R., Drummond, M. F., Danchenko, N., Despiegel, N., & Francois, C. (2013). Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. *Clinicoecon Outcomes Res*, *5*, 565–73. doi:10.2147/CEOR.S44866

Lofland, J. H., Pizzi, L., & Frick, K. D. (2004). A review of health-related workplace productivity loss instruments. *Pharmacoeconomics*, *22*(3), 165–84. doi:10.2165/00019053-200422030-00003

Lucchini, R. G., & London, L. (2014). Global occupational health: Current challenges and the need for urgent action. *Ann Glob Health*, *80*(4), 251–6.

Luyten, J., Steel, J., & Godderis, L. (2017). Economic evaluation of occupational health services: Necessary, challenging and promising. *Occup Environ Med*, *74*(11), 847–848. doi:10.1136/oemed-2017-104495

Mattke, S., Balakrishnan, A., Bergamo, G., & Newberry, S. J. (2007). A review of methods to measure health-related productivity loss. *Am J Manag Care*, *13*(4), 211–7.

McCurry, J. (2017). Japanese woman 'dies from overwork' after logging 159 hours of overtime in a month. Newspaper Article.

McEachan, R. R., Lawton, R. J., Jackson, C., Conner, M., Meads, D. M., & West, R. M. (2011). Testing a workplace physical activity intervention: A cluster randomized controlled trial. *Int J Behav Nutr Phys Act*, *8*, 29. doi:10.1186/1479-5868-8-29

Meenan, R. T., Vogt, T. M., Williams, A. E., Stevens, V. J., Albright, C. L., & Nigg, C. (2010). Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in hawaii. *J Occup Environ Med*, *52*(1), S8–S13. doi:10.1097/JOM.0b013e3181c81af9

Meijster, T., van Duuren-Stuurman, B., Heederik, D., Houba, R., Koningsveld, E., Warren, N., & Tielemans, E. (2011). Cost-benefit analysis in occupational health: A comparison of intervention scenarios for occupational asthma and rhinitis among bakery workers. *Occup Environ Med*, *68*(10), 739–745. doi:10.1136/oem.2011.064709

Miller, P., & Haslam, C. (2009). Why employers spend money on employee health: Interviews with occupational health and safety professionals from british industry. *Safety Sci*, *47*(2), 163–169. doi:10.1016/j.ssci.2008.04.001

Mossink, J., & de Greef, M. (2002). *Inventory of socioeconomic costs of work accidents* (Report No. 9295007670). Office for Official Publications of the European Communities.

Mukhopadhyay, S., & Wendel, J. (2013). Evaluating an employee wellness program. *Int J Health Care Finance Econ*, *13*(3-4), 173–99. doi:10.1007/s10754-013-9127-4

Neuhaus, M., Eakin, E. G., Straker, L., Owen, N., Dunstan, D. W., Reid, N., & Healy, G. N. (2014). Reducing occupational sedentary time: A systematic review and meta-analysis of evidence on activity-permissive workstations. *Obes Rev*, *15*(10), 822–38. doi:10.1111/obr.12201

Nicholson, P. (2017). *Occupational health: The value proposition*. Society of Occupational Medicine.

Nicholson, S., Pauly, M. V., Polsky, D., Sharda, C., Szrek, H., & Berger, M. L. (2006). Measuring the effects of work loss on productivity with team production. *Health Econ*, *15*(2), 111–23. doi:10.1002/hec.1052

Niven, K. J. (2002). A review of the application of health economics to health and safety in healthcare. *Health Policy*, *61*(3), 291–304. doi:10.1016/S0168-8510(01)00224-X

Noben, C., Evers, S., Nieuwenhuijsen, K., Ketelaar, S., Gartner, F., Sluiter, J., & Smit, F. (2015). Protecting and promoting mental health of nurses in the hospital setting: Is it cost-effective from an employer's perspective? *Int J Occup Med Environ Health*, *28*(5), 891–900. doi:10.13075/ijomeh.1896.00465

Noben, C., Hoefsmit, N., Evers, S., de Rijk, A., Houkes, I., & Nijhuis, F. (2015). Economic evaluation of a new organizational rtw intervention to improve cooperation between sick-listed employees and their supervisors a field study. *J Occup Environ Med*, *57*(11), 1170–1177. doi:10.1097/Jom.0000000000000566

Noben, C., Smit, F., Nieuwenhuijsen, K., Ketelaar, S., Gärtner, F., Boon, B., ... Evers, S. (2014). Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud*, *51*(10), 1321–1331. doi:10.1016/j.ijnurstu.2014.01.017

Norrefalk, J. R., Ekholm, K., Linder, J., Borg, K., & Ekholm, J. (2008). Evaluation of a multiprofessional rehabilitation programme for persistent musculoskeletal-related pain: Economic benefits of return to work. *J Rehabil Med*, *40*(1), 15–22. doi:10.2340/16501977-0131

Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: Evidence from the university of minnesota. *Med Care*, *50*(9), 772–8. doi:10.1097/MLR.0b013e31825a8b1f

Ødegaard, F., & Roos, P. (2014). Measuring the contribution of workers' health and psychosocial work-environment on production efficiency. *Prod Oper Manag*, *23*(12), 2191–2208. doi:10.1111/poms.12242

Olsen, J. A., & Richardson, J. (1999). Production gains from health care: What should be included in cost-effectiveness analyses? *Soc Sci Med*, *49*(1), 17–26. doi:10.1016/s0277-9536(99)00116-1

Olsen, O., Albertsen, K., Nielsen, M. L., Poulsen, K. B., Gron, S. M. F., & Brunnberg, H. L. (2008). Workplace restructurings in intervention studies – a challenge for design, analysis and interpretation. *Bmc Med Res Methodol*, *8*(1), 39. doi:10.1186/1471-2288-8-39

Oostenbrink, J. B., Koopmanschap, M. A., & Rutten, F. F. (2002). Standardisation of costs: The dutch manual for costing in economic evaluations. *Pharmacoeconomics*, *20*(7), 443–54. doi:10.2165/00019053-200220070-00002

Organisation for Economic Co-operation and Development. (2001). *Measuring productivity: Measurement of aggregate and industry-level productivity growth: Oecd manual*. Paris: Organisation for Economic Co-operation and Development.

Osilla, K. C., dela Cruz, E., Miles, J. N., Zellmer, S., Watkins, K., Larimer, M. E., & Marlatt, G. A. (2010). Exploring productivity outcomes from a brief intervention for at-risk drinking in an employee assistance program. *Addict Behav*, *35*(3), 194–200. doi:10.1016/j.addbeh.2009.10.001

Osilla, K. C., Van Busum, K., Schnyer, C., Larkin, J. W., Eibner, C., & Mattke, S. (2012). Systematic review of the impact of worksite wellness programs. *Am J Manag Care*, *18*(2), e68–81.

Palumbo, M. V., Sikorski, E. A., & Liberty, B. C. (2013). Exploring the cost-effectiveness of unit-based health promotion activities for nurses. *Workplace Health Saf*, *61*(12), 514–520. doi:10.3928/21650799-20131206-01

Paté-Cornell, E. (2012). On "black swans" and "perfect storms": Risk analysis and management when statistics are not enough. *Risk Anal*, *32*(11), 1823–1833. doi:10.1111/j.1539-6924.2011.01787.x

Pauly, M. V., Nicholson, S., Xu, J., Polsky, D., Danzon, P. M., Murray, J. F., & Berger, M. L. (2002). A general model of the impact of absenteeism on employers and employees. *Health Econ*, *11*(3), 221–31. doi:10.1002/hec.648

Payne, K., McAllister, M., & Davies, L. M. (2013). Valuing the economic benefits of complex interventions: When maximising health is not sufficient. *Health Econ*, *22*(3), 258–271. doi:10.1002/hec.2795

Pedersen, M. S., & Arendt, J. N. (2014). Bargaining for health: A case study of a collective agreement-based health program for manual workers. *J Health Econ*, *37*(1), 123–36. doi:10.1016/j.jhealeco.2014.06.004

Pelletier, K. R. (2009). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update vii 2004-2008. *J Occup Environ Med*, *51*(7), 822–37. doi:10.1097/JOM.0b013e3181a7de5a

Pelletier, K. R. (2011). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update viii 2008 to 2010. *J Occup Environ Med*, *53*(11), 1310–31. doi:10.1097/JOM.0b013e3182337748

Phillips, C. J., Phillips Nee Buck, R., Main, C. J., Watson, P. J., Davies, S., Farr, A., ... Hale, J. (2012). The cost effectiveness of nhs physiotherapy support for occupational health (oh) services. *BMC Musculoskelet Disord*, *13*, 29. doi:10.1186/1471-2474-13-29

Piko, B. F. (2006). Burnout, role conflict, job satisfaction and psychosocial health among hungarian health care staff: A questionnaire survey. *Int J Nurs Stud*, *43*(3), 311–318. doi:10.1016/j.ijnurstu.2005.05.003

Ponsonby, W. (2017). Global occupational health. *Occup Med (Lond)*, *67*(5), 331–333. doi:10.1093/occmed/kqx068

Prasad, M., Wahlqvist, P., Shikiar, R., & Shih, Y. C. (2004). A review of self-report instruments measuring health-related work productivity: A patient-reported outcomes perspective. *PharmacoEconomics*, *22*(4), 225–44.

Rebergen, D. S., Bruinvels, D. J., van Tulder, M. W., van der Beek, A. J., & van Mechelen, W. (2009). Cost-effectiveness of guideline-based care for workers with mental health problems. *J Occup Environ Med*, *51*(3), 313–22. doi:10.1097/JOM.0b013e3181990d8e

Roelofs, P. D., Bierma-Zeinstra, S. M., van Poppel, M. N., van Mechelen, W., Koes, B. W., & van Tulder, M. W. (2010). Cost-effectiveness of lumbar supports for home care workers with recurrent low back pain: An economic evaluation alongside a randomized-controlled trial. *Spine (Phila Pa 1976)*, *35*(26), E1619–26. doi:10.1097/BRS.0b013e3181cf7244

Rosen, H. S., & Gayer, T. (2010). *Public finance* (9th edition). New York: Mc Graw Hill.

Sancroft. (2017). *Maximising the value of occupational health and safety and workplace wellness reporting for a global workforce: A practical guide for internationally operating employers*. International SOS Foundation.

Schene, A. H., Koeter, M. W., Kikkert, M. J., Swinkels, J. A., & McCrone, P. (2007). Adjuvant occupational therapy for work-related major depression works: Randomized trial including economic evaluation. *Psychol Med*, *37*(3), 351–62. doi:10.1017/S0033291706009366

Schultz, A. B., Chen, C. Y., & Edington, D. W. (2009). The cost and impact of health conditions on presenteeism to employers: A review of the literature. *PharmacoEconomics*, *27*(5), 365–78. doi:10.2165/00019053-200927050-00002

Seeley, P. A., Marklin, R. W., Usher, D., & Yager, J. (2008). Case study report: Business case for implementing battery-powered tools for direct-bury line workers at an electric power utility. *J Occup Environ Hyg*, *5*(8), D86–D91.

Şen, C., Karli, N., Zarifoğlu, M., Pala, K., Seferoğlu, M., & Özkaya, G. (2012). Impact of migraine treatment on economic burden and workplace productivity among migraineurs employed in schools: A prospective study. *J Neurol Sci*, *28*(4), 465–474.

Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? implications for health economic evaluation. *BMJ*, *336*(7656), 1281. doi:10.1136/bmj.39569.510521.AD

Shrestha, N., Kukkonen-Harjula, K. T., Verbeek, J. H., Ijaz, S., Hermans, V., & Bhaumik, S. (2016). Workplace interventions for reducing sitting at work. *Cochrane Database Syst Rev*, (3). doi:10.1002/14651858.CD010912.pub3

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. doi:10.1126/science.3563507

Smallman, C., & John, G. (2001). British directors perspectives on the impact of health and safety on corporate performance. *Safety Sci*, *38*(3), 227–239. doi:10.1016/S0925-7535(01)00003-0

Smedley, J., Harris, E. C., Cox, V., Ntani, G., & Coggon, D. (2013). Evaluation of a case management service to reduce sickness absence. *Occup Med*, *63*(2), 89–95. doi:10.1093/occmed/kqs223

Smith, K. J., Raymund, M., Nowalk, M. P., Roberts, M. S., & Zimmerman, R. K. (2010). Cost-effectiveness of pneumococcal polysaccharide vaccine among healthcare workers during an influenza pandemic. *Am J Manag Care*, *16*(3), 200–6.

Spekle, E. M., Heinrich, J., Hoozemans, M. J., Blatter, B. M., van der Beek, A. J., van Dieen, J. H., & van Tulder, M. W. (2010). The cost-effectiveness of the rsi quickscan intervention programme for computer workers: Results of an economic evaluation alongside a randomised controlled trial. *BMC Musculoskelet Disord*, *11*, 259. doi:10.1186/1471-2474-11-259

Squires, H., Rick, J., Carroll, C., & Hillage, J. (2012). Cost-effectiveness of interventions to return employees to work following long-term sickness absence due to musculoskeletal disorders. *J Public Health-Uk*, *34*(1), 115–124. doi:10.1093/pubmed/fdr057

Steel, J., Godderis, L., & Luyten, J. (2018a). Methodological challenges in the economic evaluation of occupational health and safety programmes. *International journal of environmental research and public health*, *15*(11), 2606. doi:10.3390/ijerph15112606

Steel, J., Godderis, L., & Luyten, J. (2018b). Productivity estimation in economic evaluations of occupational health and safety interventions: A systematic review. *Scand J Work Environ Health*, *44*(5). doi:10.5271/sjweh.3715

Steel, J., Luyten, J., & Godderis, L. (2018). *Occupational health: The global evidence and value*. Society of Occupational Medicine (SOM); London.

Steen, A., & Welch, D. (2011). Are accounting metrics applicable to human resources? the case of return on valuing assignments. *Australasian Accounting, Business and Finance Journal*, *5*(3), 57–72.

Strömberg, C., Aboagye, E., Hagberg, J., Bergstrom, G., & Lohela-Karlsson, M. (2017). Estimating the effect and economic impact of absenteeism, presenteeism, and work environment-related problems on reductions in productivity from a managerial perspective. *Value Health*, *20*(8), 1058–1064. doi:10.1016/j.jval.2017.05.008

Syverson, C. (2011). What determines productivity? *J Econ Lit*, *49*(2), 326–365. doi:10.1257/jel.49.2.326

Taimela, S., Justen, S., Aronen, P., Sintonen, H., Laara, E., Malmivaara, A., ... Aro, T. (2008). An occupational health intervention programme for workers at high risk for sickness absence. cost effectiveness analysis based

on a randomised controlled trial. *Occup Environ Med*, *65*(4), 242–248. doi:10.1136/oem.2007.033167

Tallant, A., Porter, C. K., Putnam, S. D., Tribble, D. R., Hooper, T. I., & Riddle, M. S. (2014). Relative cost-effectiveness of a norovirus vaccine in the deployed military setting compared to a vaccine against campylobacter sp., etec, and shigella sp. *Vaccine*, *32*(40), 5156–62. doi:10.1016/j.vaccine.2014.07.070

Tan, S. S., Bouwmans, C. A., Rutten, F. F., & Hakkaart-van Roijen, L. (2012). Update of the dutch manual for costing in economic evaluations. *Int J Technol Assess Health Care*, *28*(2), 152–8. doi:10.1017/S0266462312000062

Tang, K. (2015). Estimating productivity costs in health economic evaluations: A review of instruments and psychometric evidence. *Pharmacoeconomics*, *33*(1), 31–48.

Thijs, K. M., de Boer, A. G., Vreugdenhil, G., van de Wouw, A. J., Houterman, S., & Schep, G. (2012). Rehabilitation using high-intensity physical training and long-term return-to-work in cancer survivors. *J Occup Rehabil*, *22*(2), 220–9. doi:10.1007/s10926-011-9341-1

Tompa, E., Dolinschi, R., Alamgir, H., Sarnocinska-Hart, A., & Guzman, J. (2016). A cost-benefit analysis of peer coaching for overhead lift use in the long-term care sector in canada. *Occup Environ Med*, *73*(5), 308–14. doi:10.1136/oemed-2015-103134

Tompa, E., Dolinschi, R., & de Oliveira, C. (2006). Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil*, *16*(3), 375–400. doi:10.1007/s10926-006-9035-2

Tompa, E., Dolinschi, R., de Oliveira, C., & Irvin, E. (2009). A systematic review of occupational health and safety interventions with economic analyses. *J Occup Environ Med*, *51*(9), 1004–23. doi:10.1097/JOM.0b013e3181b34f60

Tompa, E., Dolinschi, R., & Natale, J. (2013). Economic evaluation of a participatory ergonomics intervention in a textile plant. *Appl Ergon*, *44*(3), 480–7. doi:10.1016/j.apergo.2012.10.019

Tompa, E., Verbeek, J., van Tulder, M., & de Boer, A. (2010). Developing guidelines for good practice in the economic evaluation of occupational safety and health interventions. *Scand J Work Environ Health*, *36*(4), 313–8. doi:10.5271/sjweh.3009

Tompa, E., Culyer, A. J., & Dolinschi, R. (2008). *Economic evaluation of interventions for occupational health and safety: Developing good practice*. doi:10.5271/sjweh.3009

Torbeyns, T., Bailey, S., Bos, I., & Meeusen, R. (2014). Active workstations to fight sedentary behaviour. *Sports Med*, *44*(9), 1261–73. doi:10.1007/s40279-014-0202-x

Trogdon, J. G., Finkelstein, E. A., Hylands, T., Dellea, P. S., & Kamal-Bahl, S. J. (2008). Indirect costs of obesity: A review of the current literature. *Obes Rev*, *9*(5), 489–500. doi:10.1111/j.1467-789X.2008.00472.x

Troman, G., & Woods, P. (2000). Careers under stress: Teacher adaptations at a time of intensive reform. *J Educ Change*, *1*(3), 253–275. doi:10.1007/978-90-481-2358-2_7

Uegaki, K., de Bruijne, M. C., Lambeek, L., Anema, J. R., van der Beek, A. J., van Mechelen, W., & van Tulder, M. W. (2010). Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. *Scand J Work Environ Health*, *36*(4), 273–88. doi:10.5271/sjweh.3017

Uegaki, K., de Bruijne, M. C., van der Beek, A. J., van Mechelen, W., & van Tulder, M. W. (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*(1), 90–9. doi:10.1007/s10926-010-9258-0

Van Buynder, P. G., Konrad, S., Kersteins, F., Preston, E., Brown, P. D., Keen, D., & Murray, N. J. (2015). Healthcare worker influenza immunization vaccinate or mask policy: Strategies for cost effective implementation and subsequent reductions in staff absenteeism due to illness. *Vaccine*, *33*(13), 1625–1628.

van der Molen, H. F., Stocks, S. J., & Frings-Dresen, M. H. (2016). Exploring study designs for evaluation of interventions aimed to reduce occupational diseases and injuries. *Saf Health Work*, *7*(1), 83–85. doi:10.1016/j.shaw.2015.09.002

van Dongen, J. M., Proper, K. I., van Wier, M. F., van der Beek, A. J., Bongers, P. M., van Mechelen, W., & van Tulder, M. W. (2011). Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev*, *12*(12), 1031–49. doi:10.1111/j.1467-789X.2011.00925.x

van Dongen, J. M., Proper, K. I., van Wier, M. F., van der Beek, A. J., Bongers, P. M., van Mechelen, W., & van Tulder, M. W. (2012). A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health*, *38*(5), 393–408. doi:10.5271/sjweh.3275

van Dongen, J. M., Tompa, E., Clune, L., Sarnocinska-Hart, A., Bongers, P. M., van Tulder, M. W., ... van Wier, M. F. (2013). Bridging the gap between the economic evaluation literature and daily practice in occupational health: A qualitative study among decision-makers in the healthcare sector. *Implement Sci*, *8*(1), 57. doi:10.1186/1748-5908-8-57

van Eijsden, M. D., Gerhards, S. A., de Bie, R. A., & Severens, J. L. (2009). Cost-effectiveness of postural exercise therapy versus physiotherapy in computer

screen-workers with early non-specific work-related upper limb disorders (wruld); a randomized controlled trial. *Trials*, *10*, 103. doi:10.1186/1745-6215-10-103

van Wier, M. F., Verweij, L. M., Proper, K. I., Hulshof, C. T., van Tulder, M. W., & van Mechelen, W. (2013). Economic evaluation of an occupational health care guideline for prevention of weight gain among employees. *J Occup Environ Med*, *55*(9), 1100–9. doi:10.1097/JOM.0b013e31829b279a

Verbeek, J., Pulliainen, M., & Kankaanpaa, E. (2009). A systematic review of occupational safety and health business cases. *Scand J Work Environ Health*, *35*(6), 403–12.

Vermeulen, S. J., Heymans, M. W., Anema, J. R., Schellart, A. J. M., van Mechelen, W., & van der Beek, A. J. (2013). Economic evaluation of a participatory return-to-work intervention for temporary agency and unemployed workers sick-listed due to musculoskeletal disorders. *Scand J Work Environ Health*, *39*(1), 46–56. doi:10.5271/sjweh.3314

von Thiele Schwarz, U., & Hasson, H. (2012). Effects of worksite health interventions involving reduced work hours and physical exercise on sickness absence costs. *J Occup Environ Med*, *54*(5), 538–44. doi:10.1097/JOM.0b013e31824e11cd

Ware Jr, J. E., & Sherbourne, C. D. (1992). The mos 36-item short-form health survey (sf-36): I. conceptual framework and item selection. *Med Care*, 473–483. doi:10.1097/00005650-199206000-00002

Wendt, J. K., Tsai, S. P., Bhojani, F. A., & Cameron, D. L. (2010). The shell disability management program: A five-year evaluation of the impact on absenteeism and return-on-investment. *J Occup Environ Med*, *52*(5), 544–50. doi:10.1097/JOM.0b013e3181dbc8c2

Zhang, W., Bansback, N., & Anis, A. H. (2011). Measuring and valuing productivity loss due to poor health: A critical review. *Soc Sci Med*, *72*(2), 185–92. doi:10.1016/j.socscimed.2010.10.026

Zhang, W., Sun, H., Woodcock, S., & Anis, A. (2015). Illness related wage and productivity losses: Valuing 'presenteeism'. *Soc Sci Med*, *147*, 62–71. doi:10.1016/j.socscimed.2015.10.056

4.3 Suggestions for a reference case

Taking both sections together, and including more recent research, we would make the following suggestions for a reference case of economic evaluations in occupational health and safety:

- The CHEERS checklist should be completed for each economic evaluation (Husereau et al., 2013). Refer to and follow national guidelines for economic

evaluation, see Action et al., 2016 for an overview of European guidelines. Clearly report all elements, both from CHEERS, national guidelines, as well as the methods and choices made for the suggestions below.

- Include both costs and effects of the intervention, and avoid quick and simple decision-making tools.
- Health-related productivity:
  - Include productivity separately in economic evaluations, unless otherwise stated in national guidelines.
  - Using objective data on sick leave that are separable from other absences (e.g., holidays) is often the most precise. Otherwise, make use of a standardized measurement tool that includes paid and unpaid work, absenteeism, presenteeism, compensation mechanisms, and wage multipliers. The following review gives a good overview of validated instruments, and what they do (or do not) include: Tang, 2015.
  - Monetise health-related productivity using an adequate price weight (e.g. employee wage with benefits). Use both the human capital and the friction cost approach. New estimates of the friction period have recently been made for different European countries (Hanly, Ortega, Pearce, Soerjomataram, & Sharp, 2020). Adjust for the elasticity of labour-supply.
  - Add sensitivity analyses with job-dependent multipliers to indicate the absence of perfect substitutes, the effects of team production, and the penalty of not reaching an output target.
  - Correct estimates for inflation and apply discounting if the study length is more than a year.
- Intangible effects:
  - Assess whether or which intangible effects (effects on a firm's labour pool, reputation, risk and worry about health or safety) are applicable to the programme, and are relevant given the economic evaluation perspective. Take care not to double count effects (e.g. worry about health or safety could already be taken up through health measurement, effects on reputation or a firm's labour pool could be captured by productivity measurement).
  - Risk-informed decision making is to be preferred above risk-based decision making.
- Accounting for the legal and fiscal context:

  - State the location of the trial and economic evaluation, and give a brief summary of the country's relevant compensation system (worker compensation, sick pay, social insurance, accident compensation, disability insurance, employers' liability) insofar these affect the distribution of costs and effects of the programme.

- Distributive Equity

  - An economic evaluation from a societal perspective measures the (distribution of) costs and effects for all stakeholders, whereas a company's or healthcare payer's perspective only includes the costs and effects from the employer's or healthcare payer's point of view. The latter two thus take very specific views on distribution: maximising the value for employers (profit) or for healthcare payers (health). Although there is no consensus on which perspective is best to adopt (Krol & Brouwer, 2014), a strong ethical argument can be made for maximising overall well-being, and thus to adopt a societal perspective.
  - Including productivity effects in a company or societal perspective may favour treatments for productive individuals over less productive groups (e.g. elderly of children). We do not advocate their exclusion as productivity impacts welfare, but distributional weights could be used to correct for this tendency (Fleurbaey & Abi-Rafeh, 2016; Krol & Brouwer, 2014).

Chapter 5

DISCUSSION

This PhD aimed to evaluate periodic occupational health screenings, and, as a secondary objective, identified methodological challenges in economic evaluations of occupational health and safety interventions. First, a new screening instrument was developed and validated. The resulting occupational health screening tool is aimed at a broad employee population, contains questions on diverse and frequently occurring health problems, and shows adequate face validity, criterion validity, structural validity, and internal consistency. Next, this tool was used in a randomized controlled trial to assess its effectiveness in comparison to care-as-usual (a face-to-face screening with the occupational physician). The regressions and (superiority) post-hoc analyses indicate that there were some short-term effects of the intervention: trust in the physician, absenteeism and healthcare use (referrals, spontaneous consultations, and work-related consultations) were lower in the intervention group. However, these differences predominantly occurred in wave 1 and were no longer present in later measurement moments during COVID-19 peaks in the hospitals. Finally, for our secondary objective, several methodological points were made on the challenges of economic evaluations in occupational health and safety. Health-related productivity losses, intangible effects, and effects of the legal and fiscal context were shown to be important in occupational health and safety, but care should be taken when including them.

In the 20th century, health screenings were widely implemented under the widespread idea that "an ounce of prevention is worth a pound of cure". If we invest some resources now in screening, we can avoid underlying health problems from deteriorating, and spending more resources on curative care in the future. While doubts have been cast on the effectiveness of general

health checks (Krogsbøll et al., 2013), health screening has shown its value in specific domains. Famous examples include the effectiveness of the Pap test in reducing mortality due to cervical cancer, or screening for breast cancer by mammography (Morabia & Zhang, 2004).

We therefore set out on this project with the expectation that periodic health screenings are a valuable tool in the monitoring and follow-up of employees at risk, but suspected that efficiency gains were possible. Internationally as well as in Belgium, a substantial time investment has to be made, by employers, employees and occupational physicians, to follow-up each employee at risk in a face-to-face consultation. Our expectation that, in the short-term, electronic surveys and selective follow-up can save resources without unduly comprising the effectiveness of screening is confirmed in Chapter 3. The tool we developed in Chapter 2 proved at least as effective as care-as-usual on the majority of the outcomes, even during the exceptional COVID-19 circumstances, during which the hospitals' personnel was subject to exceptional stress. Our results are thus consistent with other literature, such as Krogsbøll et al., 2013, which indicated that there is no strong evidence that more general health screenings are (cost-)effective, while more focused screenings could be (MacAuley, 2012).

Some authors, such as Sluiter in the Netherlands, have made efforts to move away from general health screenings. They argue that most periodic health screenings require that '*long lists of diseases had to be checked off by the physician*' and too much focus has been placed on the classical clinical way of thinking: looking for symptoms, signs and diseases (J. K. Sluiter, 2017). Instead, they want to focus periodic health screenings on job- and task-specific health requirements, to ensure medical fitness assessments are adequately tailored to the demands of specific jobs. New surveys have been developed to support this view, by creating job- or task-specific instruments that focus on specific populations (e.g. nurses (Boezeman et al., 2016; Gärtner, Ketelaar, et al., 2011; Noben et al., 2014), ambulance workers (Van Schaaijk et al., 2017), or rail safety workers (Boschman et al., 2016)) (Ruitenburg et al., 2016). On the one hand, their specific and focused nature is beneficial to avoid unnecessary questions, and these surveys are thus able to save resources. On the other hand, creating job-specific questionnaires for all occupations is research-intensive, and general questionnaires should not be put aside too hastily. We believe a general questionnaire can still be of use for the majority of the population, as certain health and functioning complaints are frequent enough to be applicable to almost all occupations. In this PhD, we therefore chose to develop a screening survey for a general population including a range of health problems. However, this should not be seen as a return to 'long lists of diseases to be checked off by physicians': by focusing on frequent health problems, we ensure all questions are relevant for the assessment of medical fitness, while avoiding the burden of

long questionnaires on medical history, and every clinical symptom or disease. Nevertheless, the general tool we developed could be supplemented with specific risks to maximise its use for less frequent health or functioning problems that occur in specific populations. e.g. risk of violence for nurses, effects of vibration tools for construction workers, etc. Taking a task-specific approach (as suggested by Ruitenburg et al., 2016) would thus provide excellent additions to the general questionnaire proposed in this PhD.

In our secondary objective, we explored the state of economic evaluations in occupational health and safety, and identified their methodological challenges. Given the large economic consequences of occupational health and safety interventions for employers (especially when, such as in the USA, employers carry a portion of the consequences of employees' ill health), it came at no surprise that economic evaluations from the perspective of the company were frequent in this area. However, many of these economic evaluation studies lacked sufficient scientific quality, and diverged from the reference case (as described by Drummond et al., 2015) in how economic evaluations should be executed and reported. While this is partly due to a lack of knowledge, we also identified several factors that are hard to measure and monetise. Including health-related productivity or intangible effects, and accounting for the complex social, legal and fiscal context is challenging, but these aspects are simultaneously (in contrast to other domains) of key value for employers implementing occupational health and safety programmes. However, we also wish to emphasize that the relevance of these effects (e.g. effects on company reputation) is dependent on the context and the perspective taken in the economic evaluation. We advocated that clear reporting and adherence to standards (Husereau et al., 2013) would already substantially improve the methodological quality of economic evaluations in occupational health and safety.

Limitations Some limitations should be pointed out. Apart from the specific limitations of each study, outlined in the previous chapters, an important limitation is that the survey was developed simultaneously with the assessment of its effectiveness. While this is frequently done to save time and resources, it has the drawback that the instrument is still subject to change over time. However, by only removing items in subsequent questionnaire versions, we ensured the continuity of our data collection. Next, one could question the limited participation rates in the studies (e.g. 25% in the validation study, and 33% in the randomized controlled trial). However, scientifically speaking, the response rates were similar to other questionnaire-based studies, and sufficient sample sizes were obtained for hypotheses testing. Nevertheless, both the validation of the questionnaire and the effectiveness of electronic screening would benefit from further research. Finally, both the survey validation and

effectiveness studies took place in the hospital sector. While hospitals house a broad population (e.g. medical staff, technicians, administrative personnel, caretakers, cleaning staff), and are prone to a wide range of risks (physical, chemical, psychological, biological), it should be investigated more closely whether our results are also applicable outside the healthcare sector.

Future research Chapter 2 outlined the development and validation of the electronic screening tool, and although the results on the criterion validity of the tool were encouraging, the sample size was limited, and more reliability and psychometric validity tests are needed before implementing the survey on a large scale. Our sample was diverse (due to the different occupations in hospitals), but more highly educated, and contained more females than males. As outlined by De Vet et al., 2011, creating and validating a new survey is an iterative process. After the development of this first version, it would be opportune to increase the instrument's relevance by assessing the instrument's validity in other populations by taking different characteristics into account (e.g. lower education, more males, different occupational sectors). In this sense, risks specific to these employees' occupations (Ruitenburg et al., 2016; J. K. Sluiter, 2017) could be added to increase the applicability of the instrument. In addition, future research could include a second criterion analysis (with a larger sample size), and assess other validity dimensions outlined by the COSMIN taxonomy (De Vet et al., 2011; Mokkink et al., 2010). For instance, a predictive validity study could be set up, or a larger dataset would allow item scores to be weighted according to the importance of each subdomain in the survey (De Vet et al., 2011). The empirical weights for the subdomains (to achieve a weighted total score) could be calculated by estimating a MIMIC model (multiple reflective indicators and multiple formative causes) (De Vet et al., 2011). This would allow the instrument to relax the assumption of fixed weights of dimensions (each dimension now carries similar importance). Finally, it would be interesting to relax the assumption that 20% of employees need to be seen by the occupational physician. With more data, from a larger and more diverse population and linked to the assessments by occupation physicians, absolute criteria (e.g. always have a follow-up consultation when mental health score is above 8/12) could be used instead of relative criteria (20%). Instead of prioritising those that are worse off overall, the instrument could then select all employees that have worrisome problems in one or more areas.

Although we did not quantitatively assess cost-effectiveness, electronic screening is likely cost-saving in the short-run. Not only is 80% of the employees screened electronically (saving the time investment of attending a face-to-face consultation), employees that had a face-to-face consultation also showed significantly more healthcare use and absenteeism. However, it should also be

emphasized that it is as of yet unclear whether the initial investment of the control group (higher healthcare use and trust in the physician) is beneficial in the long run. It is possible that, due to the duration of the experiment (19 months), benefits of screening are yet to manifest. Furthermore, it is possible that face-to-face screenings perform better in specific subgroups, e.g. for older employees. Our sample size did not allow us to assess whether effectiveness differed across subgroups. Future research could therefore investigate the long-term effects of replacing face-to-face screenings by electronic targeted screening, and could include subgroup analyses.

Ideally, this PhD would have included a full economic evaluation of the electronic screening tool. Given the limited timing, this was not warranted within the current project. However, if the survey is tested on a larger scale, this would present an opportunity to further test its validity (see above) and perform an economic evaluation from the societal perspective at the same time. To keep the recommendations from Chapter 4 in mind, we advise to carefully measure health-related productivity (both absenteeism and presenteeism) with a validated survey (e.g. iPCQ or WHO-HPQ) and valuate it with both the human capital and friction cost approach, while accounting for the elasticity of labour supply, uncertainty, inflation, and discounting. The intangible effects we discussed in Chapter 4 (e.g. labour pool, and reputation) seem less important and relevant in this context, especially when taking a societal perspective. Similarly, the results from the trial in Chapter 3 also indicated effects on worrying about health were not significantly affected by screening. The cost-effectiveness results should also be clearly placed within the Belgian policy context, to elucidate how the results affect the distribution of burdens between the different parties (employees, employers, tax payers).

We did not discuss issues of distributive equity in depth within the PhD manuscript. However, as an extension of Chapter 4, we wish to note that an economic evaluation from a societal perspective measures the (distribution of) costs and effects for all stakeholders, whereas a company's or healthcare payer's perspective only includes the costs and effects from the employer's or healthcare payer's point of view. The latter two thus take very specific views on distribution: maximising the value for employers (profit) or for healthcare payers (health). Although there is no consensus on which perspective is best to adopt (Krol & Brouwer, 2014), a strong ethical argument can be made for maximising overall well-being, and thus to adopt a societal perspective. Furthermore, including productivity effects in a company or societal perspective may favour treatments for productive individuals over less productive groups (e.g. elderly of children). We do not advocate their exclusion as productivity impacts welfare, but distributional weights could be used to correct for this tendency (Fleurbaey & Abi-Rafeh, 2016; Krol & Brouwer, 2014). As an extension, it

can be noted that it can be beneficial to see well-being as multidimensional: individuals care about income and health, but also about their job, happiness, relations, family, and many other dimensions (Sen, 1992). Several methods have been suggested to take up these effects (Fleurbaey, Luchini, Muller, & Schokkaert, 2013; Schokkaert, 2015). Although more complex, it is thus the view of the author that an economic evaluation from a societal perspective that uses cost-benefit methods that monetise both costs and effects (in contrast to cost-effectiveness or cost-utility), employs multidimensional health, includes other dimensions (e.g. income, social relations, job satisfaction), and uses distributional weighting is to be preferred on ethical grounds (Schokkaert, 2015).

Policy consequences The goal of targeted screening is not necessarily to replace face-to-face screening, but mainly to supplement it. For many reasons, e.g. vision tests for drivers, or spirometry tests to detect occupational asthma, physical contact with the physician will remain necessary. There is also an important equity argument to be made to maintain a certain level of face-to-face screening. Our sample was more highly educated than the general population of the hospitals, suggesting that employees with lower educational attainment did not find their way to the online screening surveys. Given the generally lower socio-economic status of these employees, which is often linked to a lower health status and higher risk jobs, it is essential that their access to healthcare is maintained. For some employees, this is possibly the only visit to a physician they make all year.

We did not investigate whether stakeholders (employees, employers, and occupational physicians) prefer screening by survey or face-to-face, but can see some trends in the data. First, participation rates (e.g. 25% in the validation study, 33% in the randomized controlled trial) seem to suggest there is little interest in being screened through surveys. However, it should be noted that participation was hitherto voluntary, as was ethically required while setting up our studies. In the future, if surveys are standardly used as part of occupational health screenings, participation rates will most likely rise due to the obligated nature of periodic health screenings.

Since the start of the studies described above, Belgium has taken legal steps in the direction of implementing triage instruments ("Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017", n.d.). Before the implementation of this legislation, employees at risk were to be screened annually (although the exact frequency could vary), whereas now some employees can be followed up at two-yearly intervals with a medical questionnaire in between, which

serves to identify employees in need of closer attention. Now that research has caught up with practice, we would like to emphasize the importance of using a validated screening tool and algorithm when using targeted screening. The questionnaire developed in this PhD would be a suitable and validated candidate in this respect. In addition, we hope the implemented legislation will now further build upon the results presented in this PhD. Finally, given the generally positive results in the effectiveness study, other countries could contemplate implementing targeted screening as a way to maximise giving help to those who need it, while avoiding routine consultations.

Conclusion For the first two research objectives, a new generalized screening tool has been developed based upon literature, expert opinion, and the target population, aimed at a broad population and making use of frequent health problems. The questionnaire-based algorithm has adequate face validity, criterion validity, structural validity, and internal consistency. The recurring health and risk data from the survey could prove useful for workers' follow-up and for epidemiological studies. Next, the effectiveness study finds no indication that, in the short-term, substituting workforce screening by a physician with a quicker survey-based screening with targeted follow-up has different effects on the studied endpoints. However, as health and disease are often the result of a long-term process, more evidence is needed to determine long-term effects. For those populations where physical consultations are not strictly necessary, a lower frequency of face-to-face health screening might prove at least as effective if combined with a more targeted approach to ensure those who really need it are still offered adequate care.

Regarding our secondary objective, general guidelines on economic evaluation remain largely applicable to the field of OHS, but researchers should pay attention to the listed challenges. To ensure economic evaluations in OHS adequately inform decision-makers (especially employers), future methodological work should continue research on productivity, intangibles, and legal/fiscal effects, and work towards standardised methods of including them in economic evaluations, either directly or indirectly. In addition, the systematic review on health-related productivity shows it is a key element of the economic attractiveness of investing in OHS. Economic evaluation studies of OHS would benefit from more methodological standardization in their approach to quantifying productivity change. Future research should better account for the methodological uncertainty that occurs in estimating it, in order to demonstrate the impact that particular choices and approaches to productivity estimation can have on cost-effectiveness results.

References

Boezeman, E. J., Nieuwenhuijsen, K., & Sluiter, J. K. (2016). Predictive value and construct validity of the work functioning screener–healthcare (wfs-h). *J Occup Health*, 15-0056–OA. doi:10.1539/joh.15-0056-OA

Boschman, J., Hulshof, C., Frings-Dresen, M., & Sluiter, J. (2016). Improving fit to work assessments for rail safety workers by exploring work limitations. *Int Arch Occup Environ Health*, *89*(5), 803–811. doi:10.1007/s00420-016-1117-7

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Univ Pr.

Fleurbaey, M., & Abi-Rafeh, R. (2016). The use of distributional weights in benefit–cost analysis: Insights from welfare economics. *Review of Environmental Economics and Policy*, *10*(2), 286–307.

Fleurbaey, M., Luchini, S., Muller, C., & Schokkaert, E. (2013). Equivalent income and fair evaluation of health care. *Health Economics*, *22*(6), 711–729.

Gärtner, F. R., Ketelaar, S. M., Smeets, O., Bolier, L., Fischer, E., van Dijk, F. J., ... Sluiter, J. K. (2011). The mental vitality@ work study: Design of a randomized controlled trial on the effect of a workers' health surveillance mental module for nurses and allied health professionals. *BMC Public Health*, *11*(1), 1–13. doi:10.1186/1471-2458-11-290

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., ... Loder, E. (2013). Consolidated health economic evaluation reporting standards (cheers) statement. *The European Journal of Health Economics*, *14*(3), 367–372.

Krogsbøll, L. T., Jørgensen, K., & Gøtzsche, P. C. (2013). General health checks in adults for reducing morbidity and mortality from disease. *JAMA*, *309*(23), 2489–2490. doi:10.1001/jama.2013.5039

Krol, M., & Brouwer, W. (2014). How to estimate productivity costs in economic evaluations. *Pharmacoeconomics*, *32*(4), 335–44. doi:10.1007/s40273-014-0132-3

MacAuley, D. (2012). The value of conducting periodic health checks. British Medical Journal Publishing Group.

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... De Vet, H. C. (2010). The cosmin checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol*, *10*(1), 1–8. doi:10.1186/1471-2288-10-22

Morabia, A., & Zhang, F. F. (2004). History of medical screening: From concepts to action. *Postgraduate medical journal*, *80*(946), 463–469.

Noben, C., Smit, F., Nieuwenhuijsen, K., Ketelaar, S., Gärtner, F., Boon, B., ... Evers, S. (2014). Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud*, *51*(10), 1321–1331. doi:10.1016/j.ijnurstu.2014.01.017

Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017. (n.d.).

Ruitenburg, M. M., Frings-Dresen, M. H., & Sluiter, J. K. (2016). How to define the content of a job-specific worker's health surveillance for hospital physicians? *Saf Health Work*, *7*(1), 18–31. doi:10.1016/j.shaw.2015.08.004

Schokkaert, E. (2015). How to introduce more (or better) ethical arguments in hta? *International journal of technology assessment in health care*, *31*(3), 111–112.

Sen, A. (1992). *Inequality reexamined*. Oxford University Press.

Sluiter, J. K. (2017). The new why when designing mandatory medical examinations. *Occup Med (Lond)*, *67*(5), 325–327. doi:10.1093/occmed/kqx044

Van Schaaijk, A., Boschman, J., Frings-Dresen, M., & Sluiter, J. (2017). Appraisal of work ability in relation to job-specific health requirements in ambulance workers. *Int J Environ Res Public Health*, *90*(1), 123–131. doi:10.1007/s00420-016-1181-z

Appendix A

APPENDIX

A.1 Appendix to Chapter 2

A.1.1 Scoring algorithm

Both scoring systems provide corrections on the item- (maximum item score or number of items per dimension), dimension- (number of dimensions per category), and category-level (number of categories in the questionnaire). Scoring system (1) is an unweighted aggregation of the answers with numeric interpretation of Likert scales, and is shown in Equation 1, with Γ the maximum score of an item, I the number of items (i) in a dimension (d), D the number of dimensions in a category (c), and C the number of categories in the questionnaire:

$$\frac{1}{C}\sum_{c=1}^{C}\left(\frac{1}{D}\sum_{d=1}^{D}\left(\sum_{i=1}^{I}\frac{item_i}{\Gamma_i}\right)_d\right)_c \tag{A.1}$$

For instance, there were four dimensions in the mental health category: the General Health Questionnaire (12 items, maximum score 12), Need for Recovery Scale (11 items, maximum score 11), stress (3 items, maximum score 12), and burnout (4 items, maximum score 12). For system (1), the score on each dimension was converted to a percentage (e.g. $5/12 = 0.42$ for GHQ), after which the mental health category score was calculated as the arithmetic mean of these dimensions (e.g. $(0.42 + 0.21 + 0.18 + 0.91)/4$). The total score was then calculated as the arithmetic mean of the category scores $\frac{\sum_{i=1}^{N} category_i}{N}$.

The former (scoring system 1) uses the numeric conversion of answer categories directly, and thus assumes equal distances between answer categories. For instance, the distance between a stress score of 1 and 2 is assumed to be identical to the distance between a stress score of 11 and 12. However, this assumption is sometimes contested. Scoring system (2) shown in Equation 2 is therefore tested in addition to scoring system (1) to relax this assumption. Scoring system (2) is identical to system (1), but uses ordinal scoring, which replaces the answer of the employee by its rank vis à vis the other respondents, while ties are averaged. The calculation of the total score is similar to system (1), but uses ordinal rank scores.

$$\frac{1}{C}\sum_{c=1}^{C}\left(\frac{1}{D}\sum_{d=1}^{D}\left(\frac{1}{I}\sum_{i=1}^{I} rank(item_i)\right)_d\right)_c \tag{A.2}$$

A.1.2 Overview of categories and dimensions

Table A.1: The 12 categories and 107 dimensions identified in the systematic review and presented to the Delphi panel, and the percentage of panel members that deemed the dimension 'important' or 'very important' to include in the screening tool in round 3. Numbers in bold indicate more than 80% consensus in the Delphi, italic text indicates the 15 variables added in round 2 of the Delphi. B) Green groups indicate categories brought up by the focus group; Blue groups were prompted first. %IMP = percentage of panel members that deemed the dimension 'important' or 'very important' to include in the screening tool in round 3.

CATEGORY	DIMENSIONS	%IMP
Personal characteristics or resources		
	work centrality	44
	competence	44
	work history	22
	vulnerability	44
	social context	56
	coping	78
	age and gender	22
	Perfectionism	33
Work organisation		
	leadership and management style	44
	organisation culture	56
	organisational justice	56
Job content		
	variation	56
	risk on physical burden	78
	risk on mental or psychosocial burden	**89**
	meaningful work	78
	autonomy	67
	work pressure or fast work pace	78
	emotional demands	67
	cognitive demands	56
Employment conditions		
	work-life balance	67
	contract terms	11
	satisfaction with pay	11
	work time (shift- or nightwork)	**89**
	organisational changes	78

	development possibilities	44
	job insecurity	56
	effort-reward imbalance	22
Work environment		
	distance work-home	22
	risk on poisoning	33
	risk by physical agents	44
	risk by biological agents	44
	risk on dermatosis	56
	risk on inhalation	56
	safety risks	33
	external burden	22
	risks that demand further medical examination	11
	use personal protective equipment	44
	other risks	11
	maternal protection	44
Social work environment		
	support at work	78
	role clarity or conflicts	**89**
	personal or emotional conflicts	78
	bullying	**89**
	unwanted sexual behaviour	78
	risk on violence	**89**
Attitude towards job and organisation		
	job satisfaction	**100**
	hazardous	44
	happiness and enthusiasm at work	**100**
Productivity		
	absenteeism	**100**
	presenteeism AM	22
	presenteeism EUR	44
	productivity or work quality	44
	contra productive work behaviour	22
	organisational citizenship behaviour or extra-role behaviour	11
	creative work behaviour	0
	innovative work behaviour	0
	self-efficacy	0
Mental Health		
	rumination	22
	recovery needs after work	**89**
	mental, behavioural or neurological disorders	56
	work addiction	22
	general mental health and well-being	**89**
	stress	**100**
	traumatic events	67
	second victim	56
	boreout	22
	burnout	**100**
	apathy	11
Physical health		
	general health	**100**
	contagious or parasitic diseases	56
	neoplasms	**89**
	blood or blood-forming organ diseases	33
	immunity diseases	67
	endocrine, food, or metabolical diseases	**89**
	sleep-wake diseases	**89**
	diseases of the nervous system	78
	diseases of the ear or mastoid	33
	diseases of the circulatory system	67
	diseases of the respiratory system	**89**
	diseases of the digestive system	56
	diseases of the skin	78
	diseases of the musculoskeletal system or connective tissue	**100**
	diseases of the urogenital system	11
	sexual health	11
	pregnancy or pregnancy-desire	56
	diseases during pregnancy, birth, or post-natal	22
	other not classified symptoms, signs, or clinical findings	33
	injury, poisoning, or other external consequences	56
	external causes of morbidity or mortality	11
	allergies or contact dermatitis	**100**
	functioning	78
	pain	**100**
	other health problems (open question)	**89**
Medication and healthcare use		
	contact with healthcare providers	33
	medication use	**89**
	medication dependence	**100**
Healt(care) behaviour and literacy		

help seeking behaviour	22
trust in the occupational physician	78
health literacy	78
lifestyle	**100**
following advice on infection control	33
spontaneous consultation	**89**
practical knowledge concerning the occupational health department	67
influence of work on family	56

A.1.3 Descriptive statistics Study 1

Table A.2: Descriptive statistics Study 1 to assess criterion validity (n=58)

Variable	Levels	Count/Mean	%/sd
Gender	Male	31	53.5
	Female	27	46.5
Age	(numeric)	36	11.6(sd)
Income	Income less than €500 monthly	0	0
	Income €500.00 - €999.99 monthly	1	2.1
	Income €1000.00 - €1499.99 monthly	7	14.6
	Income €1500.00 - €1999.99 monthly	26	54.2
	Income €2000.00 - €2499.99 monthly	11	22.9
	Income €2500.00 - €2999.99 monthly	2	4.2
	Income €3000.00 - €4999.99 monthly	1	2.1
	Income €5000,00 monthly	0	0
	Income not stated	10	0.17
Nationality	Belgian	57	98.3
	Migrant inside EU	1	1.7
	Migrant outside EU	0	0
Household	Single	12	20.7
	Single with child(ren)	6	10.3
	Couple with child(ren)	26	44.8
	Couple without child(ren)	11	19
	Other	3	5.2

A.1.4 Descriptive statistics Study 2

Table A.3: Descriptive statistics Study 2 to assess internal consistency and structural validity (n=776). Numeric variables state mean and standard deviation (sd), categorical variables state count and proportion.

Variable	Levels	Mean/ Count	Proportion/ Standard Deviation
Gender	Male	143	18.50%
	Female	631	81.40%
	Other	1	0.10%
Age	Numeric [0-66]	45.39	11.2
Income	Income €1500.00 - €1999.99 monthly	240	37.10%
	Income €2000.00 - €2499.99 monthly	241	37.20%
	Income €2500.00 - €2999.99 monthly	76	11.70%
	Income €3000.00 - €4999.99 monthly	26	4%
	Income > €5000.00 monthly	5	0.80%
	Income not stated	59	9.10%
Nationality	Belgian	763	98.30%
	Migrant inside EU	9	1.20%
	Migrant outside EU	4	0.50%
Education	No degree	7	0.90%
	Primary education	23	3%
	Secondary education	141	18.20%
	Higher education	588	75.80%
	Other education	17	2.20%
Household	Single	89	11.50%
	Single with child(ren)	63	8.10%
	Couple with child(ren)	480	61.90%
	Couple without child(ren)	122	15.70%
	Other	22	2.80%
Hospital	Hospital A	197	25.40%
	Hospital B	266	34.30%
	Hospital C	214	27.60%
	Hospital D	99	12.80%
Occupation	Medical Personnel	394	52.60%
	Paramedics	20	2.70%
	Technicians	56	7.50%
	Administrative Personnel	120	16%
	Management	34	4.50%
	Cleaning Staff	34	4.50%
	Other occupations	91	12.10%
Musculoskeletal (NMQ)1	Complaint last 12m	701	90.30%
	Complaint last 7d	558	72.00%
	Functioning complaint	358	46.10%
Skin (NOSQ)2	Skin complaint	155	20.00%
General Mental Health (GHQ)3	Numeric [0-12]	2.07	2.9
Need for Recovery (NFR)4	Numeric [0-11]	3.78	2.8
EQ5D Visual Analogue Scale	Numeric [0-100]	78.95	14.6
Respiratory (mMRC)5	Occasional breathlessness	141	18.2
	Severe breathlessness	40	5.2
Alcohol Use Disorders (AUDIT-C)6	Numeric [0-12]	2.03	1.6
Burnout (COPSOQ)7	Moderate risk of burnout (50-74)	286	37.90%
	High risk of burnout (75-99)	98	12.60%
	Severe risk of burnout (100)	3	0.03%
Stress (COPSOQ)7	Numeric [0-12]	4.01	2.5
Absence (iPCQ)8	Absent at least once last 4 weeks	74	0.10%
	Days absent last 4 weeks [0-28]	1.34	5.4
Trust in Physician score	Numeric [0-55]	37.85	6.17
Relationship with the occupational physician	Knows OP for more than 5 years	327	42.20%
	Knows OP for 3-4 years	135	17.40%
	Knows OP for 1-2 years	165	21.30%
	Knows OP less than a year	147	19%

Prescribed medication use in the last two weeks	Yes	397	51.20%
Medication dependence – used recreationally	Never	718	92.60%
	Monthly or less	44	5.70%
	2 - 4 times per month	9	1.20%
	2 - 3 times per week	2	0.30%
	4 or more times per week	2	0.30%

Missing values were removed from this table.
1 NMQ=Nordic Musculoskeletal Questionnaire,
2 NOSQ= Nordic Occupational Skin Questionnaire , 3 GHQ = General Health Questionnaire,
4 NFR = Need for Recovery Scale, 5 mMRC=Medical Research Council Dyspnoe questionnaire,
6 Alcohol Use Disorders Identification Test,
7 COPSOQ=Copenhagen Psychosocial Questionnaire
(burnout questions are based on the Copenhagen Burnout Inventory),
8 iPCQ=iMTA Productivity Cost Questionnaire

A.1.5 Questionnaire

Table A.4: Questionnaire. The translation of the Dutch questions are for information purposes only, and are, due to copyright, not always the official versions of the questionnaires. The English questions were not validated in our article. Request permission for the use of the official versions of these questions by contacting the authors of the original questionnaires.

Category	Dimension	Item Code	Question (English)	Question (Dutch)	Answer (English)	Answer (Dutch)	SCORES	source
Social work environment	Role Clarity Or Conflicts	role1	I receive conflicting demands	1. Ik krijg tegenstrijdige bevelen.	Totally disagree Rather disagree Partly agree, partly disagree Rather agree Totally agree	o Helemaal niet akkoord o Eerder niet akkoord o Deels akkoord, deels niet akkoord o Eerder akkoord o Helemaal akkoord	0-4	WOCCQ
Social work environment	Role Clarity Or Conflicts	role2	I am torn apart between people who have different expectations from my work	2. Ik word heen en weer geslingerd tussen mensen die andere verwachtingen hebben in verband met mijn werk.	Totally disagree Rather disagree Partly agree, partly disagree Rather agree Totally agree	o Helemaal niet akkoord o Eerder niet akkoord o Deels akkoord, deels niet akkoord o Eerder akkoord o Helemaal akkoord	0-4	WOCCQ
Social work environment	Role Clarity Or Conflicts	role3	I receive tasks that are not part of my job	3. Ik krijg taken toegewezen die niet onder mijn functie vallen.	Totally disagree Rather disagree Partly agree, partly disagree Rather agree Totally agree	o Helemaal niet akkoord o Eerder niet akkoord o Deels akkoord, deels niet akkoord o Eerder akkoord o Helemaal akkoord	0-4	WOCCQ
Attitude towards job organisation	Job Satisfaction	JS	Regarding your work in general. How pleased are you with your job as a whole, everything taken into consideration?	Over uw werk in het algemeen: Hoe tevreden bent u over uw job in zijn geheel, alles in beschouwing genomen?	Very satisfied, Satisfied, Neither/Nor, Unsatisfied, Very unsatisfied	o Zeer tevreden o Tevreden o Neutraal o Ontevreden o Zeer ontevreden	0-4	COPSOQ III
Attitude towards job organisation	Happiness And Enthousiasm At Work	bevl1	At my work, I feel bursting with energy.	1. Op mijn werk bruis ik van energie	Never, Seldom, Sometimes, Often, Always	o Nooit o Zelden o Soms o Vaak o Altijd	0-4	COPSOQ III
Attitude towards job organisation	Happiness And Enthousiasm At Work	bevl2	I am enthusiastic about my job.	2. Ik ben enthousiast over mijn baan.	Never, Seldom, Sometimes, Often, Always	o Nooit o Zelden o Soms o Vaak o Altijd	0-4	COPSOQ III

Attitude towards job organisation	Happiness And Enthousiasm At Work	bevl3	I am immersed in my work.	3. Ik ga helemaal op in mijn werk.	Never, Seldom, Sometimes, Often, Always	o Nooit o Zelden o Soms o Vaak o Altijd	0-4	COPSOQ III
Productivity	Absenteeism	abs1d	Have you been absent from work because of illness in the last 4 weeks?	Bent u in de afgelopen 4 weken afwezig geweest van uw werk omdat u ziek was?	No Yes, i have been ... Days absent. Only count working days in the last 4 weeks.	o Nee o Ja, ik ben ... dagen afwezig geweest. Tel alleen de werkdagen in de afgelopen 4 weken	0 1	iPCQ
Physical Health	General Health	EQ5Dvas	We would like to know how good or bad your health is TODAY. This scale is numbered from 0 to 100. 100 means the best health you can imagine 0 means the worst health you can imagine. Use the marker on the scale to indicate how your health is TODAY.	We willen weten hoe goed of slecht uw gezondheid VANDAAG is. • Deze meetschaal loopt van 0 tot 100. • 100 staat voor de beste gezondheid die u zich kunt voorstellen. 0 staat voor de slechtste gezondheid die u zich kunt voorstellen. • Geef op de meetschaal aan hoe uw gezondheid VANDAAG is.	Visual Analog Scale 0-100 health: the worst health you can imagine - the best health you can imagine.	VAS 0-100 health: de slechtste gezondheid die u zich kunt voorstellen - de beste gezondheid die u zich kunt voorstellen	INVERSED: 100 - VAS	EQ-5D
Mental Health	General Mental Health And Well-Being	GHQ1	In the last few weeks, have you... been able to concentrate	In de afgelopen weken, ... 1. Hebt u zich kunnen concentreren op uw bezigheden?	Better than usual Same as usual Less than usual Much less than usual	o Beter dan gewoonlijk o Net zo goed als gewoonlijk o Slechter dan gewoonlijk o Veel slechter dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ2	loss of sleep over worry	2. Bent u door zorgen veel slaap tekortgekomen?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ3	playing a useful part	3. Hebt u het gevoel gehad zinvol bezig te zijn?	More so than usual Same as usual Less useful than usual Much less useful	o Zinvoller dan gewoonlijk o Net zo zinvol als gewoonlijk o Minder zinvol dan gewoonlijk o Veel minder zinvol dan gewoonlijk	0 0 1 1	GHQ12

Mental Health	General Mental Health And Well-Being	GHQ4	capable of making decisions	4. Voelde u zich in staat om beslissingen te nemen?	More so than usual Same as usual Less so than usual Much less capable	o Beter dan gewoonlijk o Net zo goed als gewoonlijk o Wat minder goed dan gewoonlijk o Veel minder goed dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ5	felt constantly under strain	5. Hebt u het gevoel gehad dat u voortdurend onder druk stond?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ6	couldn't overcome difficulties	6. Hebt u het gevoel gehad dat u uw moeilijkheden niet de baas kon?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ7	able to enjoy day-to-day activities	7. Hebt u plezier kunnen beleven aan uw gewone, dagelijkse bezigheden?	More so than usual Same as usual Less so than usual Much less than usual	o Meer dan gewoonlijk o Even veel als gewoonlijk o Wat minder dan gewoonlijk o Veel minder dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ8	able to face problems	8. Bent u in staat geweest uw problemen onder ogen te zien?	More so than usual Same as usual Less so than usual Much less able	o Beter dan gewoonlijk o Net zo goed als gewoonlijk o Minder goed dan gewoonlijk o Veel minder goed dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ9	feeling unhappy or depressed	9. Hebt u zich ongelukkig en neerslachtig gevoeld?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk	0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ10	losing confidence	10. Bent u het vertrouwen in uzelf kwijtgeraakt?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk	0 0 1 1	GHQ12

Mental Health	General Mental Health And Well-Being	GHQ11	thinking of self as worthless	11. Hebt u zich als een waardeloos iemand beschouwd?	Not at all No more than usual Rather more than usual Much more than usual	o Helemaal niet o Niet meer dan gewoonlijk o Wat meer dan gewoonlijk o Veel meer dan gewoonlijk		0 0 1 1	GHQ12
Mental Health	General Mental Health And Well-Being	GHQ12	feeling reasonably happy	12. Hebt u zich alles bij elkaar redelijk gelukkig gevoeld?	More so than usual About same as usual Less so than usual Much less than usual	o Gelukkiger dan gewoonlijk o Even gelukkig als gewoonlijk o Minder gelukkig dan gewoonlijk o Veel minder gelukkig dan gewoonlijk		0 0 1 1	GHQ12
Mental Health	Stress	stress1	These questions are about how you have been during the last 5 weeks. How often have you had problems relaxing?	De volgende drie vragen gaan over de afgelopen 5 weken. 1. Hoe vaak had u problemen om zich te ontspannen?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Mental Health	Stress	stress2	How often have you been irritable?	2. Hoe vaak was u prikkelbaar?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Mental Health	Stress	stress3	How often have you been tense?	3. Hoe vaak was u gespannen?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Mental Health	Burnout	burnout1	These questions are about how you have been during the last 5 weeks. How often have you felt worn out?	De volgende vier vragen gaan over de afgelopen 5 weken. 1. Hoe vaak voelde u zich afgemat?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Mental Health	Burnout	burnout2	How often have you been physically exhausted?	2. Hoe vaak was u lichamelijk uitgeput?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III

Mental Health	Burnout	burnout3	How often have you been emotionally exhausted?	3. Hoe vaak was u emotioneel uitgeput?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4	COPSOQ III
Mental Health	Burnout	burnout4	How often have you felt tired?	4. Hoe vaak voelde u zich moe?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4	COPSOQ III

Mental Health	Recovery Needs After Work	nfr	1. I find it difficult to relax at the end of a working day. 2. By the end of the working day, I feel really worn out. 3. Because of my job, at the end of the working day I feel rather exhausted. 4. After the evening meal, I generally feel in good shape. 5. In general, I only start to feel relaxed on the second non-working day. 6. I find it difficult to concentrate in my free time after work. 7. I cannot really show any interest in other people when I have just come home myself. 8. Generally, I need more than an hour before I feel completely recuperated after work. 9. When I get home from work, I need to be left in peace for a while. 10. Often, after a day's work I feel so tired that I cannot get involved in other activities. 11. A feeling of tiredness prevents me from doing my work as well as I normally would during the last part of the working day.	1. Ik vind het moeilijk om me te ontspannen aan het einde van een werkdag. 2.Aan het einde van een werkdag ben ik echt op. 3. Mijn baan maakt dat ik me aan het einde van een werkdag nogal uitgeput voel. 4. Na het avondeten voel ik me meestal nog vrij fit. 5. Ik kom meestal pas op een tweede vrije dag tot rust. 6. Het kost mij moeite om me te concentreren in mijn vrije uren na het werk. 7. Ik kan weinig belangstelling opbrengen voor andere mensen, wanneer ik zelf net thuis ben gekomen. 8. Het kost mij over het algemeen meer dan een uur voordat ik helemaal hersteld ben na mijn werk. 9. Als ik thuis kom moeten ze mij even met rust laten. 10. Het komt vaak voor dat ik na een werkdag door vermoeidheid niet meer toekom aan andere bezigheden. 11. Het komt voor dat ik tijdens het laatste deel van de werkdag door vermoeidheid mijn werk niet meer zo goed kan doen.	No Yes	Nee Ja	0 1 * 11	The Need for Recovery Scale

Physical Health	Sleep-Wake Diseases	sleep1	These questions are about how you have been during the last 4 weeks. How often have you slept badly and restlessly?	De volgende vier vragen gaan over de afgelopen 4 weken. 1. Hoe vaak heeft u slecht en onrustig geslapen?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Physical Health	Sleep-Wake Diseases	sleep2	How often have you found it hard to go to sleep?	2. Hoe vaak vond u het moeilijk om te gaan slapen?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Physical Health	Sleep-Wake Diseases	sleep3	How often have you woken up too early and not been able to get back to sleep?	3. Hoe vaak werd u te vroeg wakker en was u niet meer in staat om de slaap te hervatten?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Physical Health	Sleep-Wake Diseases	sleep4	How often have you woken up several times and found it difficult to get back to sleep?	4. Hoe vaak werd u meerdere keren wakker en vond u het moeilijk om de slaap te hervatten?	All the time; A large part of the time; Part of the time; A small part of the time; Not at all	o De hele tijd o Een groot deel van de tijd o Een deel van de tijd o Een klein deel van de tijd o Helemaal niet	0-4		COPSOQ III
Physical Health	Diseases Of The Musculoskeletal System/-Connective Tissue	musc	Have you at any time during the last 12 months been prevented from doing your normal work (at home or way from home) because of the trouble? van uw nek van uw schouders van uw ellebogen van uw polsen/handen boven in de rug onder in de rug van uw heupen/dijen van uw knieën van uw enkels/voeten	Werd u op eender welk moment gedurende de laatste 12 maanden verhinderd om uw normale werk (thuis of niet thuis) te doen door last: van uw nek van uw schouders van uw ellebogen van uw polsen/handen boven in de rug onder in de rug van uw heupen/dijen van uw knieën van uw enkels/voeten	No Yes	Nee Ja	0 1 * 9		NMQ

A.2 Appendix to Chapter 3

A.2.1 Generalised Mixed Effects Models

Generalised Linear Mixed Effects models (GLMM) were used to assess the effect of the intervention upon the outcome measures. We base the following paragraphs on (Molenberghs & Verbeke, 2006; UCLA Insitute for Digital Research and Education, n.d.; Verbeke & Molenberghs, 2000). In its non-generalised form, the linear mixed effect model (LMM) is an extension of simple linear models by allowing for both fixed and random effects. The fixed effects are introduced on the highest level of the data, are the same for all subjects, and capture between group differences. The random effects are subject-specific and capture deviations from the fixed effects, and thus capture part of the within group variability. Analogous to subjects randomly drawn from a population of subjects, the subject-specific random effects can be considered as randomly sampled from a population of regression coefficients, often with an assumed underlying distribution (e.g. a multivariate normal distribution). In its hierarchical form, the linear mixed effect model (LMM) can be represented as follows (Verbeke & Molenberghs, 2000):

$$\begin{cases} \mathbf{Y_i} = \boldsymbol{X_i}\boldsymbol{\beta} + \boldsymbol{Z_i}\boldsymbol{b_i} + \boldsymbol{\varepsilon_i} \\ \boldsymbol{b_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{D}) \\ \boldsymbol{\varepsilon_i} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma_i}) \\ \boldsymbol{b_i}, ...\boldsymbol{b_n}, \boldsymbol{\varepsilon_i}, ..., \boldsymbol{\varepsilon_n} \quad \text{independent} \end{cases} \tag{A.3}$$

with $\mathbf{Y_i}$ the response vector of repeated measurements for subject i ($1 \leq i \leq N$, N number of subjects), $\mathbf{X_i}$ and $\mathbf{Z_i}$ matrices of covariates, $\boldsymbol{\beta}$ a vector with the fixed effects, $\boldsymbol{b_i}$ a vector of random effects, and $\boldsymbol{\varepsilon_i}$ the residuals. Note how $\boldsymbol{\beta}$ does not have a subscript i (the parameters are the same for all subjects), but $\boldsymbol{b_i}$ does (parameters vary over subjects). $\mathbf{D}$ and $\boldsymbol{\Sigma_i}$ are covariance matrices; they imply a flexible and specific covariance structure that allows for highly unbalanced data. Equation (A.3) can also be stated as follows (Molenberghs & Verbeke, 2006):

$$\begin{cases} \boldsymbol{Y_i}|\boldsymbol{b_i} \sim \mathcal{N}(\boldsymbol{X_i}\boldsymbol{\beta} + \boldsymbol{Z_i}\boldsymbol{b_i}, \boldsymbol{\Sigma_i}) \\ \boldsymbol{b_i} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{D}) \\ \boldsymbol{Y_{ij}}|\boldsymbol{b_i} \quad \text{independent} \end{cases} \tag{A.4}$$

In a non-Bayesian framework, estimating (A.4) is based on the marginal distribution for $\mathbf{Y_i}$ (integrating out the random effects) through maximum

likelihood, which renders consistent, asymptotic normal, and efficient estimates. The model-fitting procedures are iterative and Newton-Raphson-based, and are described in the R package 'nlme' documentation (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2021) and Lindstrom and Bates, 1988.

Analogous to univariate standard and generalised linear models, the linear mixed model (LMM) in Equation (A.3) and (A.4) is a specific case of generalised linear mixed models (GLMM). In GLMM, $\mathbf{Y_{ij}}$ have densities $f_i(y_{ij}|\theta_{ij}, \phi)$ of the exponential family (e.g. normal, Poisson, generalised Poisson, negative binomial, or binomial distributed dependent variables) with θ a 'natural' parameter and ϕ a 'scale' parameter. GLMM models can then be represented as follows (Molenberghs & Verbeke, 2006):

$$\begin{cases} \boldsymbol{Y_i}|\boldsymbol{b_i} \sim \text{Exp}_i(\boldsymbol{\theta}, \boldsymbol{\phi}) \\ \boldsymbol{b_i} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{D}) \\ \boldsymbol{Y_{ij}}|\boldsymbol{b_i} \quad \text{independent} \end{cases} \tag{A.5}$$

To estimate (A.5), the mean μ is then modelled for a link function $\eta(\,\cdot\,)$: $\eta(\mu_{ij}) = \eta[E(Y_{ij}|\boldsymbol{b_i})] = \boldsymbol{x'_{ij}\beta} + \boldsymbol{z'_{ij}b_i}$. If the natural link function is used $\eta(\,\cdot\,) = \psi'(\,\cdot\,)$, this renders $\boldsymbol{\theta_{ij}} = \boldsymbol{x'_{ij}\beta} + \boldsymbol{z'_{ij}b_i}$, which satisfies a linear regression model.

If $\boldsymbol{Y_i}|\boldsymbol{b_i}$ is normally distributed with mean $\boldsymbol{X_i\beta} + \boldsymbol{Z_ib_i}$, and covariance $\boldsymbol{\Sigma_i}$, Equation (A.5) collapses to (A.4). The density $f_i(y_{ij})$ is then the normal distribution, with $\theta = \mu$, $\phi = \sigma^2$, and variance function $v(\mu) = 1$. If $f_i(y_{ij})$ is the Poisson density with $\theta = ln\lambda$, $\phi = 1$, variance function $v(\lambda) = \lambda$, and log link function, a Poisson mixed model can be estimated, or a negative binomial model when an extra dispersion parameter is added. If $f_i(y_{ij})$ is a Bernoulli density, and using a probit or logit link function, a probit or logit normal model can be obtained.

Equations (A.6) and (A.7) give our model for one particular employee in two levels, for the i-th measurement for the j-th employee. On the first level (L1), the β has subscripts (j) to indicate they belong to one particular employee. On the second level, each β for a particular employee (β_{pj}) can be represented as a combination of a mean estimate for that parameter (γ_{p0}), and a random effect for that measurement moment (u_{pj}). Only the intercept (β_{0j}) and time slope (β_{2j}) can vary across employees, the other β_{pj} are fixed.

$$L1 : Y_{ij} = \beta_{0j} + \beta_{1j}Group_{ij} + \beta_{2j}Time_{ij} + \beta_{3j}Group{*}Time_{ij} + \beta_{pj}Covariates_{ij} + e_{ij} \tag{A.6}$$

$$L2: \begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} \\ \beta_{2j} = \gamma_{20} + u_{2j} \\ \beta_{3j} = \gamma_{30} \\ \beta_{pj} = \gamma_{p0} \end{cases} \tag{A.7}$$

with $\beta_{pj} Covariates_{ij}$ containing age, gender, hospital, and education. In each model, fixed effects were thus included for group (intervention or control) and time (1,2,3), covariate fixed effects for age, gender, educational attainment, and hospital, and random intercept effects for each individual employee.

The fit of these models is tested by several tests for simulated residuals (using Hartig and Hartig, 2017): a Kolmogorov–Smirnov test (KS) to test distributional fit, a test for the presence of zero-inflation, and a dispersion test to verify whether the observed variance is higher or lower than the theoretical variance of the model. A generalised Poisson or negative binomial model is adopted when it is necessary to specify a separate dispersion parameter to help alleviate over- or under-dispersion. We also test whether a large number of outliers occur, and quantiles are compared to their theoretical value. Finally, the presence of autocorrelation is tested with a Durbin-Watson test on uniformly scaled residuals, and the covariance structure is adjusted (if autocorrelation is present).

Some models make use of zero-inflation, which are an extension of the generalised linear mixed model (GLMM) framework, and (just like in generalised linear models) are used if the data has more zero's than can be expected for a distribution. The choice of employing a zero-inflated count data model is made using the zero-inflation test and the Bayesian and Akaike information criteria.

We then make use of post-hoc Tukey-Kramer tests (also called honest significant difference tests) as a robustness test of our estimates.

$$Tukey = \frac{M_i - M_j}{\sqrt{\frac{MS_w}{n_h}}} \tag{A.8}$$

where $M_i - M_j$ is the difference between the estimated means, MS_w is the Mean Square Within, and n is the number of employees in the group. The means are obtained by making predictions on the estimated mixed models (e.g. the mean stress score for the control and intervention group in time 1, 2, and 3) with their standard deviation. The pairwise differences of these six (3 times * 2 groups) means are then compared. Results are averaged over the levels of the covariates (hospital, gender, education), and a p-value adjustment is used by

the Tukey method for comparing a family of 6 estimates. Because the sample sizes are unequal (unbalanced panel data), the Tukey-Kramer adjustment is used to estimate separate standard deviations for each pairwise comparison.

A.2.2 Descriptive statistics by wave

Table A.5: Summary Statistics by wave

Time	1 (N=776)	2 (N=418)	3 (N=588)	Total (N=1782)	p value
Gender					0.883[1]
N-Miss	2	16	27	45	
Male	143 (18.5%)	79 (19.7%)	107 (19.1%)	329 (18.9%)	
Female	631 (81.5%)	323 (80.3%)	454 (80.9%)	1408 (81.1%)	
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Age					0.032[2]
N-Miss	0	15	26	41	
Mean (SD)	45.387 (11.197)	46.757 (11.203)	46.797 (10.686)	46.159 (11.052)	
Range	17.000 - 66.000	18.000 - 65.000	18.000 - 65.000	17.000 - 66.000	
Income					0.003[3]
N-Miss	5	1	34	40	
Income less than 1500 monthly	123 (16.0%)	47 (11.3%)	80 (14.4%)	250 (14.4%)	
Income 1500,00 - 1999,99 monthly	240 (31.1%)	127 (30.5%)	167 (30.1%)	534 (30.7%)	
Income 2000,00 - 2499,99 monthly	241 (31.3%)	127 (30.5%)	186 (33.6%)	554 (31.8%)	
Income 2500,00 - 2999,99 monthly	76 (9.9%)	48 (11.5%)	62 (11.2%)	186 (10.7%)	
Income 3000,00 - 4999,99 monthly	26 (3.4%)	12 (2.9%)	22 (4.0%)	60 (3.4%)	
Income > 5000,00 monthly	5 (0.6%)	3 (0.7%)	2 (0.4%)	10 (0.6%)	
Income not stated	60 (7.8%)	53 (12.7%)	35 (6.3%)	148 (8.5%)	
Nationality					0.463[1]
N-Miss	0	0	17	17	
Belgian	763 (98.3%)	399 (99.0%)	557 (99.1%)	1719 (98.7%)	
Migrant inside EU	9 (1.2%)	3 (0.7%)	5 (0.9%)	17 (1.0%)	
Migrant outside EU	4 (0.5%)	1 (0.2%)	0 (0.0%)	5 (0.3%)	
Education					0.263[3]
N-Miss	0	15	27	42	
No degree	7 (0.9%)	0 (0.0%)	1 (0.2%)	8 (0.5%)	
Primary education	23 (3.0%)	9 (2.2%)	17 (3.0%)	49 (2.8%)	
Secondary education	148 (19.1%)	75 (18.6%)	98 (17.5%)	321 (18.4%)	
Higher education	593 (76.4%)	316 (78.4%)	441 (78.6%)	1350 (77.6%)	
Other education	5 (0.6%)	3 (0.7%)	4 (0.7%)	12 (0.7%)	
Household					0.864[1]
N-Miss	0	15	26	41	
Single	91 (11.7%)	42 (10.4%)	55 (9.8%)	188 (10.8%)	
Single with child(ren)	63 (8.1%)	26 (6.5%)	43 (7.7%)	132 (7.6%)	
Couple with child(ren)	486 (62.6%)	265 (65.8%)	373 (66.4%)	1124 (64.6%)	
Couple without child(ren)	130 (16.8%)	66 (16.4%)	88 (15.7%)	284 (16.3%)	
Other	6 (0.8%)	4 (1.0%)	3 (0.5%)	13 (0.7%)	
Occupation					0.513[1]
N-Miss	27	4	4	35	
Medical Personnel	394 (52.6%)	212 (51.2%)	319 (54.6%)	925 (52.9%)	
Paramedics	20 (2.7%)	20 (4.8%)	24 (4.1%)	64 (3.7%)	
Technicians	56 (7.5%)	35 (8.5%)	40 (6.8%)	131 (7.5%)	
Administrative Personnel	120 (16.0%)	75 (18.1%)	93 (15.9%)	288 (16.5%)	
Management	34 (4.5%)	22 (5.3%)	27 (4.6%)	83 (4.8%)	
Cleaning Staff	34 (4.5%)	12 (2.9%)	24 (4.1%)	70 (4.0%)	
Other	91 (12.1%)	38 (9.2%)	57 (9.8%)	186 (10.6%)	
ESPPW					0.007[1]

LS	197 (25.4%)	101 (24.2%)	175 (29.8%)	473 (26.5%)	
DW	480 (61.9%)	238 (56.9%)	340 (57.8%)	1058 (59.4%)	
ME	99 (12.8%)	79 (18.9%)	73 (12.4%)	251 (14.1%)	
Hospital					< 0.001[1]
Hospital A	197 (25.4%)	101 (24.2%)	175 (29.8%)	473 (26.5%)	
Hospital B	266 (34.3%)	113 (27.0%)	147 (25.0%)	526 (29.5%)	
Hospital C	214 (27.6%)	125 (29.9%)	193 (32.8%)	532 (29.9%)	
Hospital D	99 (12.8%)	79 (18.9%)	73 (12.4%)	251 (14.1%)	
Group					0.027[1]
Non-random control	340 (43.8%)	187 (44.7%)	266 (45.2%)	793 (44.5%)	
Random control	56 (7.2%)	15 (3.6%)	42 (7.1%)	113 (6.3%)	
Non-random intervention	53 (6.8%)	44 (10.5%)	35 (6.0%)	132 (7.4%)	
Random intervention	327 (42.1%)	172 (41.1%)	245 (41.7%)	744 (41.8%)	
Language					0.869[1]
N-Miss	434	169	0	603	
Dutch	330 (96.5%)	240 (96.4%)	569 (96.8%)	1139 (96.6%)	
French	3 (0.9%)	3 (1.2%)	3 (0.5%)	9 (0.8%)	
Other	9 (2.6%)	6 (2.4%)	16 (2.7%)	31 (2.6%)	
Absenteeism					0.067[2]
N-Miss	4	8	11	23	
Mean (SD)	4.807 (31.922)	1.229 (4.720)	1.007 (5.034)	2.727 (21.535)	
Range	0.000 - 526.000	0.000 - 57.000	0.000 - 90.000	0.000 - 526.000	
Prescribed.medication.use..last.2.weeks.					0.612[3]
N-Miss	0	1	1	2	
No	379 (48.8%)	199 (47.7%)	298 (50.8%)	876 (49.2%)	
Yes	397 (51.2%)	218 (52.3%)	289 (49.2%)	904 (50.8%)	
Non.prescribed.medication.use..last.2.weeks.					0.013[3]
N-Miss	1	1	0	2	
No	525 (67.7%)	256 (61.4%)	357 (60.7%)	1138 (63.9%)	
Yes	250 (32.3%)	161 (38.6%)	231 (39.3%)	642 (36.1%)	
Year.of.last.PHS					< 0.001[3]
N-Miss	22	16	46	84	
Never	12 (1.6%)	5 (1.2%)	7 (1.3%)	24 (1.4%)	
2009 or earlier	1 (0.1%)	3 (0.7%)	0 (0.0%)	4 (0.2%)	
2010	0 (0.0%)	1 (0.2%)	0 (0.0%)	1 (0.1%)	
2011	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
2012	2 (0.3%)	0 (0.0%)	0 (0.0%)	2 (0.1%)	
2013	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
2014	4 (0.5%)	0 (0.0%)	0 (0.0%)	4 (0.2%)	
2015	6 (0.8%)	2 (0.5%)	1 (0.2%)	9 (0.5%)	
2016	13 (1.7%)	5 (1.2%)	5 (0.9%)	23 (1.4%)	
2017	88 (11.7%)	17 (4.2%)	15 (2.8%)	120 (7.1%)	
2018	314 (41.6%)	123 (30.6%)	86 (15.9%)	523 (30.8%)	
2019	314 (41.6%)	246 (61.2%)	252 (46.5%)	812 (47.8%)	
2020	0 (0.0%)	0 (0.0%)	176 (32.5%)	176 (10.4%)	
Time.since.last.PMO					< 0.001[1]
N-Miss	34	21	53	108	
0	314 (42.3%)	0 (0.0%)	176 (32.9%)	490 (29.3%)	
1	314 (42.3%)	246 (62.0%)	252 (47.1%)	812 (48.5%)	
2	88 (11.9%)	123 (31.0%)	86 (16.1%)	297 (17.7%)	
3	13 (1.8%)	17 (4.3%)	15 (2.8%)	45 (2.7%)	
4	6 (0.8%)	5 (1.3%)	5 (0.9%)	16 (1.0%)	
5	4 (0.5%)	2 (0.5%)	1 (0.2%)	7 (0.4%)	
7	2 (0.3%)	0 (0.0%)	0 (0.0%)	2 (0.1%)	

10	1 (0.1%)	1 (0.3%)	0 (0.0%)	2 (0.1%)	
11	0 (0.0%)	3 (0.8%)	0 (0.0%)	3 (0.2%)	
Turnover.intention					0.270[2]
N-Miss	8	2	4	14	
Mean (SD)	0.924 (1.155)	0.945 (1.230)	0.841 (1.154)	0.902 (1.173)	
Range	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	
Relation.with.OP					< 0.001[3]
N-Miss	2	0	3	5	
I have not yet met the occupational physician	43 (5.6%)	14 (3.3%)	23 (3.9%)	80 (4.5%)	
Less than 1 year	104 (13.4%)	34 (8.1%)	34 (5.8%)	172 (9.7%)	
1 to 2 years	165 (21.3%)	81 (19.4%)	100 (17.1%)	346 (19.5%)	
3 to 4 years	135 (17.4%)	84 (20.1%)	111 (19.0%)	330 (18.6%)	
5 years or more	327 (42.2%)	205 (49.0%)	317 (54.2%)	849 (47.8%)	
Work.related.HC					0.596[2]
N-Miss	0	0	2	2	
Mean (SD)	1.384 (2.115)	1.299 (2.098)	1.314 (2.074)	1.341 (2.097)	
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	
Referrals.by.OP.to.other.HC					0.016[2]
N-Miss	1	0	2	3	
Mean (SD)	0.263 (1.003)	0.165 (0.801)	0.213 (0.966)	0.224 (0.947)	
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	
Spontaneous.consultations.with.OP					0.124[2]
Mean (SD)	0.232 (0.750)	0.144 (0.557)	0.190 (0.632)	0.198 (0.671)	
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	
Health.Literacy.Score					0.129[2]
N-Miss	25	8	19	52	
Mean (SD)	74.779 (10.407)	74.825 (10.133)	76.030 (11.336)	75.201 (10.669)	
Range	37.500 - 100.000	25.000 - 100.000	25.000 - 100.000	25.000 - 100.000	
Trust.in.Physician.Score					0.048[2]
N-Miss	51	16	29	96	
Mean (SD)	37.851 (6.170)	37.418 (6.341)	38.338 (6.491)	37.909 (6.324)	
Range	14.000 - 55.000	11.000 - 55.000	11.000 - 55.000	11.000 - 55.000	
Stress.score					0.040[2]
Mean (SD)	4.013 (2.528)	4.201 (2.484)	4.366 (2.441)	4.173 (2.493)	
Range	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	
General.Mental.Health					0.100[2]
Mean (SD)	2.067 (2.900)	2.153 (3.030)	2.468 (3.222)	2.219 (3.043)	
Range	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	
Need.for.Recovery					0.279[2]
Mean (SD)	3.772 (2.787)	3.632 (2.659)	3.952 (2.873)	3.799 (2.788)	
Range	0.000 - 11.000	0.000 - 10.000	0.000 - 11.000	0.000 - 11.000	
Sleep.score					0.161[2]
Mean (SD)	5.174 (3.673)	4.998 (3.649)	5.427 (3.731)	5.216 (3.688)	
Range	0.000 - 16.000	0.000 - 16.000	0.000 - 16.000	0.000 - 16.000	
Role.conflicts					0.976[2]
Mean (SD)	3.620 (2.548)	3.651 (2.575)	3.612 (2.586)	3.625 (2.566)	
Range	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000	
Job.satisfaction					0.923[2]
N-Miss	2	0	0	2	
Mean (SD)	15.235 (2.377)	15.187 (2.483)	15.224 (2.317)	15.220 (2.382)	
Range	6.000 - 20.000	7.000 - 20.000	7.000 - 20.000	6.000 - 20.000	
Musculoskeletal.Functioning					< 0.001[2]
Mean (SD)	1.146 (1.715)	1.019 (1.689)	2.386 (2.168)	1.525 (1.966)	

Range	0.000 - 9.000	0.000 - 9.000	0.000 - 9.000	0.000 - 9.000	
EQ5D VAS					0.483[2]
Mean (SD)	79.178 (13.079)	77.794 (14.210)	78.198 (13.427)	78.530 (13.471)	
Range	2.000 - 100.000	8.000 - 100.000	20.000 - 100.000	2.000 - 100.000	

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01
1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test 3. Trend test for ordinal variables

Table A.6: Summary Statistics by wave, complete cases

Time	**1 (N=208)**	**2 (N=208)**	**3 (N=208)**	**Total (N=624)**
Gender				
N-Miss	1	1	1	3
Male	49 (23.7%)	49 (23.7%)	49 (23.7%)	147 (23.7%)
Female	158 (76.3%)	158 (76.3%)	158 (76.3%)	474 (76.3%)
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Age				
Mean (SD)	47.553 (10.454)	48.553 (10.454)	48.553 (10.454)	48.220 (10.448)
Range	17.000 - 63.000	18.000 - 64.000	18.000 - 64.000	17.000 - 64.000
Income				
Income less than 1500 monthly	25 (12.0%)	25 (12.0%)	25 (12.0%)	75 (12.0%)
Income 1500,00 - 1999,99 monthly	54 (26.0%)	54 (26.0%)	54 (26.0%)	162 (26.0%)
Income 2000,00 - 2499,99 monthly	74 (35.6%)	74 (35.6%)	74 (35.6%)	222 (35.6%)
Income 2500,00 - 2999,99 monthly	30 (14.4%)	30 (14.4%)	30 (14.4%)	90 (14.4%)
Income 3000,00 - 4999,99 monthly	7 (3.4%)	7 (3.4%)	7 (3.4%)	21 (3.4%)
Income $>$ 5000,00 monthly	1 (0.5%)	1 (0.5%)	1 (0.5%)	3 (0.5%)
Income not stated	17 (8.2%)	17 (8.2%)	17 (8.2%)	51 (8.2%)
Nationality				
Belgian	207 (99.5%)	207 (99.5%)	207 (99.5%)	621 (99.5%)
Migrant inside EU	1 (0.5%)	1 (0.5%)	1 (0.5%)	3 (0.5%)
Migrant outside EU	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Education				
No degree	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Primary education	5 (2.4%)	5 (2.4%)	5 (2.4%)	15 (2.4%)
Secondary education	38 (18.3%)	38 (18.3%)	38 (18.3%)	114 (18.3%)
Higher education	165 (79.3%)	165 (79.3%)	165 (79.3%)	495 (79.3%)
Other education	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Household				
Single	25 (12.0%)	25 (12.0%)	25 (12.0%)	75 (12.0%)
Single with child(ren)	17 (8.2%)	17 (8.2%)	17 (8.2%)	51 (8.2%)
Couple with child(ren)	138 (66.3%)	138 (66.3%)	138 (66.3%)	414 (66.3%)
Couple without child(ren)	27 (13.0%)	27 (13.0%)	27 (13.0%)	81 (13.0%)
Other	1 (0.5%)	1 (0.5%)	1 (0.5%)	3 (0.5%)
Occupation				
N-Miss	1	1	1	3
Medical Personnel	101 (48.8%)	101 (48.8%)	101 (48.8%)	303 (48.8%)
Paramedics	9 (4.3%)	9 (4.3%)	9 (4.3%)	27 (4.3%)
Technicians	21 (10.1%)	21 (10.1%)	21 (10.1%)	63 (10.1%)
Administrative Personnel	41 (19.8%)	41 (19.8%)	41 (19.8%)	123 (19.8%)
Management	12 (5.8%)	12 (5.8%)	12 (5.8%)	36 (5.8%)
Cleaning Staff	6 (2.9%)	6 (2.9%)	6 (2.9%)	18 (2.9%)
Other	17 (8.2%)	17 (8.2%)	17 (8.2%)	51 (8.2%)
ESPPW				

LS	22 (10.6%)	22 (10.6%)	22 (10.6%)	66 (10.6%)
DW	151 (72.6%)	151 (72.6%)	151 (72.6%)	453 (72.6%)
ME	35 (16.8%)	35 (16.8%)	35 (16.8%)	105 (16.8%)
Hospital				
Hospital A	22 (10.6%)	22 (10.6%)	22 (10.6%)	66 (10.6%)
Hospital B	71 (34.1%)	71 (34.1%)	71 (34.1%)	213 (34.1%)
Hospital C	80 (38.5%)	80 (38.5%)	80 (38.5%)	240 (38.5%)
Hospital D	35 (16.8%)	35 (16.8%)	35 (16.8%)	105 (16.8%)
Group				
Non-random control	93 (44.7%)	93 (44.7%)	93 (44.7%)	279 (44.7%)
Random control	7 (3.4%)	7 (3.4%)	7 (3.4%)	21 (3.4%)
Non-random intervention	18 (8.7%)	18 (8.7%)	18 (8.7%)	54 (8.7%)
Random intervention	90 (43.3%)	90 (43.3%)	90 (43.3%)	270 (43.3%)
Language				
Dutch	200 (96.2%)	200 (96.2%)	200 (96.2%)	600 (96.2%)
French	3 (1.4%)	3 (1.4%)	3 (1.4%)	9 (1.4%)
Other	5 (2.4%)	5 (2.4%)	5 (2.4%)	15 (2.4%)
Absenteeism				
N-Miss	1	2	3	6
Mean (SD)	1.130 (7.055)	1.083 (3.867)	1.049 (4.141)	1.087 (5.221)
Range	0.000 - 63.000	0.000 - 28.000	0.000 - 30.000	0.000 - 63.000
Absenteeism..28				
N-Miss	5	2	5	12
Mean (SD)	0.192 (1.349)	1.083 (3.867)	0.764 (2.989)	0.681 (2.951)
Range	0.000 - 14.000	0.000 - 28.000	0.000 - 21.000	0.000 - 28.000
Prescribed medication use last 2 weeks				
N-Miss	0	1	0	1
No	97 (46.6%)	93 (44.9%)	97 (46.6%)	287 (46.1%)
Yes	111 (53.4%)	114 (55.1%)	111 (53.4%)	336 (53.9%)
Non-prescribed medication use last 2 weeks				
N-Miss	0	1	0	1
No	144 (69.2%)	129 (62.3%)	133 (63.9%)	406 (65.2%)
Yes	64 (30.8%)	78 (37.7%)	75 (36.1%)	217 (34.8%)
Year.of.last.PHS				
N-Miss	2	9	21	32
Never	1 (0.5%)	1 (0.5%)	1 (0.5%)	3 (0.5%)
2009 or earlier	0 (0.0%)	3 (1.5%)	0 (0.0%)	3 (0.5%)
2010	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2011	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2012	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2013	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2014	1 (0.5%)	0 (0.0%)	0 (0.0%)	1 (0.2%)
2015	1 (0.5%)	0 (0.0%)	0 (0.0%)	1 (0.2%)
2016	2 (1.0%)	2 (1.0%)	2 (1.1%)	6 (1.0%)
2017	23 (11.2%)	11 (5.5%)	8 (4.3%)	42 (7.1%)
2018	90 (43.7%)	72 (36.2%)	35 (18.7%)	197 (33.3%)
2019	88 (42.7%)	110 (55.3%)	82 (43.9%)	280 (47.3%)
2020	0 (0.0%)	0 (0.0%)	59 (31.6%)	59 (10.0%)
Time.since.last.PMO				
N-Miss	3	10	22	35
0	88 (42.9%)	0 (0.0%)	59 (31.7%)	147 (25.0%)
1	90 (43.9%)	110 (55.6%)	82 (44.1%)	282 (47.9%)
2	23 (11.2%)	72 (36.4%)	35 (18.8%)	130 (22.1%)
3	2 (1.0%)	11 (5.6%)	8 (4.3%)	21 (3.6%)

4	1 (0.5%)	2 (1.0%)	2 (1.1%)	5 (0.8%)
5	1 (0.5%)	0 (0.0%)	0 (0.0%)	1 (0.2%)
7	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
10	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
11	0 (0.0%)	3 (1.5%)	0 (0.0%)	3 (0.5%)
Turnover.intention				
N-Miss	0	1	0	1
Mean (SD)	0.846 (1.106)	0.928 (1.242)	0.678 (1.006)	0.817 (1.125)
Range	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000
Relation.with.OP				
I did not yet meet the OP	11 (5.3%)	7 (3.4%)	8 (3.8%)	26 (4.2%)
< 1 year	29 (13.9%)	16 (7.7%)	14 (6.7%)	59 (9.5%)
1 - 2 years	41 (19.7%)	47 (22.6%)	41 (19.7%)	129 (20.7%)
3 - 4 years	41 (19.7%)	50 (24.0%)	45 (21.6%)	136 (21.8%)
5 years or more	86 (41.3%)	88 (42.3%)	100 (48.1%)	274 (43.9%)
Work.related.HC				
Mean (SD)	1.202 (2.028)	1.231 (2.068)	0.957 (1.853)	1.130 (1.986)
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000
Referrals.by.OP.to.other.HC				
N-Miss	0	0	1	1
Mean (SD)	0.202 (0.883)	0.144 (0.721)	0.198 (0.947)	0.181 (0.855)
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000
Spontaneous.consultations.with.OP				
Mean (SD)	0.188 (0.735)	0.149 (0.575)	0.168 (0.692)	0.168 (0.670)
Range	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000	0.000 - 6.000
Health.Literacy.Score				
N-Miss	4	1	4	9
Mean (SD)	76.884 (10.825)	75.853 (9.335)	77.788 (10.828)	76.837 (10.362)
Range	37.500 - 100.000	39.062 - 100.000	45.312 - 100.000	37.500 - 100.000
Trust.in.Physician.Score				
N-Miss	13	8	8	29
Mean (SD)	37.554 (6.544)	37.700 (6.439)	38.740 (6.773)	38.002 (6.597)
Range	14.000 - 55.000	13.000 - 55.000	11.000 - 55.000	11.000 - 55.000
Stress.score				
Mean (SD)	3.635 (2.353)	3.995 (2.340)	3.966 (2.343)	3.865 (2.347)
Range	0.000 - 10.000	0.000 - 9.000	0.000 - 10.000	0.000 - 10.000
General.Mental.Health				
Mean (SD)	1.692 (2.601)	1.986 (2.913)	2.260 (3.221)	1.979 (2.927)
Range	0.000 - 10.000	0.000 - 12.000	0.000 - 12.000	0.000 - 12.000
Need.for.Recovery				
Mean (SD)	3.553 (2.764)	3.606 (2.693)	3.615 (2.785)	3.591 (2.743)
Range	0.000 - 10.000	0.000 - 10.000	0.000 - 11.000	0.000 - 11.000
Sleep.score				
Mean (SD)	4.885 (3.672)	4.817 (3.437)	5.101 (3.593)	4.934 (3.565)
Range	0.000 - 16.000	0.000 - 16.000	0.000 - 16.000	0.000 - 16.000
Role.conflicts				
Mean (SD)	3.529 (2.582)	3.562 (2.490)	3.433 (2.460)	3.508 (2.508)
Range	0.000 - 12.000	0.000 - 10.000	0.000 - 11.000	0.000 - 12.000
Job.satisfaction				
Mean (SD)	15.284 (2.229)	15.115 (2.415)	15.149 (2.337)	15.183 (2.326)
Range	8.000 - 20.000	7.000 - 20.000	7.000 - 20.000	7.000 - 20.000
Musculoskeletal.Functioning				
Mean (SD)	0.981 (1.588)	0.909 (1.509)	2.178 (2.152)	1.356 (1.863)
Range	0.000 - 9.000	0.000 - 9.000	0.000 - 9.000	0.000 - 9.000

1. Pearson's Chi-squared test; 2. Kruskal-Wallis rank sum test; 3. Trend test for ordinal variables

A.2.3 Descriptive statistics by group

Table A.7: Summary Statistics by group (Wave 1)

	control (N=396)	intervention (N=380)	Total (N=776)	p value
Gender				0.352[1]
N-Miss	1	1	2	
Male	78 (19.7%)	65 (17.2%)	143 (18.5%)	
Female	317 (80.3%)	314 (82.8%)	631 (81.5%)	
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Age				0.535[2]
Mean (SD)	45.634 (11.215)	45.129 (11.188)	45.387 (11.197)	
Range	22.000 - 66.000	17.000 - 64.000	17.000 - 66.000	
Income				0.801[3]
N-Miss	3	2	5	
Income less than 1500 monthly	69 (17.6%)	54 (14.3%)	123 (16.0%)	
Income 1500,00 - 1999,99 monthly	117 (29.8%)	123 (32.5%)	240 (31.1%)	
Income 2000,00 - 2499,99 monthly	124 (31.6%)	117 (31.0%)	241 (31.3%)	
Income 2500,00 - 2999,99 monthly	34 (8.7%)	42 (11.1%)	76 (9.9%)	
Income 3000,00 - 4999,99 monthly	15 (3.8%)	11 (2.9%)	26 (3.4%)	
Income > 5000,00 monthly	3 (0.8%)	2 (0.5%)	5 (0.6%)	
Income not stated	31 (7.9%)	29 (7.7%)	60 (7.8%)	
Nationality				0.640[1]
Belgian	388 (98.0%)	375 (98.7%)	763 (98.3%)	
Migrant inside EU	6 (1.5%)	3 (0.8%)	9 (1.2%)	
Migrant outside EU	2 (0.5%)	2 (0.5%)	4 (0.5%)	
Education				0.061[3]
No degree	3 (0.8%)	4 (1.1%)	7 (0.9%)	
Primary education	14 (3.5%)	9 (2.4%)	23 (3.0%)	
Secondary education	85 (21.5%)	63 (16.6%)	148 (19.1%)	
Higher education	294 (74.2%)	299 (78.7%)	593 (76.4%)	
Other education	0 (0.0%)	5 (1.3%)	5 (0.6%)	
Household				0.640[1]
Single	41 (10.4%)	50 (13.2%)	91 (11.7%)	
Single with child(ren)	34 (8.6%)	29 (7.6%)	63 (8.1%)	
Couple with child(ren)	247 (62.4%)	239 (62.9%)	486 (62.6%)	
Couple without child(ren)	70 (17.7%)	60 (15.8%)	130 (16.8%)	
Other	4 (1.0%)	2 (0.5%)	6 (0.8%)	
Occupation				0.048[1]
N-Miss	17	10	27	
Medical Personnel	185 (48.8%)	209 (56.5%)	394 (52.6%)	
Paramedics	12 (3.2%)	8 (2.2%)	20 (2.7%)	
Technicians	31 (8.2%)	25 (6.8%)	56 (7.5%)	
Administrative Personnel	55 (14.5%)	65 (17.6%)	120 (16.0%)	
Management	18 (4.7%)	16 (4.3%)	34 (4.5%)	
Cleaning Staff	24 (6.3%)	10 (2.7%)	34 (4.5%)	
Other	54 (14.2%)	37 (10.0%)	91 (12.1%)	
ESPPW				0.035[1]
LS	116 (29.3%)	81 (21.3%)	197 (25.4%)	
DW	234 (59.1%)	246 (64.7%)	480 (61.9%)	
ME	46 (11.6%)	53 (13.9%)	99 (12.8%)	
Hospital				0.012[1]
Hospital A	116 (29.3%)	81 (21.3%)	197 (25.4%)	
Hospital B	141 (35.6%)	125 (32.9%)	266 (34.3%)	
Hospital C	93 (23.5%)	121 (31.8%)	214 (27.6%)	
Hospital D	46 (11.6%)	53 (13.9%)	99 (12.8%)	
Group				< 0.001[1]
Non-random control	340 (85.9%)	0 (0.0%)	340 (43.8%)	
Random control	56 (14.1%)	0 (0.0%)	56 (7.2%)	
Non-random intervention	0 (0.0%)	53 (13.9%)	53 (6.8%)	
Random intervention	0 (0.0%)	327 (86.1%)	327 (42.1%)	

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01.
1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test 3. Trend test for ordinal variables

Table A.8: Summary Statistics by group (Wave 2)

	control (N=202)	intervention (N=216)	Total (N=418)	p value
Gender				0.638[1]
N-Miss	8	8	16	
Male	40 (20.6%)	39 (18.8%)	79 (19.7%)	
Female	154 (79.4%)	169 (81.2%)	323 (80.3%)	
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Age				0.842[2]
N-Miss	7	8	15	
Mean (SD)	46.949 (11.087)	46.577 (11.333)	46.757 (11.203)	
Range	23.000 - 65.000	18.000 - 64.000	18.000 - 65.000	
Income				0.527[3]
N-Miss	1	0	1	
Income less than 1500 monthly	23 (11.4%)	24 (11.1%)	47 (11.3%)	
Income 1500,00 - 1999,99 monthly	63 (31.3%)	64 (29.6%)	127 (30.5%)	
Income 2000,00 - 2499,99 monthly	58 (28.9%)	69 (31.9%)	127 (30.5%)	
Income 2500,00 - 2999,99 monthly	19 (9.5%)	29 (13.4%)	48 (11.5%)	
Income 3000,00 - 4999,99 monthly	7 (3.5%)	5 (2.3%)	12 (2.9%)	
Income > 5000,00 monthly	2 (1.0%)	1 (0.5%)	3 (0.7%)	
Income not stated	29 (14.4%)	24 (11.1%)	53 (12.7%)	
Nationality				0.477[1]
Belgian	192 (98.5%)	207 (99.5%)	399 (99.0%)	
Migrant inside EU	2 (1.0%)	1 (0.5%)	3 (0.7%)	
Migrant outside EU	1 (0.5%)	0 (0.0%)	1 (0.2%)	
Education				0.183[3]
N-Miss	7	8	15	
No degree	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Primary education	4 (2.1%)	5 (2.4%)	9 (2.2%)	
Secondary education	43 (22.1%)	32 (15.4%)	75 (18.6%)	
Higher education	147 (75.4%)	169 (81.2%)	316 (78.4%)	
Other education	1 (0.5%)	2 (1.0%)	3 (0.7%)	
Household				0.270[1]
N-Miss	7	8	15	
Single	18 (9.2%)	24 (11.5%)	42 (10.4%)	
Single with child(ren)	18 (9.2%)	8 (3.8%)	26 (6.5%)	
Couple with child(ren)	126 (64.6%)	139 (66.8%)	265 (65.8%)	
Couple without child(ren)	31 (15.9%)	35 (16.8%)	66 (16.4%)	
Other	2 (1.0%)	2 (1.0%)	4 (1.0%)	
Occupation				0.119[1]
N-Miss	2	2	4	
Medical Personnel	89 (44.5%)	123 (57.5%)	212 (51.2%)	
Paramedics	12 (6.0%)	8 (3.7%)	20 (4.8%)	
Technicians	20 (10.0%)	15 (7.0%)	35 (8.5%)	
Administrative Personnel	40 (20.0%)	35 (16.4%)	75 (18.1%)	
Management	12 (6.0%)	10 (4.7%)	22 (5.3%)	
Cleaning Staff	9 (4.5%)	3 (1.4%)	12 (2.9%)	
Other	18 (9.0%)	20 (9.3%)	38 (9.2%)	
ESPPW				0.542[1]
LS	53 (26.2%)	48 (22.2%)	101 (24.2%)	
DW	114 (56.4%)	124 (57.4%)	238 (56.9%)	
ME	35 (17.3%)	44 (20.4%)	79 (18.9%)	
Hospital				0.691[1]
Hospital A	53 (26.2%)	48 (22.2%)	101 (24.2%)	
Hospital B	56 (27.7%)	57 (26.4%)	113 (27.0%)	
Hospital C	58 (28.7%)	67 (31.0%)	125 (29.9%)	
Hospital D	35 (17.3%)	44 (20.4%)	79 (18.9%)	
Group				< 0.001[1]
Non-random control	187 (92.6%)	0 (0.0%)	187 (44.7%)	
Random control	15 (7.4%)	0 (0.0%)	15 (3.6%)	
Non-random intervention	0 (0.0%)	44 (20.4%)	44 (10.5%)	
Random intervention	0 (0.0%)	172 (79.6%)	172 (41.1%)	

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01.
1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test 3. Trend test for ordinal variables

Table A.9: Summary Statistics by group (Wave 3)

	control (N=308)	intervention (N=280)	Total (N=588)	p value
Gender				0.810[1]
N-Miss	15	12	27	
Male	57 (19.5%)	50 (18.7%)	107 (19.1%)	
Female	236 (80.5%)	218 (81.3%)	454 (80.9%)	
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Age				1.000[2]
N-Miss	14	12	26	
Mean (SD)	46.789 (10.788)	46.806 (10.595)	46.797 (10.686)	
Range	23.000 - 65.000	18.000 - 64.000	18.000 - 65.000	
Income				0.681[3]
N-Miss	19	15	34	
Income less than 1500 monthly	41 (14.2%)	39 (14.7%)	80 (14.4%)	
Income 1500,00 - 1999,99 monthly	88 (30.4%)	79 (29.8%)	167 (30.1%)	
Income 2000,00 - 2499,99 monthly	99 (34.3%)	87 (32.8%)	186 (33.6%)	
Income 2500,00 - 2999,99 monthly	28 (9.7%)	34 (12.8%)	62 (11.2%)	
Income 3000,00 - 4999,99 monthly	11 (3.8%)	11 (4.2%)	22 (4.0%)	
Income > 5000,00 monthly	0 (0.0%)	2 (0.8%)	2 (0.4%)	
Income not stated	22 (7.6%)	13 (4.9%)	35 (6.3%)	
Nationality				0.580[1]
N-Miss	9	8	17	
Belgian	292 (99.3%)	265 (98.9%)	557 (99.1%)	
Migrant inside EU	2 (0.7%)	3 (1.1%)	5 (0.9%)	
Migrant outside EU	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Education				0.380[3]
N-Miss	14	13	27	
No degree	1 (0.3%)	0 (0.0%)	1 (0.2%)	
Primary education	7 (2.4%)	10 (3.7%)	17 (3.0%)	
Secondary education	58 (19.7%)	40 (15.0%)	98 (17.5%)	
Higher education	227 (77.2%)	214 (80.1%)	441 (78.6%)	
Other education	1 (0.3%)	3 (1.1%)	4 (0.7%)	
Household				0.298[1]
N-Miss	14	12	26	
Single	26 (8.8%)	29 (10.8%)	55 (9.8%)	
Single with child(ren)	28 (9.5%)	15 (5.6%)	43 (7.7%)	
Couple with child(ren)	188 (63.9%)	185 (69.0%)	373 (66.4%)	
Couple without child(ren)	50 (17.0%)	38 (14.2%)	88 (15.7%)	
Other	2 (0.7%)	1 (0.4%)	3 (0.5%)	
Occupation				0.203[1]
N-Miss	1	3	4	
Medical Personnel	157 (51.1%)	162 (58.5%)	319 (54.6%)	
Paramedics	13 (4.2%)	11 (4.0%)	24 (4.1%)	
Technicians	28 (9.1%)	12 (4.3%)	40 (6.8%)	
Administrative Personnel	50 (16.3%)	43 (15.5%)	93 (15.9%)	
Management	13 (4.2%)	14 (5.1%)	27 (4.6%)	
Cleaning Staff	16 (5.2%)	8 (2.9%)	24 (4.1%)	
Other	30 (9.8%)	27 (9.7%)	57 (9.8%)	
ESPPW				0.226[1]
LS	101 (32.8%)	74 (26.4%)	175 (29.8%)	
DW	169 (54.9%)	171 (61.1%)	340 (57.8%)	
ME	38 (12.3%)	35 (12.5%)	73 (12.4%)	
Hospital				0.112[1]
Hospital A	101 (32.8%)	74 (26.4%)	175 (29.8%)	
Hospital B	81 (26.3%)	66 (23.6%)	147 (25.0%)	
Hospital C	88 (28.6%)	105 (37.5%)	193 (32.8%)	
Hospital D	38 (12.3%)	35 (12.5%)	73 (12.4%)	
Group				< 0.001[1]
Non-random control	266 (86.4%)	0 (0.0%)	266 (45.2%)	
Random control	42 (13.6%)	0 (0.0%)	42 (7.1%)	
Non-random intervention	0 (0.0%)	35 (12.5%)	35 (6.0%)	
Random intervention	0 (0.0%)	245 (87.5%)	245 (41.7%)	

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01.
1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test 3. Trend test for ordinal variables

A.2.4 Sample quality

Table A.10: Comparison of sample estimates in time 1 for hospital B C, and D to population data from hospital C, B, and D

Variable	Population B-C (n=1402)	Sample B-C (n=480)	Population D (n=862)	Sample D (n=99)
Gender				
Missing	2.2%	0.2%	0.0%	0.0%
Male	17.8%	20.0%	19.8%	16.16%
Female	80.0%	79.8%	80.2%	83.83%
Age			44.48(12.20)	47.46(11.15)
Missing	2.2%	0.0%		
<25	4.6%	2.1%		
25 - 34	23.7%	21.0%		
35 - 44	22.5%	21.5%		
45 - 54	25.1%	29.2%		
>55	22.0%	26.3%		

A.2.5 Full mixed models, including covariates

Table A.11: Regressions mixed methods: group + time + interaction, and covariates

Dependent Var Range	EQ5D vas (0-100)	Musculo (log) (1-10)	Stress (0-12)	Burnout (0-16)	Sleep Problems (0-16)	GHQ(log) (1-13)	NFR(log) (1-12)
Intercept	-8.19 (-8.23– -8.15)	0.54 *** (0.36 – 0.73)	2.96 *** (2.26 – 3.66)	4.07 *** (3.04 – 5.10)	4.63 *** (3.59 – 5.67)	0.56 *** (0.34 – 0.79)	1.12 *** (0.94 – 1.30)
Intervention group (ref.=control)	0.16 (0.15– 0.17)	-0.06 (-0.15 – 0.02)	0.02 (-0.32 – 0.36)	0.13 (-0.36 – 0.63)	0.36 (-0.14 – 0.86)	0.07 (-0.04 – 0.19)	0.00 (-0.09 – 0.09)
Time 2 (ref.=Time1)	0.3 (0.29– 0.31)	-0.12 *** (-0.20 – -0.04)	0.46 *** (0.16 – 0.77)	0.59 *** (0.15 – 1.03)	0.07 (-0.34 – 0.48)	0.08 (-0.04 – 0.19)	0.04 (-0.03 – 0.12)
Time 3	0.26 (0.25– 0.26)	0.45 *** (0.36 – 0.54)	0.48 *** (0.19 – 0.77)	0.69 *** (0.25 – 1.13)	0.53 ** (0.08 – 0.98)	0.17 *** (0.05 – 0.28)	0.08 ** (0.00 – 0.15)
Hospital B(ref. = hosp A)	-0.65 (-0.66– -0.64)	-0.08 (-0.17 – 0.02)	-0.38 ** (-0.74 – -0.01)	-0.81 *** (-1.35 – -0.27)	-0.82 *** (-1.37 – -0.27)	-0.23 *** (-0.35 – -0.11)	-0.11 ** (-0.21 – -0.02)
Hospital C	-0.51 (-0.52– -0.5)	-0.11 ** (-0.20 – -0.01)	-0.16 (-0.53 – 0.21)	-0.41 (-0.96 – 0.13)	-0.49 * (-1.05 – 0.07)	-0.11 * (-0.23 – 0.01)	-0.06 (-0.16 – 0.03)
Hospital D	-0.61 (-0.62– -0.6)	-0.14 ** (-0.26 – -0.02)	-0.32 (-0.77 – 0.13)	-0.42 (-1.08 – 0.25)	0.02 (-0.66 – 0.70)	-0.07 (-0.21 – 0.08)	-0.03 (-0.15 – 0.08)
Age <25 (ref. = Age > 55)	0.06 (0.04– 0.09)	-0.07 (-0.31 – 0.17)	0.11 (-0.81 – 1.03)	0.46 (-0.89 – 1.81)	-0.08 (-1.42 – 1.27)	-0.11 (-0.41 – 0.20)	0.06 (-0.17 – 0.30)
Age 25 - 34	-0.02 (-0.03– -0.01)	-0.05 (-0.16 – 0.06)	1.05 *** (0.63 – 1.47)	1.43 *** (0.81 – 2.05)	-0.40 (-1.02 – 0.23)	0.29 *** (0.15 – 0.42)	0.13 ** (0.03 – 0.24)
Age 35 - 44	-0.24 (-0.25– -0.23)	-0.05 (-0.15 – 0.05)	0.54 *** (0.14 – 0.93)	0.76 ** (0.18 – 1.34)	0.04 (-0.54 – 0.63)	0.08 (-0.05 – 0.21)	0.10 ** (0.00 – 0.20)
Age 45 - 54	-0.03 (-0.04– -0.03)	0.01 (-0.08 – 0.10)	0.18 (-0.17 – 0.53)	0.43 (-0.09 – 0.95)	0.20 (-0.32 – 0.72)	0.06 (-0.06 – 0.17)	0.06 (-0.03 – 0.15)
Female (ref. = Male)	0.22 (0.22– 0.23)	0.18 *** (0.08 – 0.27)	0.50 *** (0.14 – 0.86)	0.93 *** (0.40 – 1.47)	0.66 ** (0.12 – 1.20)	0.10 (-0.02 – 0.21)	0.08 (-0.02 – 0.17)
Education (linear)	-0.18 (-0.22– -0.14)	-0.12 (-0.53 – 0.29)	1.50 * (-0.09 – 3.08)	0.30 (-2.04 – 2.63)	1.39 (-0.98 – 3.75)	-0.50 * (-1.01 – 0.02)	0.20 (-0.21 – 0.61)
Education (quadratic)	-1.08 (-1.11– -1.05)	-0.14 (-0.49 – 0.21)	-1.42 ** (-2.77 – -0.07)	-2.47 ** (-4.46 – -0.49)	-1.47 (-3.49 – 0.54)	-0.40 * (-0.84 – 0.04)	-0.42 ** (-0.76 – -0.07)
Education (cubic)	0.84 (0.82– 0.87)	0.09 (-0.15 – 0.33)	0.99 ** (0.06 – 1.92)	0.57 (-0.81 – 1.94)	1.58 ** (0.19 – 2.97)	-0.08 (-0.39 – 0.22)	-0.01 (-0.25 – 0.23)
Education (4th degree)	-0.57 (-0.58– -0.55)	0.02 (-0.12 – 0.16)	-0.62 ** (-1.16 – -0.09)	-0.91 ** (-1.70 – -0.13)	-0.81 ** (-1.60 – -0.01)	-0.19 ** (-0.36 – -0.01)	-0.02 (-0.15 – 0.12)
intervention:time2	-0.05 (-0.06– -0.04)	0.13 ** (0.02 – 0.24)	-0.29 (-0.72 – 0.14)	-0.25 (-0.87 – 0.36)	-0.27 (-0.84 – 0.30)	-0.10 (-0.26 – 0.06)	-0.08 (-0.18 – 0.03)
intervention:time3	-0.06 (-0.07– -0.05)	0.01 (-0.12 – 0.14)	-0.21 (-0.62 – 0.20)	-0.57 * (-1.20 – 0.06)	-0.53 (-1.18 – 0.11)	-0.15 * (-0.32 – 0.01)	-0.09 * (-0.20 – 0.02)
Random Effects							
σ^2	3.144	0.03	0.43	0.94	0.91	0.06	0.03
τ_{00}	1.765 ID	0.36 ID	5.73 ID	12.10 ID	12.30 ID	0.60 ID	0.37 ID
τ_{11}		0.28 ID.time2 0.50 ID.time3	4.23 ID.time2 4.66 ID.time3	8.43 ID.time2 11.21 ID.time3	6.75 ID.time2 11.83 ID.time3	0.64 ID.time2 0.90 ID.time3	0.23 ID.time2 0.31 ID.time3
ρ_{01}		-0.45 -0.47	-0.44 -0.49	-0.43 -0.47	-0.39 -0.46	-0.5 -0.57	-0.45 -0.44
ICC		0.92	0.93	0.93	0.93	0.91	0.93
N		1042 ID	1042 ID	1042 ID	1042 ID	1042 ID	1042 ID

Observations	1037	1736	1736	1736	1736	1736	1736
Marginal R2/ Conditional R2		0.132 / 0.932	0.048 / 0.932	0.046 / 0.931	0.025 / 0.934	0.037 / 0.916	0.023 / 0.930

Note: *p<0.1; **p<0.05; *** p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept ; τ_{11} = variance of random slope; ρ_{01} = correlation between random intercept and slope

Table A.12: Regressions mixed methods: group + time + interaction, and covariates

Dependent Var	Health literacy	Trust in physician	Job Satisfaction	Role conflicts	Turnover intention	Presenteeism weighted
Range	(0-100)	(0-55)	(0-20)	(0-12)	(0-4)	(0-14)
Intercept	76.25 ***	40.59 ***	15.53 ***	3.65 ***	0.37 **	0.70 ***
	(73.27 – 79.23)	(38.85 – 42.33)	(14.86 – 16.20)	(2.95 – 4.36)	(0.05 – 0.68)	(0.46 – 0.93)
Intervention group (ref.=control)	-0.97	-0.94 **	-0.07	0.30 *	-0.02	0.04
	(-2.43 – 0.49)	(-1.78 – -0.10)	(-0.38 – 0.25)	(-0.04 – 0.64)	(-0.17 – 0.13)	(-0.07 – 0.16)
Time 2 (ref.=Time1)	-1.31 *	-0.32	-0.09	0.16	0.19 ***	-0.05
	(-2.62 – -0.00)	(-0.99 – 0.36)	(-0.35 – 0.17)	(-0.13 – 0.46)	(0.05 – 0.32)	(-0.16 – 0.07)
Time 3	1.22 *	-0.04	-0.06	0.13	-0.15 **	0.02
	(-0.19 – 2.63)	(-0.79 – 0.72)	(-0.33 – 0.21)	(-0.17 – 0.43)	(-0.28 – -0.02)	(-0.10 – 0.14)
Hospital B(ref. = hosp A)	0.49	-0.16	0.31 *	-0.79 ***	-0.16 *	-0.05
	(-1.10 – 2.07)	(-1.09 – 0.77)	(-0.04 – 0.67)	(-1.16 – -0.42)	(-0.33 – 0.00)	(-0.17 – 0.07)
Hospital C	-0.03	0.53	0.14	-0.02	0.19 **	-0.02
	(-1.63 – 1.57)	(-0.40 – 1.47)	(-0.22 – 0.50)	(-0.39 – 0.36)	(0.02 – 0.35)	(-0.14 – 0.10)
Hospital D	-3.36 ***	-4.77 ***	0.24	-0.20	-0.03	0.03
	(-5.31 – -1.41)	(-5.89 – -3.64)	(-0.20 – 0.68)	(-0.66 – 0.25)	(-0.24 – 0.18)	(-0.12 – 0.19)
Age <25(ref. = Age > 55)	-5.44 ***	-2.55 **	0.34	0.91 *	0.47 **	0.17
	(-9.45 – -1.43)	(-4.83 – -0.28)	(-0.51 – 1.20)	(-0.01 – 1.83)	(0.05 – 0.89)	(-0.13 – 0.47)
Age 25 - 34	-1.29	-0.98 *	-0.44 **	0.99 ***	0.84 ***	0.11
	(-3.08 – 0.50)	(-2.03 – 0.07)	(-0.85 – -0.04)	(0.57 – 1.41)	(0.65 – 1.03)	(-0.03 – 0.25)
Age 35 - 44	-0.55	-0.55	-0.26	0.63 ***	0.65 ***	0.12 *
	(-2.26 – 1.15)	(-1.54 – 0.44)	(-0.63 – 0.12)	(0.23 – 1.03)	(0.47 – 0.83)	(-0.02 – 0.25)
Age 45 - 54	-0.11	0.08	-0.28 *	-0.11	0.48 ***	0.09
	(-1.63 – 1.41)	(-0.79 – 0.95)	(-0.61 – 0.05)	(-0.46 – 0.24)	(0.32 – 0.64)	(-0.04 – 0.22)
Female(ref. = Male)	-0.74	-0.69	0.53 ***	-0.46 **	-0.23 ***	-0.05
	(-2.28 – 0.80)	(-1.60 – 0.22)	(0.18 – 0.88)	(-0.82 – -0.09)	(-0.40 – -0.07)	(-0.18 – 0.08)
Education (linear)	1.60	1.43	0.55	-1.68 **	-0.23	-0.47 *
	(-5.16 – 8.35)	(-2.51 – 5.37)	(-0.97 – 2.08)	(-3.28 – -0.07)	(-0.95 – 0.49)	(-0.95 – 0.00)
Education (quadratic)	1.27	1.75	0.83	-0.10	-0.54 *	0.01
	(-4.48 – 7.03)	(-1.61 – 5.11)	(-0.47 – 2.13)	(-1.47 – 1.26)	(-1.15 – 0.07)	(-0.35 – 0.37)
Education (cubic)	-0.65	1.81	1.16 **	-1.01 **	-0.42 *	-0.03
	(-4.64 – 3.35)	(-0.52 – 4.13)	(0.26 – 2.05)	(-1.95 – -0.07)	(-0.84 – 0.01)	(-0.21 – 0.14)
Education (4th degree)	0.15	-0.01	-0.10	-0.08	0.01	
	(-2.16 – 2.46)	(-1.35 – 1.33)	(-0.62 – 0.41)	(-0.62 – 0.46)	(-0.23 – 0.26)	
intervention:time2	1.55 *	0.32	0.08	-0.27	-0.18 *	0.01
	(-0.27 – 3.38)	(-0.61 – 1.25)	(-0.28 – 0.44)	(-0.68 – 0.14)	(-0.37 – 0.00)	(-0.16 – 0.17)
intervention:time3	-0.36	0.76	0.05	-0.22	0.15	0.03
	(-2.38 – 1.65)	(-0.32 – 1.85)	(-0.33 – 0.43)	(-0.65 – 0.22)	(-0.04 – 0.34)	(-0.14 – 0.21)
Random Effects						
σ^2	8.57	2.39	0.36	0.44	0.09	
τ_{00}	99.95 ID	33.41 ID	5.13 ID	5.70 ID	1.11 ID	

τ_{11}	72.66 ID.time2	15.86 ID.time2	2.48 ID.time2	3.60 ID.time2	0.70 ID.time2	
	113.80 ID.time3	31.01 ID.time3	3.86 ID.time3	5.22 ID.time3	0.95 ID.time3	
ρ_{00}	-0.49	-0.36	-0.28	-0.43	-0.29	
	-0.45	-0.45	-0.46	-0.46	-0.45	
ICC	0.92	0.93	0.93	0.93	0.93	
N	1017 ID	1001 ID	1040 ID	1042 ID	1034 ID	526 ID
Observations	1688	1642	1734	1736	1723	714
Marginal R2 Conditional R2	0.027 / 0.926	0.094 / 0.940	0.024 / 0.936	0.058 / 0.932	0.098 / 0.938	

Note: *p<0.1; **p<0.05; ***p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance; τ_{00} = variance of random intercept ; τ_{11} = variance of random slope; ρ_{01} = correlation between random intercept and slope

Table A.13: Regressions mixed methods: group + time + interaction, and covariates

Dependent Var	EQ5D vas OC	Absenteeism GP	Absenteeism last 4 weeks GP	Work-related consultations B	Referrals NB	Spontaneous consultations NB	Prescribed medication B	Non-prescribed medication B
Range	(0-100)	(0-526)	(0-28)	(no-yes)	(0-5)	(0-5)	(no-yes)	(no-yes)
Predictors	*OC*	*IRR*	*IRR*	*OR*	*IRR*	*IRR*	*OR*	*OR*
Intercept	-8.19	2.23 *	5.12 ***	0.53 **	0.64	0.39 ***	2.49 **	0.14 ***
	(-8.23– -8.15)	(0.93 – 5.35)	(1.93 – 13.54)	(0.31 – 0.89)	(0.05 – 8.44)	(0.24 – 0.64)	(1.19 – 5.20)	(0.07 – 0.24)
Intervention group (ref.=control)	0.16	0.59 **	0.30 ***	0.69 **	0.46 ***	0.35 ***	1.53 *	1.23
	(0.15– 0.17)	(0.37 – 0.95)	(0.14 – 0.63)	(0.47 – 1.00)	(0.28 – 0.76)	(0.23 – 0.55)	(0.93 – 2.50)	(0.83 – 1.82)
Time 2 (ref.=Time1)	0.3	1.45	1.36	0.80	0.58 *	0.54 ***	1.30	1.51 *
	(0.29– 0.31)	(0.91 – 2.29)	(0.74 – 2.51)	(0.52 – 1.23)	(0.33 – 1.00)	(0.34 – 0.86)	(0.78 – 2.17)	(0.97 – 2.34)
Time 3	0.26	0.83	1.38	0.62 **	0.48 ***	0.59 ***	0.88	1.32
	(0.25– 0.26)	(0.52 – 1.33)	(0.74 – 2.59)	(0.43 – 0.92)	(0.29 – 0.80)	(0.40 – 0.87)	(0.56 – 1.40)	(0.89 – 1.96)
Hospital B (ref. = hosp A)	-0.65	0.85	1.75 ***	0.64 **	0.88	0.79	0.94	0.85
	(-0.66– -0.64)	(0.59 – 1.23)	(1.16 – 2.64)	(0.45 – 0.92)	(0.55 – 1.42)	(0.57 – 1.11)	(0.57 – 1.56)	(0.58 – 1.23)
Hospital C	-0.51	0.78	0.78	0.79	1.13	0.82	0.59 **	1.01
	(-0.52– -0.5)	(0.54 – 1.13)	(0.49 – 1.23)	(0.55 – 1.14)	(0.74 – 1.74)	(0.57 – 1.16)	(0.35 – 0.99)	(0.70 – 1.47)
Hospital D	-0.61	0.46 ***	0.75	0.80	0.74	0.58 **	0.65	0.96
	(-0.62– -0.6)	(0.27 – 0.80)	(0.30 – 1.86)	(0.51 – 1.25)	(0.40 – 1.34)	(0.36 – 0.95)	(0.34 – 1.23)	(0.61 – 1.51)
Age <25 (ref. = Age > 55)	0.06	0.90	0.53	0.87	1.67	1.95 *	0.14 ***	2.04
	(0.04– 0.09)	(0.28 – 2.89)	(0.14 – 2.00)	(0.33 – 2.28)	(0.61 – 4.62)	(0.91 – 4.15)	(0.04 – 0.55)	(0.77 – 5.42)
Age 25 - 34	-0.02	1.45 *	0.97	0.99	0.58 **	0.97	0.18 ***	1.07
	(-0.03– -0.01)	(0.96 – 2.21)	(0.59 – 1.59)	(0.66 – 1.50)	(0.34 – 0.99)	(0.65 – 1.45)	(0.09 – 0.34)	(0.70 – 1.65)
Age 35 - 44	-0.24	1.06	1.47	0.80	0.75	1.00	0.20 ***	1.52 **
	(-0.25– -0.23)	(0.69 – 1.61)	(0.86 – 2.52)	(0.54 – 1.18)	(0.47 – 1.20)	(0.69 – 1.46)	(0.11 – 0.36)	(1.02 – 2.27)
Age 45 - 54	-0.03	0.99	1.37	0.98	0.91	0.79	0.43 ***	1.09
	(-0.04– -0.03)	(0.67 – 1.47)	(0.88 – 2.14)	(0.69 – 1.40)	(0.61 – 1.37)	(0.55 – 1.12)	(0.26 – 0.72)	(0.75 – 1.57)
Female (ref. = Male)	0.22	1.83 ***	1.03	1.97 ***	1.24	1.32	1.29	1.87 ***
	(0.22– 0.23)	(1.17 – 2.87)	(0.47 – 2.25)	(1.36 – 2.84)	(0.80 – 1.92)	(0.91 – 1.92)	(0.78 – 2.14)	(1.28 – 2.72)
Education (linear)	-0.18	0.78	0.54 ***	1.16	0.54 ***	0.73	0.93	1.78 **
	(-0.22– -0.14)	(0.50 – 1.24)	(0.35 – 0.84)	(0.71 – 1.89)	(0.34 – 0.83)	(0.48 – 1.11)	(0.47 – 1.84)	(1.04 – 3.04)
intervention:time2	-0.05	1.38	4.11 ***	1.20	1.16	2.51 **	0.56	0.97

	(-0.06– -0.04)	(0.68 – 2.79)	(1.69 – 10.01)	(0.66 – 2.18)	(0.46 – 2.88)	(1.20 – 5.27)	(0.28 – 1.16)	(0.53 – 1.79)
intervention:time3	-0.06	1.91 *	2.96 **	1.73 *	1.78	2.47 ***	0.95	1.27
	(-0.07– -0.05)	(0.94 – 3.85)	(1.11 – 7.89)	(1.00 – 2.99)	(0.80 – 3.97)	(1.29 – 4.75)	(0.49 – 1.83)	(0.73 – 2.21)
Zero-Inflated Model								
(Intercept)			7.19 ***		0.24			
Dispersion Model								
overdispersion parameter		521	7.19		3.5	1.11		
Random Effects								
σ^2	3.144			3.29			3.29	3.29
τ_{00}	1.765			1.37 ID			4.30 ID	1.44 ID
ICC				0.29			0.57	0.31
N				1042 ID			1042 ID	1042 ID
Observations	1037	1714	1684	1735	1734	1736	1734	1734
R2 conditional/ R2 marginal	NA	NA / 0.410	NA / 0.344	0.028 / 0.313	NA / 0.077	NA / 0.084	0.073 / 0.598	0.046 / 0.337

Table A.14: Regressions mixed methods: group + time + interaction, and covariates

Dependent Var	Weighted Presenteeism		
Range	(0-140)		
Predictors	*Incidence Rate Ratios*		
Intercept	20.30 ***	**Dispersion Model**	
	(15.47 – 26.63)	Intercept	12.47 ***
Intervention group (ref.=control)	1.15		(7.96 – 19.53)
	(0.95 – 1.39)	Hospital B (ref. = hosp A)	0.93
Time 2 (ref.=Time1)	0.98		(0.56 – 1.55)
	(0.77 – 1.25)	Hospital C	0.77
Time 3	1.07		(0.46 – 1.30)
	(0.87 – 1.31)	Hospital D	1.25
Hospital B (ref. = hosp A)	0.97		(0.66 – 2.36)
	(0.78 – 1.20)	Age <25 (ref. = Age > 55)	3.24
Hospital C	0.83 *		(0.67 – 15.64)
	(0.67 – 1.03)	Age 25 - 34	1.48
Hospital D	1.09		(0.83 – 2.61)
	(0.83 – 1.44)	Age 35 - 44	1.44
Age <25 (ref. = Age > 55)	1.29		(0.84 – 2.44)
	(0.62 – 2.71)	Age 45 - 54	1.80 **
Age 25 - 34	1.04		(1.08 – 3.00)
	(0.82 – 1.33)	**Zero-Inflated Model**	
Age 35 - 44	1.12	Intercept	1.90 ***
	(0.90 – 1.40)		(1.45 – 2.48)
Age 45 - 54	1.13	Hospital B (ref. = hosp A)	1.73 ***

	(0.91 – 1.40)		(1.31 – 2.29)
Female (ref. = Male)	0.83 **	Hospital C	1.33 **
	(0.70 – 0.99)		(1.01 – 1.75)
Education (linear)	0.88	Hospital D	1.44 **
	(0.71 – 1.08)		(1.02 – 2.02)
Education (quadratic)	0.96	Age <25 (ref. = Age > 55)	0.56
	(0.81 – 1.14)		(0.27 – 1.17)
intervention:time2	0.97	Age 25 - 34	0.48 ***
	(0.69 – 1.35)		(0.35 – 0.66)
intervention:time3	0.88	Age 35 - 44	0.69 **
	(0.66 – 1.18)		(0.51 – 0.93)
		Age 45 - 54	0.71 **
			(0.54 – 0.94)
Observations	1638		
R2 conditional / R2 marginal	NA / 0.026		

Table A.15: Regressions mixed methods: group + time + interaction, and covariates.

Dependent Var Range	**Worry intensity** (1-4)	**Worry frequency** (1-5)	**Worry weighted score** (0-1.55)
Intercept	1.41 ***	1.82 ***	0.72 ***
	(1.19 – 1.64)	(1.51 – 2.14)	(0.61 – 0.83)
group2 [intervention]	-0.06	-0.03	-0.02
	(-0.17 – 0.06)	(-0.18 – 0.13)	(-0.07 – 0.03)
time [2]	-0.14 **	0.01	-0.03
	(-0.24 – -0.03)	(-0.12 – 0.15)	(-0.08 – 0.01)
time [3]	-0.06	0.24 ***	0.03
	(-0.16 – 0.04)	(0.11 – 0.38)	(-0.01 – 0.08)
hosp [Heilig Hart Tienen]	-0.11 *	-0.28 ***	-0.08 ***
	(-0.23 – 0.01)	(-0.44 – -0.11)	(-0.14 – -0.03)
hosp [Heilig Hart Leuven]	-0.15 **	-0.20 **	-0.08 ***
	(-0.27 – -0.03)	(-0.37 – -0.03)	(-0.14 – -0.02)
hosp [AZ Vesalius]	-0.13 *	-0.38 ***	-0.11 ***
	(-0.28 – 0.01)	(-0.58 – -0.17)	(-0.18 – -0.04)
agecat [Age <25]	-0.09	0.11	-0.00
	(-0.38 – 0.21)	(-0.30 – 0.52)	(-0.14 – 0.14)
agecat [Age 25 - 34]	-0.23 ***	-0.12	-0.08 **
	(-0.37 – -0.10)	(-0.31 – 0.07)	(-0.15 – -0.02)
agecat [Age 35 - 44]	-0.26 ***	-0.17 *	-0.10 ***
	(-0.39 – -0.14)	(-0.35 – 0.01)	(-0.16 – -0.04)
agecat [Age 45 - 54]	-0.09	-0.15 *	-0.05 *
	(-0.20 – 0.02)	(-0.31 – 0.01)	(-0.11 – 0.00)
sex [Vrouw]	0.04	0.15 *	0.04
	(-0.08 – 0.15)	(-0.02 – 0.31)	(-0.02 – 0.09)
educs [linear]	-0.14	-0.40	-0.12
	(-0.65 – 0.37)	(-1.12 – 0.32)	(-0.36 – 0.13)
educs [quadratic]	0.06	0.30	0.07
	(-0.37 – 0.50)	(-0.32 – 0.91)	(-0.14 – 0.28)
educs [cubic]	0.14	-0.16	0.00

	(-0.16 – 0.44)	(-0.58 – 0.27)	(-0.14 – 0.15)
	0.03	0.08	0.02
educs [4th degree]	(-0.14 – 0.20)	(-0.16 – 0.33)	(-0.06 – 0.11)
	0.20 ***	0.10	0.07 **
intervention:time2	(0.05 – 0.35)	(-0.08 – 0.29)	(0.01 – 0.13)
	0.03	-0.07	-0.00
intervention:time3	(-0.11 – 0.18)	(-0.27 – 0.12)	(-0.07 – 0.06)
Random Effects			
σ^2	0.05	0.09	0.01
τ_{00}	0.62 ID	1.15 ID	0.14 ID
τ_{11}	0.46 ID.time2	0.72 ID.time2	0.08 ID.time2
	0.61 ID.time3	1.07 ID.time3	0.12 ID.time3
ρ_{01}	-0.42	-0.39	-0.36
	-0.57	-0.48	-0.5
ICC	0.92	0.93	0.93
N	1042 ID	1042 ID	1042 ID
Observations	1736	1735	1735
Marginal R2 / Conditional R2	0.037 / 0.922	0.030 / 0.930	0.031 / 0.931

Note: *p<0.1; **p<0.05; *** p<0.01. ICC=between/(between+within) variance. σ^2 = within-group variance;
τ_{00} = variance of random intercept ; τ_{11} = variance of random slope;
ρ_{01} = correlation between random intercept and slope

Bibliography

Acheson, D., Barker, D., Chambers, J., Drever, F., Earwicker, R., Graham, H., ... Whitehead, M. (1998). *Independent inquiry into inequalities in health report*. The Stationery Office, London.

Action, E. J., Heintz, E., Gerber-Grote, A., Ghabri, S., Hamers, F. F., Rupel, V. P., ... Davidson, T., et al. (2016). Is there a european view on health economic evaluations? results from a synopsis of methodological guidelines used in the eunethta partner countries. *Pharmacoeconomics*, *34*(1), 59–76.

Adepoju, O. E., Bolin, J. N., Ohsfeldt, R. L., Phillips, C. D., Zhao, H., Ory, M. G., & Forjuoh, S. N. (2014). Can chronic disease management programs for patients with type 2 diabetes reduce productivity-related indirect costs of the disease? evidence from a randomized controlled trial. *Popul Health Manag*, *17*(2), 112–120.

Alamgir, H., Yu, S., Fast, C., Hennessy, S., Kidd, C., & Yassi, A. (2008). Efficiency of overhead ceiling lifts in reducing musculoskeletal injury among carers working in long-term care institutions. *Injury*, *39*(5), 570–7. doi:10.1016/j.injury.2007.11.420

All Party Parliamentary Group on Occupational Safety and Health. (2016). Occupational medical workforce crisis. the need for action to keep the uk workforce healthy. All Party Parliamentary Group on Occupational Safety and Health London.

Altin, S. V., Finke, I., Kautz-Freimuth, S., & Stock, S. (2014). The evolution of health literacy assessment tools: A systematic review. *BMC public health*, *14*(1), 1207.

Ambler, T., & Roberts, J. H. (2008). Assessing marketing performance: Don't settle for a silver metric. *J Market Manag*, *24*(7-8), 733–750. doi:10.1362/026725708X345498

Anders, B., Ommen, O., Pfaff, H., Lüngen, M., Lefering, R., Thüm, S., & Janssen, C. (2013). Direct, indirect, and intangible costs after severe trauma up to occupational reintegration–an empirical analysis of 113 seriously injured patients. *GMS Psycho-Social-Medicine*, *10*.

Apostolakis, G. E. (2004). How useful is quantitative risk assessment? *Risk Anal*, *24*(3), 515–520. doi:10.1111/j.0272-4332.2004.00455.x

Arena, R., Arnett, D. K., Terry, P. E., Li, S., Isaac, F., Mosca, L., ... Whitsel, L. P. (2014). The role of worksite health screening: A policy statement from the american heart association. *Circulation*, *130*(8), 719–34. doi:10.1161/CIR.0000000000000079

Arends, I., Bultmann, U., van Rhenen, W., Groen, H., & van der Klink, J. J. (2013). Economic evaluation of a problem solving intervention to prevent recurrent sickness absence in workers with common mental disorders. *PLoS One*, *8*(8), e71937. doi:10.1371/journal.pone.0071937

Arneson, H., & Ekberg, K. (2006). Measuring empowerment in working life: A review. *Work*, *26*(1), 37–46.

Asche, F., & Aven, T. (2004). On the economic value of safety. *Risk Decis Pol*, *9*(3), 253–267.

At'kov, O. Y., Azarov, A. V., Zhukov, D. A., Nicoloyannis, N., & Durand, L. (2011). Influenza vaccination in healthy working adults in russia: Observational study of effectiveness and return on investment for the employer. *Appl Health Econ Health Policy*, *9*(2), 89–99. doi:10.2165/11538680-000000000-00000

Aven, T. (2010). On how to define, understand and describe risk. *Reliab Eng Syst Saf*, *95*(6), 623–631. doi:10.1016/j.ress.2010.01.011

Aven, T., & Zio, E. (2014). Foundational issues in risk assessment and risk management. *Risk Anal*, *34*(7), 1164–1172. doi:10.1111/risa.12132

Baker, J., Mitchell, R., Lawson, K., & Pell, J. (2013). Ethnic differences in the cost-effectiveness of targeted and mass screening for high cardiovascular risk in the uk: Cross-sectional study. *Heart*, *99*(23), 1766–71. doi:10.1136/heartjnl-2013-304625

Baker, K. M., Goetzel, R. Z., Pei, X. F., Weiss, A. J., Bowen, J., Tabrizi, M. J., ... Thompson, E. (2008). Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *J Occup Environ Med*, *50*(9), 981–990. doi:10.1097/JOM.0b013e318184a489

Barbosa, C., Bray, J. W., Dowd, W. N., Mills, M. J., Moen, P., Wipfli, B., ... Kelly, E. L. (2015). Return on investment of a work-family intervention: Evidence from the work, family, and health network. *J Occup Environ Med*, *57*(9), 943–51. doi:10.1097/JOM.0000000000000520

Barling, J., & Hutchinson, I. (2000). Commitment vs. control-based safety practices, safety reputation, and perceived safety climate. *Can J Adm Sci*, *17*(1), 76–84. doi:10.1111/j.1936-4490.2000.tb00208.x

Behling, C., Kelly, R., Ruppert, J., Whitsel, L., Hudson, M., Terry, P., ... Loeppke, R. (2013). Biometric health screening for employers consensus statement of the health enhancement research organization, american

college of occupational and environmental medicine, and care continuum alliance. *J. Occup. Environ. Med.*, *55*(10), 1244–1251.

Bernaards, C. M., Bosmans, J. E., Hildebrandt, V. H., van Tulder, M. W., & Heymans, M. W. (2011). The cost-effectiveness of a lifestyle physical activity intervention in addition to a work style intervention on recovery from neck and upper limb symptoms and pain reduction in computer workers. *Occup Environ Med*, *68*(4), 265–72. doi:10.1136/oem.2008.045450

Black, C., Bagust, A., Boland, A., Walker, S., McLeod, C., De Verteuil, R., ... Waugh, N. (2006). The clinical effectiveness and cost-effectiveness of computed tomography screening for lung cancer: Systematic reviews. *Health Technol Assess*, *10*(3), iii–iv, ix–x, 1–90.

Boezeman, E. J., Nieuwenhuijsen, K., & Sluiter, J. K. (2016). Predictive value and construct validity of the work functioning screener–healthcare (wfs-h). *J Occup Health*, 15-0056–OA. doi:10.1539/joh.15-0056-OA

Boschman, J., Hulshof, C., Frings-Dresen, M., & Sluiter, J. (2016). Improving fit to work assessments for rail safety workers by exploring work limitations. *Int Arch Occup Environ Health*, *89*(5), 803–811. doi:10.1007/s00420-016-1117-7

Bouwmans, C., Krol, M., Severens, H., Koopmanschap, M., Brouwer, W., & Hakkaart-van Roijen, L. (2015). The imta productivity cost questionnaire: A standardized instrument for measuring and valuing health-related productivity losses. *Value in health*, *18*(6), 753–758.

Braun, T., Bambra, C., Booth, M., Adetayo, K., & Milne, E. (2015). Better health at work? an evaluation of the effects and cost-benefits of a structured workplace health improvement programme in reducing sickness absence. *J Public Health-Uk*, *37*(1), 138–142.

Briggs, A., Sculpher, M., & Claxton, K. (2006). *Decision modelling for health economic evaluation*. Oxford: Oxford Univ Pr.

Brouwer, W. B., & Koopmanschap, M. A. (2005). The friction-cost method : Replacement for nothing and leisure for free? *PharmacoEconomics*, *23*(2), 105–11. doi:10.2165/00019053-200523020-00002

Brown, H. E., Gilson, N. D., Burton, N. W., & Brown, W. J. (2011). Does physical activity impact on presenteeism and other indicators of workplace well-being? *Sports Med*, *41*(3), 249–62. doi:10.2165/11539180-000000000-00000

Brown, J., Mackay, D., Demou, E., Craig, J., Frank, J., & Macdonald, E. B. (2015). The easy (early access to support for you) sickness absence service: A four-year evaluation of the impact on absenteeism. *Scand J Work Environ Health*, *41*(2), 204–15. doi:10.5271/sjweh.3480

Brownlee, S., Chalkidou, K., Doust, J., Elshaug, A. G., Glasziou, P., Heath, I., ... Chalmers, K. (2017). Evidence for overuse of medical services around the world. *The Lancet*.

Brymer, R. A., Molloy, J. C., & Gilbert, B. A. (2014). Human capital pipelines: Competitive implications of repeated interorganizational hiring. *J Manage*, *40*(2), 483–508. doi:10.1177/0149206313516797

Buchan, J., & Campbell, J. (2013). Challenges posed by the global crisis in the health workforce. *BMJ*, *347:f6201.* doi:10.1136/bmj.f6201

Bulterys, S., Johannik, K., Vlamings, J., & Moens, G. (2006). An employee satisfaction survey in occupational health care—how do employees evaluate their occupational health service. In *International congress series* (Vol. 1294, pp. 89–92). Elsevier.

Bültmann, U., Sherson, D., Olsen, J., Hansen, C. L., Lund, T., & Kilsgaard, J. (2009). Coordinated and tailored work rehabilitation: A randomized controlled trial with economic evaluation undertaken with workers on sick leave due to musculoskeletal disorders. *J Occup Rehabil*, *19*(1), 81–93. doi:10.1007/s10926-009-9162-7

Burden, A. C., & Rouse, A. (2010). Cardiovascular prevention. go for targeted screening. *BMJ*, *340*, c2866. doi:10.1136/bmj.c2866

Burdorf, A. (2007). Economic evaluation in occupational health—its goals, challenges, and opportunities. *Scand J Work Environ Health*, 161–164. doi:10.5271/sjweh.1127

Burton, J., & World Health Organization. (2010). *Who healthy workplace framework and model: Background and supporting literature and practices* (Report No. Report No.: 9241500247). World Health Organization.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., & Bradley, K. A. (1998). The audit alcohol consumption questions (audit-c): An effective brief screening test for problem drinking. *Archives of internal medicine*, *158*(16), 1789–1795.

Cagno, E., Micheli, G. J. L., Masi, D., & Jacinto, C. (2013). Economic evaluation of osh and its way to smes: A constructive review. *Safety Sci*, *53*, 134–152. doi:10.1016/j.ssci.2012.08.016

Carroll, C., Rick, J., Pilgrim, H., Cameron, J., & Hillage, J. (2010). Workplace involvement improves return to work rates among employees with back pain on long-term sick leave: A systematic review of the effectiveness and cost-effectiveness of interventions. *Disabil Rehabil*, *32*(8), 607–21. doi:10.3109/09638280903186301

Centers for Disease Control and Prevention (CDC). (2018). Workplace health in america 2017. Atlanta, GA. Centers for Disease Control, Prevention, U.S. Department of Health, and Human Services.

Chan, A. L., Shie, H. J., Lee, Y. J., & Lin, S. J. (2008). The evaluation of free influenza vaccination in health care workers in a medical center in taiwan. *Pharm World Sci*, *30*(1), 39–43. doi:10.1007/s11096-007-9137-8

Chandra, A., & Skinner, J. (2012). Technology growth and expenditure growth in health care. *J Econ Lit*, *50*(3), 645–80. doi:10.1257/jel.50.3.645

Chapman, L. S. (2012). Meta-evaluation of worksite health promotion economic return studies: 2012 update. *Am J Health Promot*, *26*(4), TAHP1–TAHP12. doi:10.4278/ajhp.26.4.tahp

Chartered Institute of Personnel and Development. (2007). *Recruitment, retention and turnover: Annual survey report 2007*. CIPD.

Chen, Y.-C., Guo, Y.-L. L., Lin, L.-C., Lee, Y.-J., Hu, P.-Y., Ho, J.-J., & Shiao, J. S.-C. (2020). Development of the nurses' occupational stressor scale. *Int J Environ Res Public Health*, *17*(2), 649. doi:10.3390/ijerph17020649

Cherry, N., Meyer, J., Holt, D., Chen, Y., & McDonald, J. (2000). Surveillance of work-related diseases by occupational physicians in the uk: Opra 1996–1999. *Occupational Medicine*, *50*(7), 496–503.

Chowdhury, M. F., & Tanim, T. R. (2016). Industrial accidents in bangladesh apparel manufacturing sector: An analysis of the two most deadliest accidents in history. *Asian J Soc Sci Manag Stud*, *3*(2), 115–126. doi:10.20448/journal.500/2016.3.2/500.2.115.126

Clarke, V. (2013). Honesty and integrity: Is this the edge businesses need in a stagnant economy? Web Page.

Cleemput, I., Van Wilder, P., Vrijens, F., Huybrechts, M., & Ramaekers, D. (2008). *Guidelines for pharmacoeconomic evaluations in belgium. health technology assessment (hta)* (Report No. KCE Reports 78C (D/2008/10.273/27)). Federal Health Care Knowledge Centre (KCE).

Colosio, C., Mandic-Rajcevic, S., Godderis, L., van der Laan, G., Hulshof, C., & van Dijk, F. (2017). Workers' health surveillance: Implementation of the directive 89/391/eec in europe. *Occupational Medicine*, *67*(7), 574–578.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., & Medical Research Council, G. (2008). Developing and evaluating complex interventions: The new medical research council guidance. *BMJ*, *337*, a1655. doi:10.1136/bmj.a1655

Cravens, K., Oliver, E. G., & Ramamoorti, S. (2003). The reputation index: Measuring and managing corporate reputation. *Eur Manag J*, *21*(2), 201–212.

Dale, A. M., Gardner, B. T., Zeringue, A., Werner, R., Franzblau, A., & Evanoff, B. (2014). The effectiveness of post-offer pre-placement nerve conduction screening for carpal tunnel syndrome. *J Occup Environ Med*, *56*(8), 840–7. doi:10.1097/JOM.0000000000000185

Dallat, M. A., Hunter, R. F., Tully, M. A., Cairns, K. J., & Kee, F. (2013). A lesson in business: Cost-effectiveness analysis of a novel financial incentive intervention for increasing physical activity in the workplace. *BMC Public Health*, *13*, 953. doi:10.1186/1471-2458-13-953

De Croon, E. M., Sluiter, J. K., & Frings-Dresen, M. H. (2006). Psychometric properties of the need for recovery after work scale: Test-retest reliability

and sensitivity to detect change. *Occupational and environmental medicine*, *63*(3), 202–206.

De Loecker, J., & Goldberg, P. K. (2014). Firm performance in a global market. *Annu Rev Econ*, *6*(1), 201–227.

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

Dekker, S. W. (2014). The bureaucratization of safety. *Safety Sci*, *70*, 348–357. doi:10.1016/j.ssci.2014.07.015

Desapriya, E., Harjee, R., Brubacher, J., Chan, H., Hewapathirane, D. S., Subzwari, S., & Pike, I. (2014). Vision screening of older drivers for preventing road traffic injuries and fatalities. *Cochrane Database Syst Rev*, (2), CD006252. doi:10.1002/14651858.CD006252.pub4

Désiron, H. A., Donceel, P., Godderis, L., Van Hoof, E., & de Rijk, A. (2015). What is the value of occupational therapy in return to work for breast cancer patients? a qualitative inquiry among experts. *European journal of cancer care*, *24*(2), 267–280.

Dewa, C. S., & Hoch, J. S. (2014a). Estimating the net benefit of a specialized return-to-work program for workers on short-term disability related to a mental disorder: An example exploring investment in collaborative care. *J Occup Environ Med*, *56*(6), 628–31. doi:10.1097/JOM.0000000000000157

Dewa, C. S., & Hoch, J. S. (2014b). When could a stigma program to address mental illness in the workplace break even? *Can J Psychiatry*, *59*(10), S34–S39.

Dewa, C. S., Hoch, J. S., Carmen, G., Guscott, R., & Anderson, C. (2009). Cost, effectiveness, and cost-effectiveness of a collaborative mental health care program for people receiving short-term disability benefits for psychiatric disorders. *Can J Psychiatry*, *54*(6), 379–88. doi:10.1177/070674370905400605

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533–534. doi:10.1016/S1473-3099(20)30120-1

Dowd, B., McGrail, M., Lohman, W. H., Sheasby, B., O'Connor, H., Calasanz, M., ... Parente, S. (2010). The economic impact of a disability prevention program. *J Occup Environ Med*, *52*(1), 15–21. doi:10.1097/JOM.0b013e3181c5c358

Driessen, M., Bosmans, J., Proper, K., Anema, J., Bongers, P., & van der Beek, A. (2012). The economic evaluation of a participatory ergonomics programme to prevent low back and neck pain. *Work*, *41 Suppl 1*, 2315–20. doi:10.3233/WOR-2012-0458-2315

Drummond, M. F., & McGuire, A. (2001). *Economic evaluation in health care: Merging theory with practice*. Oxford: Oxford University Press.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Univ Pr.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). Methods for the economic evaluation of health care programmes. Book. Oxford: Oxford Univ Pr.

Dutheil, F., Delaire, P., Boudet, G., Rouffiac, K., Djeriri, K., Souweine, B., & Chamoux, A. (2008). [cost/effectiveness comparison of the vaccine campaign and reduction of sick leave, after vaccination against influenza among the clermont-ferrand university hospital staff]. *Med Maladies Infect*, *38*(11), 567–573. doi:10.1016/j.medmal.2008.09.019

Edwards, R. T., Charles, J. M., & Lloyd-Williams, H. (2013). Public health economics: A systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Publ Health*, *13*(1), 1001. doi:10.1186/1471-2458-13-1001

Ektor-Andersen, J., Ingvarsson, E., Kullendorff, M., & Orbaek, P. (2008). High cost-benefit of early team-based biomedical and cognitive-behaviour intervention for long-term pain-related sickness absence. *J Rehabil Med*, *40*(1), 1–8. doi:10.2340/16501977-0127

Elsler, D., Treutlein, D., Rydlewska, I., Frusteri, L., Kruger, H., Veerman, T., ... Taylor, T. N. (2010). A review of case studies evaluating economic incentives to promote occupational safety and health. *Scand J Work Environ Health*, *36*(4), 289–98.

Esler, D., Eeckelaert, L., Knight, A., Treutlein, D., Pecillo, M., Elo-Schäfer, J., ... Dontas, S. (2010). *Economic incentives to improve occupational safety and health: A review from the european perspective*.

EU. (1989a). *Council directive 89/391/eec of 12 june 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work*. Official Journal of the European Communities.

EU. (1989b). Council directive 89/391/eec of 12 june 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work. *Official Journal of the European Communities*, *183*, 1–8.

EU. (2017). Commission staff working document, ex-post evaluation of the european union occupational safety and health directives (refit evaluation). *Official Journal of the European Communities*, *SWD 10*, 21.

European Agency for Safety and Health at Work. (2017). The economics of occupational safety and health - the value of osh to society. Web Page.

EuroQol Research Foundation. (2019). *Eq-5d-5l user guide*.

Evangelinos, K., Fotiadis, S., Skouloudis, A., Khan, N., Konstandakopoulou, F., Nikolaou, I., & Lundy, S. (2018). Occupational health and safety disclosures in sustainability reports: An overview of trends among corporate leaders. *Corp Soc Resp Env Ma*. doi:10.1002/csr.1512

Evanoff, B., & Kymes, S. (2010). Modeling the cost-benefit of nerve conduction studies in pre-employment screening for carpal tunnel syndrome. *Scand J Work Environ Health*, *36*(4), 299–304.

Evans-Lacko, S., Koeser, L., Knapp, M., Longhitano, C., Zohar, J., & Kuhn, K. (2016). Evaluating the economic impact of screening and treatment for depression in the workplace. *Eur Neuropsychopharmacol*, *26*(6), 1004–13. doi:10.1016/j.euroneuro.2016.03.005

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. *Int J Technol Assess Health Care*, *21*(2), 240–5.

Fisk, W. J., Black, D., & Brunner, G. (2011). Benefits and costs of improved ieq in u.s. offices. *Indoor Air*, *21*(5), 357–67. doi:10.1111/j.1600-0668.2011.00719.x

Fletcher, C. M., Elmes, P. C., Fairbairn, A. S., & Wood, C. H. (1959). Significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *British medical journal*, *2*(5147), 257.

Fleurbaey, M., & Abi-Rafeh, R. (2016). The use of distributional weights in benefit–cost analysis: Insights from welfare economics. *Review of Environmental Economics and Policy*, *10*(2), 286–307.

Fleurbaey, M., Luchini, S., Muller, C., & Schokkaert, E. (2013). Equivalent income and fair evaluation of health care. *Health Economics*, *22*(6), 711–729.

Furlan, A. D., Gnam, W. H., Carnide, N., Irvin, E., Amick, 3., B. C., DeRango, K., ... Bultmann, U. (2012). Systematic review of intervention practices for depression in the workplace. *J Occup Rehabil*, *22*(3), 312–21. doi:10.1007/s10926-011-9340-2

Gagné, M., Forest, J., Gilbert, M.-H., Aubé, C., Morin, E., & Malorni, A. (2010). The motivation at work scale: Validation evidence in two languages. *Educ Psychol Meas*, *70*(4), 628–646. doi:10.1177/0013164409355698

Gärtner, F. R., Ketelaar, S. M., Smeets, O., Bolier, L., Fischer, E., van Dijk, F. J., ... Sluiter, J. K. (2011). The mental vitality@ work study: Design of a randomized controlled trial on the effect of a workers' health surveillance mental module for nurses and allied health professionals. *BMC Public Health*, *11*(1), 1–13. doi:10.1186/1471-2458-11-290

Gärtner, F. R., Nieuwenhuijsen, K., Van Dijk, F. J., & Sluiter, J. K. (2011). Psychometric properties of the nurses work functioning questionnaire (nwfq). *PLoS One*, *6*(11), e26565. doi:10.1371/journal.pone.0026565

Gärtner, F. R., Nieuwenhuijsen, K., van Dijk, F. J., & Sluiter, J. K. (2012). Impaired work functioning due to common mental disorders in nurses and allied health professionals: The nurses work functioning questionnaire.

International archives of occupational and environmental health, *85*(2), 125–138.

Geraedts, A. S., van Dongen, J. M., Kleiboer, A. M., Wiezer, N. M., van Mechelen, W., Cuijpers, P., & Bosmans, J. E. (2015). Economic evaluation of a web-based guided self-help intervention for employees with depressive symptoms: Results of a randomized controlled trial. *J Occup Environ Med*, *57*(6), 666–75. doi:10.1097/JOM.0000000000000423

Gerbert, B., Bronstone, A., Pantilat, S., McPhee, S., Allerton, M., & Moe, J. (1999). When asked, patients tell: Disclosure of sensitive health-risk behaviors. *Medical Care*, 104–111.

Glasziou, P., Straus, S., Brownlee, S., Trevena, L., Dans, L., Guyatt, G., ... Saini, V. (2017). Evidence for underuse of effective medical services around the world. *The Lancet*.

Godderis, L., Mylle, G., Coene, M., Verbeek, C., Viaene, B., Bulterys, S., & Schouteden, M. (2015). Data warehouse for detection of occupational diseases in ohs data. *Occup Med (Lond)*, *65*(8), 651–658.

Goetzel, R. Z., Tabrizi, M., Henke, R. M., Benevent, R., Brockbank, C. V., Stinson, K., ... Newman, L. S. (2014). Estimating the return on investment from a health risk management program offered to small colorado-based employers. *J Occup Environ Med*, *56*(5), 554–60. doi:10.1097/JOM.0000000000000152

Gold, M., Siegel, J., Russell, L., Weinstein, M., & editors. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford Universitary Press.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness.* Swindon, Wiltshire, UK: Oxford U. Press.

Goldberg, D. (1978). *Manual of the general health questionnaire*. Windsor, England: NFER Publishing.

Goldberg, D., & Williams, P. (2000). General health questionnaire (ghq). *Swindon, Wiltshire, UK: nferNelson*.

Goorden, M., Vlasveld, M. C., Anema, J. R., van Mechelen, W., Beekman, A. T., Hoedeman, R., ... Hakkaart-van Roijen, L. (2014). Cost-utility analysis of a collaborative care intervention for major depressive disorder in an occupational healthcare setting. *J Occup Rehabil*, *24*(3), 555–62. doi:10.1007/s10926-013-9483-4

Goossens, E., Fleck, D., Canobbio, M. M., Harrison, J. L., & Moons, P. (2013). Development of an international research agenda for adult congenital heart disease nursing. *European Journal of Cardiovascular Nursing*, *12*(1), 7–16.

Greene, B. L., Miller, J. D., Brown, T. M., Harshman, R. S., Richerson, G. T., & Doyle, J. J. (2009). Economic impact of the bp downshift program

on blood pressure control among commercial driver license employees. *J Occup Environ Med*, *51*(5), 542–553. doi:10.1097/JOM.0b013e3181a2fec7

Greene, C. J., Morland, L. A., Durkalski, V. L., & Frueh, B. C. (2008). Noninferiority and equivalence designs: Issues and implications for mental health research. *Journal of traumatic stress*, *21*(5), 433–439.

Groeneveld, I. F., van Wier, M. F., Proper, K. I., Bosmans, J. E., van Mechelen, W., & van der Beek, A. J. (2011). Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med*, *53*(6), 610–617. doi:10.1097/JOM.0b013e31821b9c24

Guimaraes, L. B., Ribeiro, J. L., & Renner, J. S. (2012). Cost-benefit analysis of a socio-technical intervention in a brazilian footwear company. *Appl Ergon*, *43*(5), 948–57. doi:10.1016/j.apergo.2012.01.003

Hackl, F., Halla, M., Hummer, M., & Pruckner, G. J. (2015). The effectiveness of health screening. *Health Econ*, *24*(8), 913–35. doi:10.1002/hec.3072

Hak, E., Knol, L. M., Wilschut, J. C., & Postma, M. J. (2010). [influenza vaccination of hospital healthcare staff from the perspective of the employer: A positive balance]. *Ned Tijdschr Geneeskd*, *154*, A1188.

Hamberg-van Reenen, H. H., Proper, K. I., & van den Berg, M. (2012). Worksite mental health interventions: A systematic review of economic evaluations. *Occup Environ Med*, *69*(11), 837–45. doi:10.1136/oemed-2012-100668

Hanly, P., Ortega, M. O., Pearce, A., Soerjomataram, I., & Sharp, L. (2020). Advances in the methodological approach to friction period estimation: A european perspective. *Social Science & Medicine*, *264*, 113289.

Hansez, I. (2008). The "working conditions and control questionnaire"(woccq): Towards a structural model of subjective stress. *European review of Applied psychology*, *58*(4), 253–262.

Hartig, F., & Hartig, M. F. (2017). Package 'dharma'.

Hassink, W., & van den Berg, B. (2017). Methods and instruments for the estimation of production changes in economic evaluations. In D. Razzouk (Ed.), *Mental health economics* (Chap. Methods and Instruments for the Estimation of Production Changes in Economic Evaluations, pp. 431–440). Springer, Cham.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the delphi survey technique. *J Adv Nurs*, *32*(4), 1008–1015. doi:10.1046/j.1365-2648.2000.t01-1-01567.x

Haun, J. N., Valerio, M. A., McCormack, L. A., Sørensen, K., & Paasche-Orlow, M. K. (2014). Health literacy measurement: An inventory and descriptive summary of 51 instruments. *Journal of health communication*, *19*(sup2), 302–333.

Herman, A. M., & Jeffress, C. N. (2000). *Screening and surveillance: A guide to osha standards*. Department of labor Washington DC occupational safety and health administration.

Herman, P. M., Szczurko, O., Cooley, K., & Seely, D. (2014). A naturopathic approach to the prevention of cardiovascular disease: Cost-effectiveness analysis of a pragmatic multi-worksite randomized clinical trial. *J Occup Environ Med*, *56*(2), 171–6. doi:10.1097/JOM.0000000000000066

Heuvel, S., Zwaan, L., Dam, L. v., Oude Hengel, K., Eekhout, I., van Emmerik, M., ... Wilhelm, C. (2017). *Estimating the costs of work-related accidents and ill-health: An analysis of european data sources* (Report No. 9292409972). European Agency for Safety and Health at Work (EU-OSHA).

Hildebrandt, V., Bongers, P., Van Dijk, F., Kemper, H., & Dul, J. (2001). Dutch musculoskeletal questionnaire: Description and basic qualities. *Ergonomics*, *44*(12), 1038–1055. doi:10.1080/00140130110087437

Hlobil, H., Uegaki, K., Staal, J. B., de Bruyne, M. C., Smid, T., & van Mechelen, W. (2007). Substantial sick-leave costs savings due to a graded activity intervention for workers with non-specific sub-acute low back pain. *Eur Spine J*, *16*(7), 919–924. doi:10.1007/s00586-006-0283-9

Hoch, J. S., & Dewa, C. S. (2014). Advantages of the net benefit regression framework for economic evaluations of interventions in the workplace: A case study of the cost-effectiveness of a collaborative mental health care program for people receiving short-term disability benefits for psychiatric disorders. *J Occup Environ Med*, *56*(4), 441–5. doi:10.1097/JOM.0000000000000130

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.

Hsu, C.-C., & Sandford, B. A. (2007). The delphi technique: Making sense of consensus. *Practical assessment, research & evaluation*, *12*(10), 1–8.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., ... Loder, E. (2013). Consolidated health economic evaluation reporting standards (cheers) statement. *The European Journal of Health Economics*, *14*(3), 367–372.

IJzelenberg, H., Meerding, W. J., & Burdorf, A. (2007). Effectiveness of a back pain prevention program: A cluster randomized controlled trial in an occupational setting. *Spine (Phila Pa 1976)*, *32*(7), 711–9. doi:10.1097/01.brs.0000259072.14859.d9

International Labour Office. (1998). *Technical and ethical guidelines for workers' health surveillance (osh no. 72)*. International Labour Office (ILO).

Ison, T. G. (1998). Worker's compensation systems. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th edition). Geneva: International Labour Organization.

Jackson, C. (2007). The general health questionnaire. *Occup Med (Lond)*, *57*(1), 79–79. doi:10.1093/occmed/kql169

Jeanrenaud, C., & Pellegrini, S. (2007). Valuing intangible costs of alcohol dependence: A contingent valuation study. *Rev Econ Polit*, *117*(5), 813–825. doi:10.3917/redp.175.0813

Jensen, C., Nielsen, C. V., Jensen, O. K., & Petersen, K. D. (2013). Cost-effectiveness and cost-benefit analyses of a multidisciplinary intervention compared with a brief intervention to facilitate return to work in sick-listed patients with low back pain. *Spine (Phila Pa 1976)*, *38*(13), 1059–67. doi:10.1097/BRS.0b013e31828ca0af

Jensen, I. B., Busch, H., Bodin, L., Hagberg, J., Nygren, A., & Bergstrom, G. (2009). Cost effectiveness of two rehabilitation programmes for neck and back pain patients: A seven year follow-up. *Pain*, *142*(3), 202–8. doi:10.1016/j.pain.2008.12.015

Jensen, J. D. (2011). Can worksite nutritional interventions improve productivity and firm profitability? a literature review. *Perspect Public Health*, *131*(4), 184–92. doi:10.1177/1757913911408263

Jinnett, K., Schwatka, N., Tenney, L., Brockbank, C. V., & Newman, L. S. (2017). Chronic conditions, workplace safety, and job demands contribute to absenteeism and job performance. *Health Aff (Millwood)*, *36*(2), 237–244. doi:10.1377/hlthaff.2016.1151

Johannesson, M., & Karlsson, G. (1997). The friction cost method: A comment. *J Health Econ*, *16*(2), 249–55, discussion 257–9. doi:http://dx.doi.org/10.1016/S0167-6296(97)00006-4

Jones, C. (2018). *The economics of presenteeism in the context of rheumatoid arthritis, ankylosing spondylitis and psoriatic arthritis* (Thesis).

Kankaanpaa, E., Suhonen, A., & Valtonen, H. (2008). Promoting prevention with economic arguments - the case of finnish occupational health services. *BMC Public Health*, *8*, 130. doi:10.1186/1471-2458-8-130

Karrholm, J., Ekholm, K., Ekholm, J., Bergroth, A., & Ekholm, K. S. (2008). Systematic co-operation between employer, occupational health service and social insurance office: A 6-year follow-up of vocational rehabilitation for people on sick-leave, including economic benefits. *J Rehabil Med*, *40*(8), 628–36. doi:10.2340/16501977-0233

Kersten, P., Küçükdeveci, A. A., & Tennant, A. (2012). The use of the visual analogue scale (vas) in rehabilitation outcomes. *Journal of rehabilitation medicine*, *44*(7), 609–610.

Kigozi, J., Jowett, S., Lewis, M., Barton, P., & Coast, J. (2016). Estimating productivity costs using the friction cost approach in practice: A systematic review. *Eur J Health Econ*, *17*(1), 31–44. doi:10.1007/s10198-014-0652-y

Kigozi, J., Jowett, S., Lewis, M., Barton, P., & Coast, J. (2017a). The estimation and inclusion of presenteeism costs in applied economic evaluation: A

systematic review. *Value Health*, *20*(3), 496–506. doi:10.1016/j.jval.2016.12.006

Kigozi, J., Jowett, S., Lewis, M., Barton, P., & Coast, J. (2017b). Valuing productivity costs using the friction-cost approach: Estimating friction-period estimates by occupational classifications for the uk. *Health Econ*. doi:10.1002/hec.3513

Koopmanschap, M., Burdorf, A., Jacob, K., Meerding, W. J., Brouwer, W., & Severens, H. (2005). Measuring productivity changes in economic evaluation: Setting the research agenda. *PharmacoEconomics*, *23*(1), 47–54. doi:10.2165/00019053-200523010-00004

Koopmanschap, M. A., & Rutten, F. F. (1996). A practical guide for calculating indirect costs of disease. *PharmacoEconomics*, *10*(5), 460–6.

Koopmanschap, M. A., Rutten, F. F., van Ineveld, B. M., & van Roijen, L. (1995). The friction cost method for measuring indirect costs of disease. *J Health Econ*, *14*(2), 171–89. doi:10.1016/0167-6296(94)00044-5

Koopmanschap, M. A., Rutten, F. F. H., van Ineveld, B. M., & van Roijen, L. (1997). Reply to johanneson's and karlsson's comment. *J Health Econ*, *16*(2), 257–259. doi:10.1016/s0167-6296(96)00524-3

Kristensen, T. S., Hannerz, H., Høgh, A., & Borg, V. (2005). The copenhagen psychosocial questionnaire-a tool for the assessment and improvement of the psychosocial work environment. *Scand J Work Environ Health*, 438–449.

Krogsbøll, L. T., Jørgensen, K., & Gøtzsche, P. C. (2013). General health checks in adults for reducing morbidity and mortality from disease. *JAMA*, *309*(23), 2489–2490. doi:10.1001/jama.2013.5039

Krol, M., & Brouwer, W. (2014). How to estimate productivity costs in economic evaluations. *Pharmacoeconomics*, *32*(4), 335–44. doi:10.1007/s40273-014-0132-3

Krol, M., Brouwer, W. B., Severens, J. L., Kaper, J., & Evers, S. M. (2012). Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects. *Soc Sci Med*, *75*(11), 1981–8. doi:10.1016/j.socscimed.2012.07.012

Krol, M., Stolk, E., & Brouwer, W. (2014). Predicting productivity based on eq-5d: An explorative study. *Eur J Health Econ*, *15*(5), 465–75. doi:10.1007/s10198-013-0487-y

Krol, M., Brouwer, W., & Rutten, F. (2013). Productivity costs in economic evaluations: Past, present, future. *PharmacoEconomics*, *31*(7), 537–549. doi:10.1007/s40273-013-0056-3

Kucherov, D., & Zavyalova, E. (2012). Hrd practices and talent management in the companies with the employer brand. *European Journal of Training and Development*, *36*(1), 86–104. doi:10.1108/03090591211192647

Kuorinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sorensen, F., Andersson, G., & Jorgensen, K. (1987). Standardised nordic questionnaires for the analysis of musculoskeletal symptoms. *Appl Ergon*, *18*. doi:10.1016/0003-6870(87)90010-x

Lahiri, S., Latif, S., & Punnett, L. (2013). An economic analysis of a safe resident handling program in nursing homes. *Am J Ind Med*, *56*(4), 469–78. doi:10.1002/ajim.22139

Lahiri, S., Low, C., & Barry, M. (2011). A business case evaluation of workplace engineering noise control: A net-cost model. *J Occup Environ Med*, *53*(3), 329–37. doi:10.1097/JOM.0b013e31820d19c0

Lahiri, S., Tempesti, T., & Gangopadhyay, S. (2016). Is there an economic case for training intervention in the manual material handling sector of developing countries? *J Occup Environ Med*, *58*(2), 207–214.

Lalloo, D., Demou, E., Smedley, J., Madan, I., Asanati, K., & Macdonald, E. B. (2018). Current research priorities for uk occupational physicians and occupational health researchers: A modified delphi study. *Occup Environ Med*, oemed-2018–105114. doi:10.1136/oemed-2018-105114

Lambrechts, M.-C., Ketterer, F., Symons, L., Mairiaux, P., Peremans, L., Remmen, R., ... Godderis, L. (2015). The approach taken to substance abuse by occupational physicians: A qualitative study on influencing factors. *Journal of occupational and environmental medicine*, *57*(11), 1228–1235.

Larsen, A. K., Holtermann, A., Mortensen, O. S., Punnett, L., Rod, M. H., & Jørgensen, M. B. (2015). Organizing workplace health literacy to reduce musculoskeletal pain and consequences. *BMC nursing*, *14*(1), 46.

Lawson, K. D., Fenwick, E. A., Pell, A. C., & Pell, J. P. (2010). Comparison of mass and targeted screening strategies for cardiovascular risk: Simulation of the effectiveness, cost-effectiveness and coverage using a cross-sectional survey of 3921 people. *Heart*, *96*(3), 208–212.

Lensberg, B. R., Drummond, M. F., Danchenko, N., Despiegel, N., & Francois, C. (2013). Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. *Clinicoecon Outcomes Res*, *5*, 565–73. doi:10.2147/CEOR.S44866

Lindstrom, M. J., & Bates, D. M. (1988). Newton—raphson and em algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, *83*(404), 1014–1022.

Lofland, J. H., Pizzi, L., & Frick, K. D. (2004). A review of health-related workplace productivity loss instruments. *Pharmacoeconomics*, *22*(3), 165–84. doi:10.2165/00019053-200422030-00003

Lucchini, R. G., & London, L. (2014). Global occupational health: Current challenges and the need for urgent action. *Ann Glob Health*, *80*(4), 251–6.

Luyten, J., Van de Sande, S., de Schrijver, K., Van Damme, P., & Beutels, P. (2012). Cost-effectiveness of hepatitis a vaccination for adults in belgium. *Vaccine*, *30*(42), 6070–80. doi:10.1016/j.vaccine.2012.07.049

Luyten, J., Steel, J., & Godderis, L. (2017). Economic evaluation of occupational health services: Necessary, challenging and promising. *Occup Environ Med*, *74*(11), 847–848. doi:10.1136/oemed-2017-104495

MacAuley, D. (2012). The value of conducting periodic health checks. British Medical Journal Publishing Group.

Mairiaux, P., Chanteux, A., Schleich, E., & Delwick, B. (2010). *Financement des missions de prévention des services externes de prévention et de protection au travail*. SPF Emploi, travail et concertation sociale.

Makrides, L., Smith, S., Allt, J., Farquharson, J., Szpilfogel, C., Curwin, S., ... Edington, D. (2011). The healthy lifeworks project: A pilot study of the economic analysis of a comprehensive workplace wellness program in a canadian government department. *J Occup Environ Med*, *53*(7), 799–805. doi:10.1097/JOM.0b013e318222af67

Manuguerra, M., & Heller, G. Z. (2010). Ordinal regression models for continuous scales. *The International Journal of Biostatistics*, *6*(1).

Manuguerra, M., Heller, G. Z., & Ma, J. (2020). Continuous ordinal regression for analysis of visual analogue scales: The r package ordinalcont. *Journal of Statistical Software*, *96*(1), 1–25.

Mascha, E. J., & Sessler, D. I. (2011). Equivalence and noninferiority testing in regression models and repeated-measures designs. *Anesthesia & Analgesia*, *112*(3), 678–687.

Mattke, S., Balakrishnan, A., Bergamo, G., & Newberry, S. J. (2007). A review of methods to measure health-related productivity loss. *Am J Manag Care*, *13*(4), 211–7.

McCurry, J. (2017). Japanese woman 'dies from overwork' after logging 159 hours of overtime in a month. Newspaper Article.

McEachan, R. R., Lawton, R. J., Jackson, C., Conner, M., Meads, D. M., & West, R. M. (2011). Testing a workplace physical activity intervention: A cluster randomized controlled trial. *Int J Behav Nutr Phys Act*, *8*, 29. doi:10.1186/1479-5868-8-29

Meenan, R. T., Vogt, T. M., Williams, A. E., Stevens, V. J., Albright, C. L., & Nigg, C. (2010). Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in hawaii. *J Occup Environ Med*, *52*(1), S8–S13. doi:10.1097/JOM.0b013e3181c81af9

Meijster, T., van Duuren-Stuurman, B., Heederik, D., Houba, R., Koningsveld, E., Warren, N., & Tielemans, E. (2011). Cost-benefit analysis in occupational health: A comparison of intervention scenarios for occupational asthma and rhinitis among bakery workers. *Occup Environ Med*, *68*(10), 739–745. doi:10.1136/oem.2011.064709

Mewes, J. C., Steuten, L. M., Groeneveld, I. F., de Boer, A. G., Frings-Dresen, M. H., MJ, I., & van Harten, W. H. (2015). Return-to-work intervention for cancer survivors: Budget impact and allocation of costs and returns in the netherlands and six major eu-countries. *BMC Cancer*, *15*(1), 899. doi:10.1186/s12885-015-1912-7

Miller, P., & Haslam, C. (2009). Why employers spend money on employee health: Interviews with occupational health and safety professionals from british industry. *Safety Sci*, *47*(2), 163–169. doi:10.1016/j.ssci.2008.04.001

Miller, T. R., Zaloshnja, E., & Spicer, R. S. (2007). Effectiveness and benefit-cost of peer-based workplace substance abuse prevention coupled with random testing. *Accid Anal Prev*, *39*(3), 565–73. doi:10.1016/j.aap.2006.10.001

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS medicine*, *6*(7), e1000097.

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... De Vet, H. C. (2010). The cosmin checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol*, *10*(1), 1–8. doi:10.1186/1471-2288-10-22

Molenberghs, G., & Verbeke, G. (2006). *Models for discrete longitudinal data*. New York, NY: Springer New York.

Morabia, A., & Zhang, F. F. (2004). History of medical screening: From concepts to action. *Postgraduate medical journal*, *80*(946), 463–469.

Mossink, J., & de Greef, M. (2002). *Inventory of socioeconomic costs of work accidents* (Report No. 9295007670). Office for Official Publications of the European Communities.

Mukhopadhyay, S., & Wendel, J. (2013). Evaluating an employee wellness program. *Int J Health Care Finance Econ*, *13*(3-4), 173–99. doi:10.1007/s10754-013-9127-4

Neuhaus, M., Eakin, E. G., Straker, L., Owen, N., Dunstan, D. W., Reid, N., & Healy, G. N. (2014). Reducing occupational sedentary time: A systematic review and meta-analysis of evidence on activity-permissive workstations. *Obes Rev*, *15*(10), 822–38. doi:10.1111/obr.12201

Nicholson, P. (2017). *Occupational health: The value proposition*. Society of Occupational Medicine.

Nicholson, S., Pauly, M. V., Polsky, D., Sharda, C., Szrek, H., & Berger, M. L. (2006). Measuring the effects of work loss on productivity with team production. *Health Econ*, *15*(2), 111–23. doi:10.1002/hec.1052

Niven, K. J. (2002). A review of the application of health economics to health and safety in healthcare. *Health Policy*, *61*(3), 291–304. doi:10.1016/S0168-8510(01)00224-X

Noben, C., Evers, S., Nieuwenhuijsen, K., Ketelaar, S., Gartner, F., Sluiter, J., & Smit, F. (2015). Protecting and promoting mental health of nurses in the hospital setting: Is it cost-effective from an employer's perspective? *Int J Occup Med Environ Health*, *28*(5), 891–900. doi:10.13075/ijomeh.1896.00465

Noben, C., Hoefsmit, N., Evers, S., de Rijk, A., Houkes, I., & Nijhuis, F. (2015). Economic evaluation of a new organizational rtw intervention to improve cooperation between sick-listed employees and their supervisors a field study. *J Occup Environ Med*, *57*(11), 1170–1177. doi:10.1097/Jom.0000000000000566

Noben, C., Smit, F., Nieuwenhuijsen, K., Ketelaar, S., Gärtner, F., Boon, B., ... Evers, S. (2014). Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud*, *51*(10), 1321–1331. doi:10.1016/j.ijnurstu.2014.01.017

Norrefalk, J. R., Ekholm, K., Linder, J., Borg, K., & Ekholm, J. (2008). Evaluation of a multiprofessional rehabilitation programme for persistent musculoskeletal-related pain: Economic benefits of return to work. *J Rehabil Med*, *40*(1), 15–22. doi:10.2340/16501977-0131

Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: Evidence from the university of minnesota. *Med Care*, *50*(9), 772–8. doi:10.1097/MLR.0b013e31825a8b1f

Ødegaard, F., & Roos, P. (2014). Measuring the contribution of workers' health and psychosocial work-environment on production efficiency. *Prod Oper Manag*, *23*(12), 2191–2208. doi:10.1111/poms.12242

Olsen, J. A., & Richardson, J. (1999). Production gains from health care: What should be included in cost-effectiveness analyses? *Soc Sci Med*, *49*(1), 17–26. doi:10.1016/s0277-9536(99)00116-1

Olsen, O., Albertsen, K., Nielsen, M. L., Poulsen, K. B., Gron, S. M. F., & Brunnberg, H. L. (2008). Workplace restructurings in intervention studies – a challenge for design, analysis and interpretation. *Bmc Med Res Methodol*, *8*(1), 39. doi:10.1186/1471-2288-8-39

Oostenbrink, J. B., Koopmanschap, M. A., & Rutten, F. F. (2002). Standardisation of costs: The dutch manual for costing in economic evaluations. *Pharmacoeconomics*, *20*(7), 443–54. doi:10.2165/00019053-200220070-00002

Organisation for Economic Co-operation and Development. (2001). *Measuring productivity: Measurement of aggregate and industry-level productivity growth: Oecd manual*. Paris: Organisation for Economic Co-operation and Development.

Osilla, K. C., dela Cruz, E., Miles, J. N., Zellmer, S., Watkins, K., Larimer, M. E., & Marlatt, G. A. (2010). Exploring productivity outcomes from a brief intervention for at-risk drinking in an employee assistance program. *Addict Behav*, *35*(3), 194–200. doi:10.1016/j.addbeh.2009.10.001

Osilla, K. C., Van Busum, K., Schnyer, C., Larkin, J. W., Eibner, C., & Mattke, S. (2012). Systematic review of the impact of worksite wellness programs. *Am J Manag Care*, *18*(2), e68–81.

Oude Hengel, K. M., Bosmans, J. E., Van Dongen, J. M., Bongers, P. M., Van der Beek, A. J., & Blatter, B. M. (2014). Prevention program at construction worksites aimed at improving health and work ability is cost-saving to the employer: Results from an rct. *Am J Ind Med*, *57*(1), 56–68.

Palumbo, M. V., Sikorski, E. A., & Liberty, B. C. (2013). Exploring the cost-effectiveness of unit-based health promotion activities for nurses. *Workplace Health Saf*, *61*(12), 514–520. doi:10.3928/21650799-20131206-01

Paté-Cornell, E. (2012). On "black swans" and "perfect storms": Risk analysis and management when statistics are not enough. *Risk Anal*, *32*(11), 1823–1833. doi:10.1111/j.1539-6924.2011.01787.x

Pauly, M. V., Nicholson, S., Xu, J., Polsky, D., Danzon, P. M., Murray, J. F., & Berger, M. L. (2002). A general model of the impact of absenteeism on employers and employees. *Health Econ*, *11*(3), 221–31. doi:10.1002/hec.648

Payne, K., McAllister, M., & Davies, L. M. (2013). Valuing the economic benefits of complex interventions: When maximising health is not sufficient. *Health Econ*, *22*(3), 258–271. doi:10.1002/hec.2795

Pedersen, M. S., & Arendt, J. N. (2014). Bargaining for health: A case study of a collective agreement-based health program for manual workers. *J Health Econ*, *37*(1), 123–36. doi:10.1016/j.jhealeco.2014.06.004

Pelletier, K. R. (2009). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update vii 2004-2008. *J Occup Environ Med*, *51*(7), 822–37. doi:10.1097/JOM.0b013e3181a7de5a

Pelletier, K. R. (2011). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update viii 2008 to 2010. *J Occup Environ Med*, *53*(11), 1310–31. doi:10.1097/JOM.0b013e3182337748

Philip, B. K. (1990). Parametric statistics for evaluation of the visual analog scale. *Anesthesia & Analgesia*, *71*(6), 710.

Phillips, C. J., Phillips Nee Buck, R., Main, C. J., Watson, P. J., Davies, S., Farr, A., ... Hale, J. (2012). The cost effectiveness of nhs physiotherapy support for occupational health (oh) services. *BMC Musculoskelet Disord*, *13*, 29. doi:10.1186/1471-2474-13-29

Piko, B. F. (2006). Burnout, role conflict, job satisfaction and psychosocial health among hungarian health care staff: A questionnaire survey. *Int J Nurs Stud*, *43*(3), 311–318. doi:10.1016/j.ijnurstu.2005.05.003

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-152.

Ponsonby, W. (2017). Global occupational health. *Occup Med (Lond)*, *67*(5), 331–333. doi:10.1093/occmed/kqx068

Prasad, M., Wahlqvist, P., Shikiar, R., & Shih, Y. C. (2004). A review of self-report instruments measuring health-related work productivity: A patient-reported outcomes perspective. *PharmacoEconomics*, *22*(4), 225–44.

Prosser, L. A., O'Brien, M. A., Molinari, N. A., Hohman, K. H., Nichol, K. L., Messonnier, M. L., & Lieu, T. A. (2008). Non-traditional settings for influenza vaccination of adults: Costs and cost effectiveness. *PharmacoEconomics*, *26*(2), 163–78.

Ravesteijn, B., Kippersluis, H. v., & van Doorslaer, E. (2018). The wear and tear on health: What is the role of occupation? *Health economics*, *27*(2), e69–e86.

Rebergen, D. S., Bruinvels, D. J., van Tulder, M. W., van der Beek, A. J., & van Mechelen, W. (2009). Cost-effectiveness of guideline-based care for workers with mental health problems. *J Occup Environ Med*, *51*(3), 313–22. doi:10.1097/JOM.0b013e3181990d8e

Ritchie, H., Ortiz-Ospina, E., Beltekian, D., Mathieu, E., Hasell, J., Macdonald, B., ... Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. https://ourworldindata.org/coronavirus.

Robbins, C. L., Keyserling, T. C., Pitts, S. B., Morrow, J., Majette, N., Sisneros, J. A., ... Dietz, P. M. (2013). Screening low-income women of reproductive age for cardiovascular disease risk factors. *J Womens Health (Larchmt)*, *22*(4), 314–21. doi:10.1089/jwh.2012.4149

Roe, R. A., & Ester, P. (1999). Values and work: Empirical findings and theoretical perspective. *Applied psychology*, *48*(1), 1–21.

Roelofs, P. D., Bierma-Zeinstra, S. M., van Poppel, M. N., van Mechelen, W., Koes, B. W., & van Tulder, M. W. (2010). Cost-effectiveness of lumbar supports for home care workers with recurrent low back pain: An economic evaluation alongside a randomized-controlled trial. *Spine (Phila Pa 1976)*, *35*(26), E1619–26. doi:10.1097/BRS.0b013e3181cf7244

Rongen, A., Robroek, S. J., van Lenthe, F. J., & Burdorf, A. (2013). Workplace health promotion: A meta-analysis of effectiveness. *American journal of preventive medicine*, *44*(4), 406–415.

Ros, M., Schwartz, S. H., & Surkiss, S. (1999). Basic individual values, work values, and the meaning of work. *Applied psychology*, *48*(1), 49–71.

Rosen, H. S., & Gayer, T. (2010). *Public finance* (9th edition). New York: Mc Graw Hill.

Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017. (n.d.).

Royal Decree of 30 January 2017 amending the Royal Decree of 28 May 2003 on the monitoring of employee health, Belgian Official Gazette, 6 February 2017. (n.d.). Legal Rule or Regulation.

Royal decree of 30 january 2017 amending the royal decree of 28 may 2003 on the monitoring of employee health, belgian official gazette, 6 february 2017. (n.d.).

Ruitenburg, M. M., Frings-Dresen, M. H., & Sluiter, J. K. (2016). How to define the content of a job-specific worker's health surveillance for hospital physicians? *Saf Health Work*, *7*(1), 18–31. doi:10.1016/j.shaw.2015.08.004

Sancroft. (2017). *Maximising the value of occupational health and safety and workplace wellness reporting for a global workforce: A practical guide for internationally operating employers.* International SOS Foundation.

Schaafsma, F. G., Mahmud, N., Reneman, M. F., Fassier, J. B., & Jungbauer, F. H. (2016). Pre-employment examinations for preventing injury, disease and sick leave in workers. *Cochrane Database Syst Rev*, (1), CD008881. doi:10.1002/14651858.CD008881.pub2

Schene, A. H., Koeter, M. W., Kikkert, M. J., Swinkels, J. A., & McCrone, P. (2007). Adjuvant occupational therapy for work-related major depression works: Randomized trial including economic evaluation. *Psychol Med*, *37*(3), 351–62. doi:10.1017/S0033291706009366

Schokkaert, E. (2015). How to introduce more (or better) ethical arguments in hta? *International journal of technology assessment in health care*, *31*(3), 111–112.

Schokkaert, E., Steel, J., & Van de Voorde, C. (2017). Out-of-pocket payments and subjective unmet need of healthcare. *Applied health economics and health policy*, *15*(5), 545–555. doi:10.1007/s40258-017-0331-0

Schultz, A. B., Chen, C. Y., & Edington, D. W. (2009). The cost and impact of health conditions on presenteeism to employers: A review of the literature. *PharmacoEconomics*, *27*(5), 365–78. doi:10.2165/00019053-200927050-00002

Seeley, P. A., Marklin, R. W., Usher, D., & Yager, J. (2008). Case study report: Business case for implementing battery-powered tools for direct-bury line workers at an electric power utility. *J Occup Environ Hyg*, *5*(8), D86–D91.

Selvarajah, S., Haniff, J., Kaur, G., Hiong, T. G., Bujang, A., Cheong, K. C., & Bots, M. L. (2013). Identification of effective screening strategies for cardiovascular disease prevention in a developing country: Using cardiovascular risk-estimation and risk-reduction tools for policy recommendations. *BMC Cardiovasc Disord*, *13*(1), 10.

Sen, A. (1992). *Inequality reexamined.* Oxford University Press.

Şen, C., Karli, N., Zarifoğlu, M., Pala, K., Seferoğlu, M., & Özkaya, G. (2012). Impact of migraine treatment on economic burden and workplace productivity among migraineurs employed in schools: A prospective study. *J Neurol Sci*, *28*(4), 465–474.

Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? implications for health economic evaluation. *BMJ*, *336*(7656), 1281. doi:10.1136/bmj.39569.510521.AD

Shirali, G., Shekari, M., & Angali, K. A. (2017). Assessing reliability and validity of an instrument for measuring resilience safety culture in sociotechnical systems. *Safety and Health at Work*.

Shrestha, N., Kukkonen-Harjula, K. T., Verbeek, J. H., Ijaz, S., Hermans, V., & Bhaumik, S. (2016). Workplace interventions for reducing sitting at work. *Cochrane Database Syst Rev*, (3). doi:10.1002/14651858.CD010912.pub3

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. doi:10.1126/science.3563507

Sluiter, J., Weel, A., & Hulshof, C. (2013a). Guidance preventive medical examination of workers (leidraad preventief medisch onderzoek van werkenden). *KMKA, NVAB, Coronel Institute for Occupational and Environmental Health, Academic Medical Center, Amsterdam*, 1–20.

Sluiter, J., Weel, A., & Hulshof, C. (2013b). Leidraad preventief medisch onderzoek van werkenden. *NVAB, Utrecht*.

Sluiter, J. K. (2017). The new why when designing mandatory medical examinations. *Occup Med (Lond)*, *67*(5), 325–327. doi:10.1093/occmed/kqx044

Smallman, C., & John, G. (2001). British directors perspectives on the impact of health and safety on corporate performance. *Safety Sci*, *38*(3), 227–239. doi:10.1016/S0925-7535(01)00003-0

Smedley, J., Harris, E. C., Cox, V., Ntani, G., & Coggon, D. (2013). Evaluation of a case management service to reduce sickness absence. *Occup Med*, *63*(2), 89–95. doi:10.1093/occmed/kqs223

Smith, K. J., Raymund, M., Nowalk, M. P., Roberts, M. S., & Zimmerman, R. K. (2010). Cost-effectiveness of pneumococcal polysaccharide vaccine among healthcare workers during an influenza pandemic. *Am J Manag Care*, *16*(3), 200–6.

Soler, R. E., Leeks, K. D., Razi, S., Hopkins, D. P., Griffith, M., Aten, A., ... Task Force on Community Preventive, S. (2010). A systematic review of selected interventions for worksite health promotion. the assessment of health risks with feedback. *Am J Prev Med*, *38*(2 Suppl), S237–62. doi:10.1016/j.amepre.2009.10.030

Sorensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H. (2012). Health literacy and public health: A systematic review

and integration of definitions and models. *BMC Public Health*, *12*, 80. doi:10.1186/1471-2458-12-80

Sørensen, K., Pelikan, J. M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., ... Uiters, E. (2015). Health literacy in europe: Comparative results of the european health literacy survey (hls-eu). *European journal of public health*, *25*(6), 1053–1058.

Spekle, E. M., Heinrich, J., Hoozemans, M. J., Blatter, B. M., van der Beek, A. J., van Dieen, J. H., & van Tulder, M. W. (2010). The cost-effectiveness of the rsi quickscan intervention programme for computer workers: Results of an economic evaluation alongside a randomised controlled trial. *BMC Musculoskelet Disord*, *11*, 259. doi:10.1186/1471-2474-11-259

Squires, H., Rick, J., Carroll, C., & Hillage, J. (2012). Cost-effectiveness of interventions to return employees to work following long-term sickness absence due to musculoskeletal disorders. *J Public Health-Uk*, *34*(1), 115–124. doi:10.1093/pubmed/fdr057

Steel, J., Godderis, L., & Luyten, J. (2017). *Research report: Economic evaluation in occupational health and safety*. Belgische Beroepsvereniging voor Arbeidsgeneesheren.

Steel, J., Godderis, L., & Luyten, J. (2018a). Methodological challenges in the economic evaluation of occupational health and safety programmes. *International journal of environmental research and public health*, *15*(11), 2606. doi:10.3390/ijerph15112606

Steel, J., Godderis, L., & Luyten, J. (2018b). Productivity estimation in economic evaluations of occupational health and safety interventions: A systematic review. *Scand J Work Environ Health*, *44*(5). doi:10.5271/sjweh.3715

Steel, J., Godderis, L., & Luyten, J. (2021a). Onderzoek in het kort: Online vragenlijst of fysiek bij de arbeidsarts? rapportering van werkgebonden gezondheidproblemen in belgische ziekenhuizen.

Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals.

Steel, J., Luyten, J., & Godderis, L. (2018). *Occupational health: The global evidence and value*. Society of Occupational Medicine (SOM); London.

Steel, J. S., Godderis, L., & Luyten, J. (2021a). Development and validation of a triage tool to follow-up workers' health. *Occup. Med. (Lond)*, *(under review)*.

Steel, J. S., Godderis, L., & Luyten, J. (2021b). Disclosure in online vs. face-to-face occupational health screenings: A cross-sectional study in belgian hospital employees. *International Journal of Environmental Research and Public Health*, *18*(4), 1460. doi:10.3390/ijerph18041460

Steen, A., & Welch, D. (2011). Are accounting metrics applicable to human resources? the case of return on valuing assignments. *Australasian Accounting, Business and Finance Journal*, *5*(3), 57–72.

Street, R. L., & Haidet, P. (2011). How well do doctors know their patients? factors affecting physician understanding of patients' health beliefs. *J Gen Intern Med*, *26*(1), 21–27. doi:10.1007/s11606-010-1453-3

Strömberg, C., Aboagye, E., Hagberg, J., Bergstrom, G., & Lohela-Karlsson, M. (2017). Estimating the effect and economic impact of absenteeism, presenteeism, and work environment-related problems on reductions in productivity from a managerial perspective. *Value Health*, *20*(8), 1058–1064. doi:10.1016/j.jval.2017.05.008

Sturkenboom, I. H., Hendriks, J. C., Graff, M. J., Adang, E. M., Munneke, M., Nijhuis-van der Sanden, M. W., & Bloem, B. R. (2015). Economic evaluation of occupational therapy in parkinson's disease: A randomized controlled trial. *Mov Disord*, *30*(8), 1059–67. doi:10.1002/mds.26217

Susitaival, P., Flyvholm, M.-A., Meding, B., Kanerva, L., Lindberg, M., Svensson, Å., & Olafsson, J. (2003). Nordic occupational skin questionnaire (nosq-2002): A new tool for surveying occupational skin diseases and exposure. *Contact dermatitis*, *49*(2), 70–76.

Svensson, E. (2000). Comparison of the quality of assessments using continuous and discrete ordinal rating scales. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *42*(4), 417–434.

Syverson, C. (2011). What determines productivity? *J Econ Lit*, *49*(2), 326–365. doi:10.1257/jel.49.2.326

Taimela, S., Justen, S., Aronen, P., Sintonen, H., Laara, E., Malmivaara, A., ... Aro, T. (2008). An occupational health intervention programme for workers at high risk for sickness absence. cost effectiveness analysis based on a randomised controlled trial. *Occup Environ Med*, *65*(4), 242–248. doi:10.1136/oem.2007.033167

Takala, J., Hämäläinen, P., Saarela, K. L., Yun, L. Y., Manickam, K., Jin, T. W., ... Lim, S., et al. (2014). Global estimates of the burden of injury and illness at work in 2012. *Journal of occupational and environmental hygiene*, *11*(5), 326–337.

Tallant, A., Porter, C. K., Putnam, S. D., Tribble, D. R., Hooper, T. I., & Riddle, M. S. (2014). Relative cost-effectiveness of a norovirus vaccine in the deployed military setting compared to a vaccine against campylobacter sp., etec, and shigella sp. *Vaccine*, *32*(40), 5156–62. doi:10.1016/j.vaccine.2014.07.070

Tan, S. S., Bouwmans, C. A., Rutten, F. F., & Hakkaart-van Roijen, L. (2012). Update of the dutch manual for costing in economic evaluations. *Int J Technol Assess Health Care*, *28*(2), 152–8. doi:10.1017/S0266462312000062

Tang, K. (2015). Estimating productivity costs in health economic evaluations: A review of instruments and psychometric evidence. *Pharmacoeconomics*, *33*(1), 31–48.

Task Force on Community Preventive Services. (2005). *The guide to community preventive services: What works to promote health?* Oxford University Press.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., ... Stewart-Brown, S. (2007). The warwick-edinburgh mental well-being scale (wemwbs): Development and uk validation. *Health and Quality of life Outcomes*, *5*(1), 63.

The World Bank Group. (2017). Databank - gdp (current us dollars). Web Page.

Theodore, B. R., Mayer, T. G., & Gatchel, R. J. (2015). Cost-effectiveness of early versus delayed functional restoration for chronic disabling occupational musculoskeletal disorders. *J Occup Rehabil*, *25*(2), 303–15. doi:10.1007/s10926-014-9539-0

Thijs, K. M., de Boer, A. G., Vreugdenhil, G., van de Wouw, A. J., Houterman, S., & Schep, G. (2012). Rehabilitation using high-intensity physical training and long-term return-to-work in cancer survivors. *J Occup Rehabil*, *22*(2), 220–9. doi:10.1007/s10926-011-9341-1

Thom, D. H., Ribisl, K. M., Stewart, A. L., Luke, D. A., & Physicians, T. S. T. S. (1999). Further validation and reliability testing of the trust in physician scale. *Med Care*, 510–517.

Tompa, E., Dolinschi, R., Alamgir, H., Sarnocinska-Hart, A., & Guzman, J. (2016). A cost-benefit analysis of peer coaching for overhead lift use in the long-term care sector in canada. *Occup Environ Med*, *73*(5), 308–14. doi:10.1136/oemed-2015-103134

Tompa, E., Dolinschi, R., & de Oliveira, C. (2006). Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil*, *16*(3), 375–400. doi:10.1007/s10926-006-9035-2

Tompa, E., Dolinschi, R., de Oliveira, C., & Irvin, E. (2009). A systematic review of occupational health and safety interventions with economic analyses. *J Occup Environ Med*, *51*(9), 1004–23. doi:10.1097/JOM.0b013e3181b34f60

Tompa, E., Dolinschi, R., & Natale, J. (2013). Economic evaluation of a participatory ergonomics intervention in a textile plant. *Appl Ergon*, *44*(3), 480–7. doi:10.1016/j.apergo.2012.10.019

Tompa, E., Trevithick, S., & McLeod, C. (2007). Systematic review of the prevention incentives of insurance and regulatory mechanisms for occupational health and safety. *Scand J Work Environ Health*, *33*(2), 85–95.

Tompa, E., Verbeek, J., van Tulder, M., & de Boer, A. (2010). Developing guidelines for good practice in the economic evaluation of occupational

safety and health interventions. *Scand J Work Environ Health*, *36*(4), 313–8. doi:10.5271/sjweh.3009

Tompa, E., Culyer, A. J., & Dolinschi, R. (2008). *Economic evaluation of interventions for occupational health and safety: Developing good practice*. doi:10.5271/sjweh.3009

Torbeyns, T., Bailey, S., Bos, I., & Meeusen, R. (2014). Active workstations to fight sedentary behaviour. *Sports Med*, *44*(9), 1261–73. doi:10.1007/s40279-014-0202-x

Torrance, I., & Heron, R. (2017a). Occupational health should be part of the nhs. British Medical Journal Publishing Group.

Torrance, I., & Heron, R. (2017b). Occupational health should be part of the nhs. Generic. British Medical Journal Publishing Group.

Trogdon, J. G., Finkelstein, E. A., Hylands, T., Dellea, P. S., & Kamal-Bahl, S. J. (2008). Indirect costs of obesity: A review of the current literature. *Obes Rev*, *9*(5), 489–500. doi:10.1111/j.1467-789X.2008.00472.x

Troman, G., & Woods, P. (2000). Careers under stress: Teacher adaptations at a time of intensive reform. *J Educ Change*, *1*(3), 253–275. doi:10.1007/978-90-481-2358-2_7

UCLA Insitute for Digital Research and Education. (n.d.). Introduction to linear mixed models. Accessed on February 19, 2021. https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/.

Uegaki, K., de Bruijne, M. C., Lambeek, L., Anema, J. R., van der Beek, A. J., van Mechelen, W., & van Tulder, M. W. (2010). Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. *Scand J Work Environ Health*, *36*(4), 273–88. doi:10.5271/sjweh.3017

Uegaki, K., de Bruijne, M. C., van der Beek, A. J., van Mechelen, W., & van Tulder, M. W. (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*(1), 90–9. doi:10.1007/s10926-010-9258-0

Valenty, M., Homère, J., Lemaitre, A., Plaine, J., Ruhlman, M., Cohidon, C., & Imbernon, E. (2015a). Surveillance programme for uncompensated work-related diseases in france. *Occupational Medicine*, *65*(8), 642–650.

Valenty, M., Homère, J., Lemaitre, A., Plaine, J., Ruhlman, M., Cohidon, C., & Imbernon, E. (2015b). Surveillance programme for uncompensated work-related diseases in france. *Occup Med (Lond)*, *65*(8), 642–650. doi:10.1093/occmed/kqv104

Van Buynder, P. G., Konrad, S., Kersteins, F., Preston, E., Brown, P. D., Keen, D., & Murray, N. J. (2015). Healthcare worker influenza immunization vaccinate or mask policy: Strategies for cost effective implementation and

subsequent reductions in staff absenteeism due to illness. *Vaccine*, *33*(13), 1625–1628.

Van Den Hout, W. B., De Buck, P. D. M., & Vliet Vlieland, T. P. M. (2007). Cost-utility analysis of a multidisciplinary job retention vocational rehabilitation program in patients with chronic arthritis at risk of job loss. *Arthritis Care Res*, *57*(5), 778–786.

van der Molen, H. F., Stocks, S. J., & Frings-Dresen, M. H. (2016). Exploring study designs for evaluation of interventions aimed to reduce occupational diseases and injuries. *Saf Health Work*, *7*(1), 83–85. doi:10.1016/j.shaw.2015.09.002

Van Duijn, M., Eijkemans, M. J., Koes, B. W., Koopmanschap, M. A., Burton, K. A., & Burdorf, A. (2010). The effects of timing on the cost-effectiveness of interventions for workers on sick leave due to low back pain. *Occup Environ Med*, *67*(11), 744–750.

Van Schaaijk, A., Boschman, J., Frings-Dresen, M., & Sluiter, J. (2017). Appraisal of work ability in relation to job-specific health requirements in ambulance workers. *Int J Environ Res Public Health*, *90*(1), 123–131. doi:10.1007/s00420-016-1181-z

Van Sprundel, M. (, Van Regenmortel, A. (, Muylaert, K. (, Borgers, E., Cambré, L., De Schryver, A., ... Vanthielen, K. (2016). *Arbeid en gezondheid: Een delicaat evenwicht?* Antwerpen: Intersentia.

van Dongen, J. M., Proper, K. I., van Wier, M. F., van der Beek, A. J., Bongers, P. M., van Mechelen, W., & van Tulder, M. W. (2011). Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev*, *12*(12), 1031–49. doi:10.1111/j.1467-789X.2011.00925.x

van Dongen, J. M., Proper, K. I., van Wier, M. F., van der Beek, A. J., Bongers, P. M., van Mechelen, W., & van Tulder, M. W. (2012). A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health*, *38*(5), 393–408. doi:10.5271/sjweh.3275

van Dongen, J. M., van Berkel, J., Boot, C. R., Bosmans, J. E., Proper, K. I., Bongers, P. M., ... van Wier, M. F. (2016). Long-term cost-effectiveness and return-on-investment of a mindfulness-based worksite intervention: Results of a randomized controlled trial. *J Occup Environ Med*, *58*(6), 550–60. doi:10.1097/JOM.0000000000000736

van Dongen, J. M., Tompa, E., Clune, L., Sarnocinska-Hart, A., Bongers, P. M., van Tulder, M. W., ... van Wier, M. F. (2013). Bridging the gap between the economic evaluation literature and daily practice in occupational health: A qualitative study among decision-makers in the healthcare sector. *Implement Sci*, *8*(1), 57. doi:10.1186/1748-5908-8-57

van Eijsden, M. D., Gerhards, S. A., de Bie, R. A., & Severens, J. L. (2009). Cost-effectiveness of postural exercise therapy versus physiotherapy in computer screen-workers with early non-specific work-related upper limb disorders (wruld); a randomized controlled trial. *Trials*, *10*, 103. doi:10.1186/1745-6215-10-103

van Oostrom, S., Heymans, M., Vet, H., Tulder, M., Mechelen, W., & Anema, J. (2010). Economic evaluation of a workplace intervention for sick-listed employees with distress. *Occup Environ Med*, *67*(9), 603–610.

van Wier, M. F., Dekkers, J. C., Bosmans, J. E., Heymans, M. W., Hendriksen, I. J., Pronk, N. P., ... van Tulder, M. W. (2012). Economic evaluation of a weight control program with e-mail and telephone counseling among overweight employees: A randomized controlled trial. *Int J Behav Nutr Phys Act*, *9*, 112. doi:10.1186/1479-5868-9-112

van Wier, M. F., Verweij, L. M., Proper, K. I., Hulshof, C. T., van Tulder, M. W., & van Mechelen, W. (2013). Economic evaluation of an occupational health care guideline for prevention of weight gain among employees. *J Occup Environ Med*, *55*(9), 1100–9. doi:10.1097/JOM.0b013e31829b279a

Verbeek, J., Pulliainen, M., & Kankaanpaa, E. (2009). A systematic review of occupational safety and health business cases. *Scand J Work Environ Health*, *35*(6), 403–12.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer.

Vermeulen, S. J., Heymans, M. W., Anema, J. R., Schellart, A. J. M., van Mechelen, W., & van der Beek, A. J. (2013). Economic evaluation of a participatory return-to-work intervention for temporary agency and unemployed workers sick-listed due to musculoskeletal disorders. *Scand J Work Environ Health*, *39*(1), 46–56. doi:10.5271/sjweh.3314

von Thiele Schwarz, U., & Hasson, H. (2012). Effects of worksite health interventions involving reduced work hours and physical exercise on sickness absence costs. *J Occup Environ Med*, *54*(5), 538–44. doi:10.1097/JOM.0b013e31824e11cd

Waddell, G., & Burton, A. K. (2006). Is work good for your health and well-being?

Walker, E., & Nowacki, A. S. (2011). Understanding equivalence and noninferiority testing. *Journal of general internal medicine*, *26*(2), 192–196.

Ware Jr, J. E., & Sherbourne, C. D. (1992). The mos 36-item short-form health survey (sf-36): I. conceptual framework and item selection. *Med Care*, 473–483. doi:10.1097/00005650-199206000-00002

Wendt, J. K., Tsai, S. P., Bhojani, F. A., & Cameron, D. L. (2010). The shell disability management program: A five-year evaluation of the impact

on absenteeism and return-on-investment. *J Occup Environ Med*, *52*(5), 544–50. doi:10.1097/JOM.0b013e3181dbc8c2

Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in nursing & health*, *13*(4), 227–236.

Wilson, C. J., Deane, F. P., Ciarrochi, J. V., & Rickwood, D. (2005). Measuring help seeking intentions: Properties of the general help seeking questionnaire.

World Health Organization. (1995). *Global strategy on occupational health for all: The way to health at work, recommendation of the second meeting of the who collaborating centres in occupational health, 11-14 october 1994, beijing, china*. World Health Organization.

Wynn, P. (2013). Periodic health checks in the workplace—is it time to change the prescription? Oxford University Press UK.

Yen, L., Schultz, A. B., Schaefer, C., Bloomberg, S., & Edington, D. W. (2010). Long-term return on investment of an employee health enhancement program at a midwest utility company from 1999 to 2007. *Int J Workplace Health Manag*, *3*(2), 79–96. doi:10.1108/17538351011054998

Zhang, W., Bansback, N., & Anis, A. H. (2011). Measuring and valuing productivity loss due to poor health: A critical review. *Soc Sci Med*, *72*(2), 185–92. doi:10.1016/j.socscimed.2010.10.026

Zhang, W., Sun, H., Woodcock, S., & Anis, A. (2015). Illness related wage and productivity losses: Valuing 'presenteeism'. *Soc Sci Med*, *147*, 62–71. doi:10.1016/j.socscimed.2015.10.056

Zimmerman, R. K., Wiringa, A. E., Nowalk, M. P., Lin, C. C. J., Rousculp, M. D., Mitgang, E. A., & Lee, B. Y. (2012). The comparative value of various employer-sponsored influenza vaccination clinics. *J Occup Environ Med*, *54*(9), 1107–1117. doi:10.1097/JOM.0b013e3182677d34

Zwerling, A., van den Hof, S., Scholten, J., Cobelens, F., Menzies, D., & Pai, M. (2012). Interferon-gamma release assays for tuberculosis screening of healthcare workers: A systematic review. *Thorax*, *67*(1), 62–70. doi:10.1136/thx.2010.143180

CURRICULUM - Jonas Steel

Leuven Institute for Healthcare Policy (LIHP)
Department of Public Health and Primary Care, KU Leuven
Kapucijnenvoer 35 bus 7001 (block D room 05.04)
3000 Leuven, Belgium

Phone: +32 16 37 30 79
Secr.: +32 16 32 12 59

Email: jonas.steel@kuleuven.be
URL: https://www.kuleuven.be/wieiswie/nl/person/00084228

Born: October 7, 1991—Beernem, Belgium
Nationality: Belgian

Current position

PhD Student, Leuven Institute for Healthcare Policy, KU Leuven

Areas of specialisation

Occupational Health; Economic Evaluation

Work experience

Leuven Institute for Healthcare Policy, KU Leuven	2017-2021
Research Associate Department of Economics, KU Leuven	2015-2016
Internship Voka Health Community	2015

Education

Msc in Economics, KU Leuven	2014-2015
Msc in Policy Economics, KU Leuven	2013-2014
MA in Philosophy, KU Leuven	2009-2013

Communication

Oral Presentation at EPICOH, Edinburgh	2017

Teaching

Seminars Health Economics[E0G63a], KU Leuven	2020

Courses

Decision Analytic Modelling for Economic Evaluation 2018 - Advanced (University of York)	2018

Last updated: August 17, 2021

List of publications

Articles

Luyten, J., Steel, J., & Godderis, L. (2017). Economic evaluation of occupational health services: Necessary, challenging and promising. *Occup Environ Med*, *74*(11), 847–848. doi:10.1136/oemed-2017-104495

Schokkaert, E., Steel, J., & Van de Voorde, C. (2017). Out-of-pocket payments and subjective unmet need of healthcare. *Applied health economics and health policy*, *15*(5), 545–555. doi:10.1007/s40258-017-0331-0

Steel, J., Godderis, L., & Luyten, J. (2018a). Methodological challenges in the economic evaluation of occupational health and safety programmes. *International journal of environmental research and public health*, *15*(11), 2606. doi:10.3390/ijerph15112606

Steel, J., Godderis, L., & Luyten, J. (2018b). Productivity estimation in economic evaluations of occupational health and safety interventions: A systematic review. *Scand J Work Environ Health*, *44*(5). doi:10.5271/sjweh.3715

Steel, J., Godderis, L., & Luyten, J. (2021b). Short-term effectiveness of face-to-face periodic occupational health screening versus electronic screening with targeted follow-up. results from a quasi-randomised controlled trial in four belgian hospitals.

Steel, J. S., Godderis, L., & Luyten, J. (2021a). Development and validation of a triage tool to follow-up workers' health. *Occup. Med. (Lond)*, *(under review)*.

Steel, J. S., Godderis, L., & Luyten, J. (2021b). Disclosure in online vs. face-to-face occupational health screenings: A cross-sectional study in belgian hospital employees. *International Journal of Environmental Research and Public Health*, *18*(4), 1460. doi:10.3390/ijerph18041460

Other

Steel, J., Godderis, L., & Luyten, J. (2017). *Research report: Economic evaluation in occupational health and safety*. Belgische Beroepsvereniging voor Arbeidsgeneesheren.

Steel, J., Godderis, L., & Luyten, J. (2021a). Onderzoek in het kort: Online vragenlijst of fysiek bij de arbeidsarts? rapportering van werkgebonden gezondheidproblemen in belgische ziekenhuizen.

Steel, J., Luyten, J., & Godderis, L. (2018). *Occupational health: The global evidence and value*. Society of Occupational Medicine (SOM); London.

Acknowledgements, personal contribution and conflict of interest statements

Acknowledgements

Chapter 2: The authors wish to thank Liesel Boutsen and Nina Smets for conducting and analysing the focus groups, and the Leuven Language Institute, a professional paid service offered by the university, for its assistance with minimising language barriers. J.S. and L.B. carried out the systematic review (screening and eligibility assessment) independently, disagreements were discussed until consensus was found. J.S. extracted data.

Chapter 3: The authors wish to express our gratitude to Annouschka Laenen for her assistance with the statistical analyses.

Chapter 4: The authors wish to thank the editor and the anonymous reviewers for their insightful and clarifying feedback, which greatly contributed to the quality of this article.

Personal Contribution

Chapter 2: Jonas Steel conducted the systematic review, Delphi panel, focus group, and analysed the validation studies. All authors were involved in the design of the study and drafting of the final paper.

Chapter 3: Jonas Steel: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - Original Draft, Writing - Review

& Editing, Visualization Lode Godderis: Conceptualization, Methodology, Recruitment of partners and participants, Writing - Review & Editing, Jeroen Luyten: Conceptualization, Methodology, Recruitment of partners and participants, Review & Editing,

Chapter 4, Section 1: Jonas Steel conducted the search strategy, data extraction and analyses. All authors were involved in the design of the study and drafting of the final paper. Section 2: Conceptualisation, J.S. and J.L. and L.G.; Funding Acquisition, J.L. and L.G.; Writing—Original Draft Preparation, J.S.; Writing—Review and Editing, J.L. and L.G.; Supervision, J.L. and L.G.

Conflict of interest statements

Chapter 2: This study was sponsored by the Belgian Association for Occupational Physicians, and three external services for prevention and protection at work: IDEWE, Liantis, and Mensura. The funders had no role in the analyses or interpretation of data; in the writing of the manuscript, nor the decision to publish the results.

Chapter 3: Dr. Godderis, Dr. Luyten and Jonas Steel report grants from Belgian Professional Association for Occupational Medicine, Idewe, Liantis & Mensura, during the conduct of the study; and Lode Godderis became CEO of the occupational health service IDEWE after the study was completed. The funders had no role in the analyses or interpretation of data; in the writing of the manuscript, nor the decision to publish the results.

Chapter 4, Section 1 and 2: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.



FACULTY OF MEDICINE
DEPARTMENT OF PUBLIC HEALTH AND PRIMARY CARE
LEUVEN INSTITUTE FOR HEALTHCARE POLICY
Kapucijnenvoer 35
B-3000 Leuven
jonas.steel@kuleuven.be
https://ligb.be