The Economics of Presenteeism in the context of Rheumatoid Arthritis, Ankylosing Spondylitis and Psoriatic Arthritis

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2017

Cheryl J. Jones
School of Health Sciences

List of Contents

List of Tables

List of Figures

Abstract

Background

Presenteeism is an economic concept that is difficult to identify, measure, and value. Rheumatoid arthritis (RA), ankylosing spondylitis (AS) and psoriatic arthritis (PsA) are three chronic auto-immune conditions that increase levels of presenteeism. Workplace interventions (WPIs) help individuals to manage their health condition at work. Existing methods used to quantify the impact of presenteeism are unable to adequately inform the employer of the productive benefits of WPIs. The overall aim of this thesis was to appraise current methods used to quantify presenteeism and to develop methods to value the impact of presenteeism suitable for use in economic evaluations (EE) of WPIs.

Methods

Two systematic reviews were conducted: 1) to assess the extent to which self-report measure of presenteeism were underpinned by economic theory; and 2) to explore if, and how, productivity was quantified and included in EE of WPIs for musculoskeletal conditions (MSDs). Thematic analysis methods were used to analyse qualitative data collected from working individuals with RA, AS or PsA (n=22) that explored the extent to which measures of health status (EQ5D; SF6D) and capability (ICECAP-A) capture the impact on ability to work caused by RA, AS or PsA. Econometric methods were used to specify prediction models that included measures of health status, capability and presenteeism, using a sample of 542 working people with RA and AS.

Results

The first systematic review identified 24 self-report measures of presenteeism; all, except one measure were not underpinned by economic theory. The second systematic review identified 20 EE of WPIs for MSDs. Absenteeism was included in all studies (n=20); however, presenteeism was included in only four. The qualitative data confirmed measures of health status and capability had the ability to capture those aspects of RA, AS and PsA that impact an individual's ability to work. The best performing prediction model used an OLS specification including SF6D, age and gender to predict presenteeism measured by the WPAI.

Conclusion

The results suggest that HRQoL measures, specifically the SF6D, can be used to capture and predict levels of presenteeism caused by RA, AS and PsA.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copy Right Statement

Abbreviations

ACR	American College of Rheumatology
AE	Absolute Error
ALWQ or ARLWQ	Angina- Related Limitations at Work Questionnaire
AMA	American Medical Association
APAWHI	American Productivity Audit and Work and Health Interview
AS/ SpA	Ankylosing Spondylitis
BWS	Best Worst Scaling
CBA	Cost Benefit Analysis
CCA	Cost Consequence Analysis
CCG	Clinical Commissioning Groups
CEA	Cost Effectiveness Analysis
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CHEC	Consensus on Health Economic Criteria
CHEERS	Consolidated Health Economics Evaluation Reporting Standards
CLAD	Censored Least Absolute Deviations
COSMIN	Consensus-based Standards for the selection of health Measurement Instruments
CRD	Centre for Reviews and Dissemination
CUA	Cost Utility Analysis
DCEs	Discrete Choice Experiments
csDMARDs	synthetic Disease Modifying Anti-Rheumatic Drugs
EWPS	Endicott Work Productivity Scale
EHCA	Employer Health Coalition of Tampa Assessment Instrument
EQ5D	EuroQol Five Dimensions
EQ5D-3L	EuroQol Five Dimensions Three Levels
EQ5D-5L	EuroQol Five Dimensions Five Levels
EULAR	European League Against Rheumatism
FAM	Framework Analysis Methods
FCA	Friction Cost Approach
FACS	Functional Ability Confidence Scale
GDP	Gross Domestic Product
GEEs	Generalised Estimating Equations
GHQ	General Health Questionnaire
GP	General Practitioner
GTM	Grounded Theory Method
HCA	Human Capital Approach
HEED	Health Economic Evaluation Database
HLQ	Health and Labour Questionnaire
HPQ	Health and Productivity Questionnaire
HRPQ-D	Health-Related Productivity Questionnaire Diary
HRQoL	Health Related Quality of Life
HTA	Health Technology Appraisal

HWQ	Health and Work Questionnaire
IA	Inflammatory Arthritis
ICECAP-A	ICEpop CAPability measure for Adults
ICECAP-O	ICEpop CAPability measure for Older people
ICER	Incremental Cost Effectiveness Ratio
iPCQ	iMTA Productivity Cost Questionnaire
LAD	Least Absolute Deviation
LEAPS	Lam Employment Absence and Productivity Scale
LIFE	Longitudinal Interval Follow-up Evaluation
LSRDs	Liebowitz Disability Self-Rating Scale
MAE	Mean Absolute Error
MDAQ	Migraine Disability Assessment Questionnaire
MIDAS	Migraine Disability Assessment Questionnaire
MSD	Musculoskeletal Disorders
MWPLQ	Migraine Work and Productivity Loss Questionnaire
NAM	Narrative Analysis Methods
NASS	National Ankylosing Spondylitis Society
NHS	National Health Service
NHS EED	National Health Service Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIMH-PQ	National Institute of Mental Health Panic Questionnaire
NPV	Net Present Value
NRAS	National Rheumatoid Arthritis Society
NSB	Net Social Benefit
OA	Osteoarthritis
OT	Osterhaus Technique
OMERACT	Outcomes Measures in Rheumatology group
ORQ	Occupational Role Questionnaire
OFS	Occupational Functioning Questionnaire
OLS	Ordinary Least Squares
ONS	Office of National Statistics
PsAZZ	Psoriasis Association
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PROMs	Patient Reported Outcomes Measures
PRODISQ	Productivity and Disease Questionnaire
PsA	Psoriatic Arthritis
PSS	Personal Social Services
QALYs	Quality Adjusted Life Years
QoL	Quality of Life
QIC	Quasi Likelihood under Independence Model Criterion
QICC	Corrected Quasi Likelihood under Independence Model Criterion
QQ Method	Quantity and Quality Method
RA	Rheumatoid Arthritis

RAPID3	Routine Assessment of Patient Index Data Three
RCT	Randomised Control Trial
RMSE	Root Mean Square Error
RUG	Rheumatoid Arthritis User Group
SAS-SR	Social Adjustment Scale Self-Report
S.A	Sensitivity Analysis
SAHAPS	Stanford/American Health Association Presenteeism Scale
SF-12	Short Form 12
SF-36	Short Form 36
SF6D	Short Form Six Dimensions
SF-HLQ	Short-Form Health and Labour Questionnaire
SPS	Stanford Presenteeism Scale
SPS – 6	Stanford Presenteeism Scale Six
SPS – 13	Stanford Presenteeism Scale Thirteen
SME	Small/ Medium Size Enterprises
TTO	Time Trade Off
UHI	Unnamed Hepatitis Instrument
UK	United Kingdom
USA	United States of America
VAS	Visual Analogue Scale
VOLP	Valuation of Lost Productivity
WAI	Work Ability Index
WALS	Work Activity Limitation Scale
WBA-P	Well-Being Assessment for Productivity
WIS	Work Instability Scale
WIS – AS	Work Instability Scale- Ankylosing Spondylitis
WIS – RA	Work Instability Scale- Rheumatoid arthritis
WLQ	Work Limitations Questionnaire
WLQ-16	Work Limitations Questionnaire-16 questions
WLQ-25	Work Limitations Questionnaire-25 questions
WLQ-J	Working Limitations Questionnaire – Japan
WPAI	Work Productivity Activity Impairment Questionnaire
WPAI:AR	Work Activity Impairment Questionnaire: Allergic Rhinitis
WPAI: CD	Work Activity Impairment Questionnaire : Crohn’s Disease
WPAI:GERD	Work Activity Impairment Questionnaire: Gastroesophageal Reflux Disease
WPAI: GH	Work Activity Impairment Questionnaire: General Health
WPAI: IBS	Work Activity Impairment Questionnaire: Irritable Bowel Syndrome
WPAI: SpA	Work Activity Impairment Questionnaire: Ankylosing Spondylitis
WPAI: RA	Work Activity Impairment Questionnaire: Rheumatoid Arthritis
WPI	Worker Productivity Index
WPIs	Workplace Interventions

WPI-MSD	Workplace Interventions for Musculoskeletal Disorders
WPS-AS	Work Productivity Survey for Ankylosing Spondylitis
WPS-RA	Work Productivity Survey for Rheumatoid Arthritis
WPSI	Work Productivity Short Inventory
WRFQ	Work Role Functioning Questionnaire
WTP	Willingness to Pay

Presentations

- Oral presentation ‘Health Economic Evaluation’, Centre for Musculoskeletal Research seminar series, The University of Manchester (January 2015)

- Oral presentation ‘The Economics of Presenteeism in the context of Rheumatoid Arthritis, Ankylosing Spondylitis and Psoriatic Arthritis; Quantitative Ideas’, Centre for Health Evaluation & Outcomes Sciences (CHEOS), UBC, Vancouver (February 2016)

- Oral Presentation ‘Exploring the impact of health status and well-being of people with inflammatory arthritis on presenteeism in the workplace: A Qualitative Study’, Health Economists’ Study Group (HESG), Birmingham January (2017)

- Oral Presentation ‘The Economics of Presenteeism in Rheumatoid Arthritis, Ankylosing Spondylitis, and Psoriatic Arthritis’. ARUK/MRC Centre for Musculoskeletal Health and Work, The University of Southampton (January 2017)

- Oral Presentation ‘Exploring the impact of health status and capability of employers with inflammatory arthritis on presenteeism in the workplace: A qualitative study’, The Lancaster Health Hub, The University of Lancaster, (October 2017)

- Poster Presentation ‘Assessing the relationship between health status and capability on the level of presenteeism suitable for economic evaluations of WPIs’, Health Economists’ Study Group (HESG), City University, London (January 2018)

Peer reviewed publication

- Jones, C., K. Payne, B. Gannon, and S.Verstappen. 2016. Economic Theory and Self-Reported Measures of Presenteeism in Musculoskeletal Disease. *Current Rheumatology Reports.* 18:53. This paper reports the findings from Chapter Three, see Appendix 1.1

- Jones, C., K, Payne, B. Gannon, S. Verstappen. 2017. Exploring the impact of health status and well-being of people with inflammatory arthritis on presenteeism in the workplace: A Qualitative Study’. *Annals of the Rheumatic Diseases*. 76 supplement 2: (437)

Acknowledgements

I must start by thanking my supervisors Professor Katherine Payne, Dr Suzanne Verstappen, and Professor Brenda Gannon. Their dedication, commitment, and unwavering level of support throughout my thesis have been instrumental towards helping me to complete this PhD programme. I cannot begin to thank them enough. I would also like to thank Alex Thompson for agreeing to step in as a technical supervisor providing me with sound advice for the quantitative component of this thesis.

My gratitude is also extended to the Centre for Musculoskeletal Health and Work (CMHW), the University of Southampton, for providing me with the thesis programme and extra funding to secure my placement at the Centre for Health Evaluation and Outcome Sciences based in Vancouver, Canada. Colleagues at the CMHW have been exceptionally supportive and encouraging of my work and progress as an early researcher during my three year thesis programme.

I would also like to give Yiu-Shing Lau a very special thank you for his generous support throughout my PhD and also for his genius STATA skills and willingness to teach me how to code efficiently.

A special thanks must go to all the researchers based at Manchester Centre for Health Economics, including Anne-Sophie Oxholm (honorary MCHE researcher), all of whom have always been there with me to celebrate my achievements as well as confide in when times got tough. Thank you for making my experience as a PhD student so enjoyable.

Finally, I would like to thank my family and friends beyond my academic studies. Their love, care and support have been crucial during the past three years. I'd also like to thank my late Grandmother, Ruth Jones (Nan), who was always there with her support, love and kindness.

"Not everything that can be counted counts and not everything that counts can be counted"

- *Albert Einstein*

Chapter 1 Introduction

1.1 Thesis Outline

The overall aim of this thesis was to identify and appraise current methods available to quantify the impact of presenteeism and, if appropriate, to develop methods suitable for use in economic evaluations of workplace interventions (WPIs). In this thesis, WPIs are defined as a set of activities, programmes, and equipment that can be used to manage conditions in terms of improving symptoms and functional ability (Public Health England 2014a). This thesis used a mixed methods approach involving both quantitative and qualitative methods. The primary focus was to explore the potential for quantifying the impact of presenteeism (hereafter shortened to presenteeism) using measures of health status and capability. This first chapter introduces: the definition of presenteeism and discusses the impact of presenteeism across diseases (section 1.2); WPIs and highlights issues of quantifying presenteeism; the selected case study for this thesis (section 1.4); the research questions (section 1.5); and the empirical approach (section 1.6). Section 1.7 outlines the structure of the thesis (section 1.7).

1.2 Definitions of Presenteeism

Productivity is defined as a ratio of inputs used per unit of output and provides a measure of technical efficiency that examines how inputs, such as labour and capital (technology), are used by firms to produce a volume of output (Sloman, Wride, and Garratt 2012). Productivity, related to a single employees or a workforce, describes the efficiency of a worker, or workforce, which may be measured by assessing the level of output produced by a worker, or workforce, over a specific time period (Smith 1994). Presenteeism is a concept that is related to productivity however there is currently no consensus regarding the definition of presenteeism (Johns 2010). A published review identified and summarised nine different definitions of presenteeism which are listed in Table 1.1 (Johns 2010).

Table 1.1 Definitions of presenteeism

Definitions of Presenteeism	Illustrative references
Attending work, as opposed to being absent	Smith (1970)
Exhibiting excellent attendance	Canfield and Soash (1955); Stolz (1993)
Working elevated hours, thus putting in ‘‘face time,’’ even when unfit	Simpson (1998); Worrall, Cooper, and Campbell (2000)
Being reluctant to work part time rather than full time	Sheridan (2004)
Being unhealthy but exhibiting no sickness absenteeism	Kivimäki et al. (2005)
Going to work despite feeling unhealthy	Aronsson, Gustafsson, and Dallner (2000); Dew, Keefe, and Small (2005)
Going to work despite feeling unhealthy or experiencing other events that might normally compel absence (e.g., child care problems)	Evans (2004); Johansson and Lundberg (2004)
Reduced productivity at work due to health problems	Turpin et al. (2004)
Reduced productivity at work due to health problems or other events that distract one from full productivity (e.g., office politics)	Hummer, Sherman, and Quinn (2002); Whitehouse (2005)

Source: *Johns (2010)*

Johns (2010) stated that the definitions, depicted in Table 1.1, share a common theme where the individual is present at work; beyond this, however, the definitions were varied and some confusing. Johns (2010) argued that the most appropriate definition of presenteeism should be ‘attending work whilst ill’. This definition goes part of the way toward explaining the concept of presenteeism but it does not say anything about lost productivity as a result of ill-health.

The definition provided by Turpin et al., (2004) refers to lost productivity at work because of problems experienced with ill-health. In an earlier review of self-reported presenteeism measurement instruments presenteeism was similarly characterised as the health-related productivity lost whilst at paid work (Loeppke et al. 2003). Loeppke et al., explained that this definition of presenteeism included several concepts: 1) time not on task (e.g. in the workplace, but not working); 2) decreased quality of work (e.g. increased injury rates, product waste, product defects); 3) decreased quantity of work; 4) unsatisfactory employee interpersonal factors (e.g. personality disorders); and 5)

unsatisfactory work culture. The definition offered by Loeppke et al., is thorough and has subsequently been adopted by the Outcomes Measures in Rheumatology group (OMERACT) (Tugwell et al. 2007) worker productivity group (Escorpizo et al. 2007). OMERACT is a research body that is involved in developing and validating clinical and imaging outcome measures in rheumatic diseases for use in clinical studies. For the purpose of this thesis, to maintain consistency with the OMERACT group, the adopted definition of presenteeism was the 'reduction in working performance whilst at work due to ill-health'.

Throughout this thesis productivity describes a worker's efficiency in terms of the output they produce. Presenteeism specifically refers to the productivity lost caused by ill-health in an employee attending work.

1.2.1 Presenteeism

The impact of poor health has been directly linked to lost productivity through absenteeism and presenteeism. Absenteeism and presenteeism, two related but distinct concepts, are terms used to describe how ill-health negatively impacts productivity. Absenteeism refers to the amount of time (measured in hours, days, weeks) spent away from work because of ill-health (Mattke et al. 2007; Krol and Brouwer 2014). There is a growing body of evidence that a number of health conditions, mental and physical, can increase levels of presenteeism (Schultz and Edington 2007). There is also evidence to suggest that presenteeism accounts for a greater proportion of the total cost of lost productivity compared with absenteeism (Schultz, Chen, and Edington 2009). Such evidence has caught the attention of academic researchers, employers, and healthcare policy makers (Johns 2010). To illustrate the magnitude of the cost of presenteeism, compared to the cost of absenteeism, a report written in 2013 quantified the impact of musculoskeletal disorders (MSDs) across six European countries (Belgium, Bulgaria, Ireland, Portugal, Spain, and the United Kingdom) and found absenteeism accounted for 12% of lost productivity and presenteeism accounted for 40% per week (Zheltoukhova 2013).

Interest in understanding and quantifying the impact of presenteeism has grown since the early 2000s, however it is still a relatively new concept when compared with absenteeism. The literature has produced an extensive body of theoretical debates and empirical evidence concerning absenteeism, but far less has been done to explore presenteeism in terms of its identification, measurement, and valuation (Heuvel et al.

2009). As a consequence, current methods of valuing the impact of presenteeism are still in their infancy (Goetzel et al. 2004). The evidence supporting the appropriateness of the methods that have been used to value presenteeism is also limited. There is a paucity of evidence to support the selection of a preferred measure for identifying and measuring presenteeism and there exists substantial uncertainty regarding what methods should be used to value presenteeism in monetary terms (chapter two, section 2.3.5). At present, there is no agreed standardised technique that can be used to quantify the impact of presenteeism (Soeren Mattke et al. 2007; Brooks et al. 2010).

The lack of a standardised method to quantify the impact of presenteeism has resulted in published estimates that vary extensively (Schultz, Chen, and Edington 2009). To illustrate this point, a systematic review by Schultz, Chen, and Edington (2009) highlighted the substantial differences between estimates of presenteeism for different health conditions. The aim of the systematic review was two-fold: 1) to identify the proportion of the total cost of healthcare attributable to presenteeism per health condition; and 2) to provide employers with an average cost of presenteeism per health condition. The total costs of each health condition included those associated with medical technologies and pharmaceuticals, absenteeism and presenteeism. The review identified four studies that incorporated all four costs (medical, pharmaceutical, absenteeism and presenteeism) to estimate the total cost per condition. Table 1.2, taken from Schultz, Chen, and Edington (2009), presents presenteeism as a percentage of the total cost of each health condition estimated by the four identified studies.

Table 1.2 Cost of presenteeism as a percentage of the total cost per health condition

	Goetzel et al. (2004) 'Average' estimate (%)	Goetzel et al. (2004) 'Low' estimate (%)	Collins et al. (2005) (%)	Loeppke et al. (2007) (%)	Average of available estimates (%)
Allergies	82%	55	74		70
Osteoarthritis	77	35	67	44	56
Asthma	73	35	72		60
Back/neck disorders			69	50	60
Breathing disorders			56		56
Cancer	53	6			30
Depression/Anxiety	71	27	81	70	62
Diabetes mellitus	62	16	56		45
Fatigue				73	73
Heart/circulatory	19	0	66		28
High cholesterol				43	43
Hypertension	63	9		33	35
Migraine	89	49	72		70
Obesity				56	56
Other chronic pain				33	33
Respiratory infection	25	3			14
Sleeping problems				66	66
Stomach/bowel			67		67

Source: Schultz, Chen, and Edington (2009)

From these data, it is clear that estimates of presenteeism as a percentage of the total cost of each health condition vary extensively across studies within the same disease area. The authors suggested that the various methods used to estimate the cost impact of presenteeism accounted for some of the variation between estimates. The authors also acknowledged that even if studies had used the same methods to identify and measure presenteeism other aspects of the studies, such as the population or the perspective adopted by the study (societal or employer) (see chapter two, section 2.2.5) would also have caused estimates to vary across studies of presenteeism.

1.3 Workplace Interventions

In the UK, work disability is defined by the Equality Act 2010 as a 'physical or mental impairment which has substantial and long-term negative effect on your ability to do normal daily activities. Inflammatory arthritis (IA) (see section 1.4) is a long-term chronic condition that is considered to be a disability by the Equality Act 2010. IA is understood to have a negative impact on individuals' ability to work and increases the likelihood of job loss or early retirement (Lenssinck et al. 2013).

Workplace Interventions (WPIs) are defined as a set of activities, programmes, and equipment that can be used to help employees manage their health condition by improving symptoms and functional ability at the workplace (Public Health England 2014a). Table 1.3 lists some examples of the different types of WPIs that are available and how they aim to improve the health and well-being of employees with MSDs. WPI-MSDs can be directed towards the individual patient, the patient's workplace, healthcare or other services for which the patient has access (Palmer et al. 2012). The definition of WPIs does not include those approaches employers offer to employees who declare themselves as 'disabled'. Interventions for those with a declared disability are classed as reasonable adjustments that employers are obligated to offer by law (Equality Act 2010).

Table 1.3 Types of workplace interventions

Intervention level	Workplace Interventions	Description
Individual	Exercise or physical therapy	To increase flexibility and strength of the muscles
	Psychological therapy	To change the behaviour and attitude of the patient. To help patients cope with their condition whilst at work
	Vocationally focussed therapy with education	Help patients overcome social barriers including attitudes and perceptions of work
	Rehearsal of safer working techniques	To avoid unnecessary strain on the body
Workplace	Ergonomic or psychosocial risk assessments aimed at the individual	To identify and control risks within the workplace
	Ergonomic changes to the physical environment	Equipment changes, such as computer displays
		Environmental changes, such as temperature and lighting
	Job modifications	Flexible working hours and lighter duties to help prevent further damage to the body
Healthcare Service	Assessment and coordinated action plan	Developed by a multidisciplinary team or case manager
	Consultation with occupational physician	-
	Education of primary-care doctors and/or occupational physicians and/or agreements between them	To improve communication between work and primary care and/or occupational physicians
	Access to external support and referral services	Further support available outside of the workplace

Source: Adapted from Palmer et al. (2012) and HSE. (2013)

Employers have recognised the potential substantial costs associated with presenteeism and realise there are gains towards improving productivity by investing in their workforce's health (Hemp 2004). Employers frequently fund WPIs to improve productivity but have limited access to reliable information about the potential productive benefits such interventions may produce (Uegaki et al. 2011). This thesis focuses on WPIs, funded by employers, as an exemplar of when it is important to take

account of the impact on presenteeism in economic evaluations (see chapter two, section 2.1) that aim to quantify the costs and benefits of such interventions; more information is provided about WPIs in chapter two .

1.4 The Case Study for this Thesis

This thesis uses inflammatory arthritic conditions (IA) as the case study to provide a focus for the empirical work. Three examples are used: rheumatoid arthritis (RA); ankylosing spondylitis (AS), and psoriatic arthritis (PsA). This section describes the selected case study and why it was chosen as a good exemplar to test and develop methods suitable to estimate the impact ill-health has on presenteeism.

In the United Kingdom (UK), the three most common forms of adult inflammatory arthritis are RA, AS, and PsA. All three conditions are autoimmune diseases causing pain, fatigue, and stiffness of the joints; if untreated, the conditions may permanently damage the joints leaving the patient permanently disabled (Scott and Kingsley 2007). The onset of each disease may occur when an individual is of working age (below the age of 65 years). Individuals who develop an inflammatory disease are at an increased risk of presenteeism, absenteeism, and becoming permanently work disabled (Lenssinck et al. 2013). All three conditions are episodic and fluctuate overtime meaning individuals are likely to experience different levels of presenteeism over time (Verstappen 2015). RA, AS, and PsA each have distinctive characteristics and affect people in different ways. The following section explains the epidemiology of each disease and the impact they have on a patient's ability to work.

1.4.1 Rheumatoid Arthritis

RA primarily attacks the lining in synovial joints (most commonly the joints in the hands) causing them to become inflamed (Scott, Wolfe, and Huizinga 2010). Synovial inflammation results in hyperplasia (swelling) of the joints and may result in bone erosion and degradation of cartilage (McInnes and Schett 2011). RA is also associated with an increased likelihood of developing co-morbidities, such as cardiovascular conditions (Symmons and Gabriel 2011)

Epidemiology

RA affects approximately 1% of the adult population in the UK (Symmons et al. 2002). RA affects sub-populations depending on their ethnicity, gender, and age. A re-estimation of UK incidence rates of RA by gender, using the 2010 American College of

Rheumatology (ACR)/European League Against Rheumatism (EULAR) criteria, following on from the previously applied 1987 ACR classification criteria, were 54/100,000 for females and 25/100,000 for males (Humphreys et al. 2013). The peak age of incidence for women was found to be between the ages of 45 and 74; typically ages where women are still working (Humphreys et al. 2013). Conversely, the likelihood of men developing RA was found to increase with age with highest rates in men at over 65 years. At present, the cause of RA remains largely unknown, but genetics and environmental factors, such as smoking, may collectively play a role (Klareskog, Catrina and Paget 2009).

Work Status

Evidence suggests that 20% to 70% of patients with RA become work disabled within 4.5 and 22 years after the initial onset of symptoms (Verstappen 2015). Previous systematic reviews have reported factors that determine work disability and unemployment in patients with RA (Verstappen et al. 2004; de Croon et al. 2004). The factors include physically demanding jobs, disease duration, severity of disease, levels of pain, and psychological factors.

1.4.2 Ankylosing Spondylitis

AS is a chronic inflammatory rheumatic disease causing inflammation severe stiffness and rigidity in the sacroiliac joints (pelvic joints), the spine, and other larger peripheral joints (neck and ribs) causing (Ebringer 2013). Common features of AS include pain and stiffness of the back and pelvic joints as well as fatigue and sickness (Ebringer 2013). In some severe cases, AS can cause individual vertebrae of the spine to fuse together limiting the patient's ability to move freely (Sieper et al. 2011). Also, up to 40% of individuals with AS develop eye conditions the most common being uveitis (iritis) (Inman 2012).

Epidemiology

A study conducted in 2014 investigated the global prevalence of AS (Dean et al. 2014). The study did not specifically look at the prevalence of AS in the UK, but did estimate the prevalence of AS across Europe and suggested it was 23.8 per 10,000 (0.02%). It has also been estimated that 80% of individuals with AS develop the condition before the age of 30 years (Sieper and Braun 2011). Evidence suggests that males are more likely to develop the condition compared to females at a ratio of 2.5:1 to 5:1 (Inman 2012). The specific cause of AS is still largely unknown; however, it is understood that genetics may play a crucial role (Inman 2012).

Work Status

Individuals with AS are three times more likely compared to the general population to withdraw from work increasing from 5% during the first year of diagnosis to over 20% at 10 years and 30% at 20 years (Martindale, Shukla, and Goodacre 2015). There is limited evidence of the key determinants of work productivity and AS; however physical factors, such as pain, stiffness, and fatigue, and psychosocial factors, including anxiety, depression and lower self-esteem, have been shown to be associated with the loss of employment (Martindale, Shukla, and Goodacre 2015).

1.4.3 Psoriatic Arthritis

PsA is a chronic inflammatory spondyloarthritis which affects the joints of the body (Duarte, Faillace, and Freire de Carvalho 2012). Psoriasis, a condition that affects the skin, is often present with the condition (Duarte, Faillace, and Freire de Carvalho 2012). Patients with PsA experience symptoms of inflammation of the sacroiliac joints (sacroilitis), swelling of the fingers and toes (dactylitis), inflammation of the tendons or ligaments (enthsitis), skin lesions (psoriasis), and alterations to the nails including pitting, thickening and discolouration. Some patients may even experience inflammation of the eye (iritis), urethra (urethritis), and bowel (Mease and Helliwell 2008; Duarte, Faillace, and Freire de Carvalho 2012).

Epidemiology

The prevalence of PsA has been estimated to affect approximately 0.19% of the UK's adult population (Ogdie et al. 2013). Disease onset occurs in patients between the ages of 20 years and 50 years and has been found to affect males and females equally (Mease and Helliwell 2008). Currently, the cause of PsA remains unexplained however, it has

been suggested that genetic and environmental factors, infections, drugs, and joint trauma may contribute (Duarte, Faillace, and Freire de Carvalho 2012).

Work Status

A systematic review conducted in 2012 aimed to establish the relationship between levels of work disability and PsA (Tillett, de-Vries, and McHugh 2012). The authors of the review found that work disability levels were high, ranging from 16% to 39%, and associated with longer disease duration, worse physical function, high joint count, low educational level, being female, erosive disease, and manual labour.

1.4.4 Treatment

At present, there is no known cure for RA, AS, or PsA. The supportive management of patients with these conditions is achieved through combinations of diverse therapies including pharmacological treatment, physical treatment, education, and counselling. Pharmacological treatment aims to relieve symptoms of RA, AS, and PsA, by inhibiting the progression of the disease. Such treatments aim to prevent further damage to the joints and enhance a patient's health related quality of life (HRQoL) through reduced pain, fatigue and functional disability. Treatment of RA, AS, and PsA is most effective when started early after disease on-set (NICE 2009).

In the UK, a combination of synthetic disease modifying anti-rheumatic drugs (csDMARDs), including methotrexate, with at least one other csDMARD or oral steroids is offered as a first line treatment (NICE 2009; Smolen et al. 2013). Although csDMARDs can be a very effective, not all patients respond adequately. Patients that do not respond to csDMARDs may be offered biologic agents as a second line treatment option. Biologic agents are protein-based treatments that are produced using elements from a living organism. The purpose of biologic treatment is to suppress key components of the immune system to inhibit the production of cells that lead to increased inflammation (Keyser 2011). Biologics, whilst proven to be an effective form of treatment (Keyser 2011), come at a substantial incremental cost compared to csDMARDs; for example, the cost of adalimumab(biologic) in the UK is £9,295 per patient, per year (NICE 2010) whereas the cost of methotrexate (csDMARD) is approximately £400 per patient per year (NICE 2017). Biologics, therefore, remain a second-line treatment strategy in most countries (Finckh 2009; Her and Kavanaugh 2012). In the United Kingdom, biologics are only offered if the patient has: 1) failed to

respond adequately to csDMARDs; and 2) fulfils the requirements of the disease severity assessment criteria (NICE 2012a).

1.4.5 Economic Impact of RA, AS and PsA

Due to extensive heterogeneity of the methods used to quantify the impact of presenteeism (see chapter two, section 2.3) it is not possible to state the exact economic impact caused by RA, AS and PSA. However, a selection of studies and systematic reviews are presented to illustrate the extent to which RA, AS, and PsA impact of productivity in terms of presenteeism and absenteeism.

Rheumatoid Arthritis

In 2012, a study was carried out to estimate the impact of absenteeism and presenteeism in patients with RA (Braakman-Jansen et al. 2012). The study used three self-reported measures including the: 1) Productivity and Disease Questionnaire (PRODISQ)(absenteeism measure) (Koopmanschap 2005) and Quantity and Quality method (QQ method) (presenteeism measure) (Brouwer, Koopmanschap, and Rutten 1999); 2) Work Productivity Activity Impairment General Health (WPAI:GH) (Reilly, Zbrozek, and Dukes 2012); and 3) Health and Labour Questionnaire (HLQ) (Leona Van Roijen et al. 1996) (see chapter three, section 3.4.3). Absenteeism was measured by the PRODISQ and WPAI-GH as the number of days absent from work within the last 3 months and seven days, respectively. Presenteeism was measured by the QQ method, WPAI-GH and HLQ as the level of reduced work performance on the last day of work, the last seven days, and the last two weeks, respectively. The study compared productivity levels of a group of RA patients, diagnosed within three months, and a control group consisting of individuals without RA matched by aged and gender. Each group was required to complete the three selected measures for presenteeism and absenteeism.

Presenteeism was reported as the percentage reduction in productivity per patient, per week. Patients with RA reported percentage level reductions in productvitiy of 60% (measured using the QQ method), 79% (WPAI:GH) and 67% (HLQ). The results were statistically significantly different compared with those reported by the control group whose reductions in productivity were 25% (QQ method), 41% (WPAI:GH) and 23% (HLQ). The percentage reductions in productivity were converted into a cost using the human capital approach (HCA) where the percentage reduction in productivity per time period is multiplied by a wage rate which was set at €37.69 (see chapter two, section

2.3.5). Table 1.4 presents estimates of the cost impact of presenteeism and absenteeism caused by RA, per patient, per week. The total cost refers to the cost of productivity, taking into account the cost of absenteeism and presenteeism.

Table 1.4 Percentage cost of lost productivity

Costs (mean, €)	**RA, cost per patient per week**	**Percentage of total cost**	**Control, cost per patient per week**	**Percentage of total cost**
PRODISQ				
Absenteeism	€120	29%	€9	5%
Presenteeism (QQ)	€299	71%	€154	95%
WPAI:GH				
Absenteeism	€116	27%	€6	8%
Presenteeism	€318	73%	€72	92%
HLQ				
Presenteeism	€79	NA	€11	NA

Source: Braakman-Jansen et al. (2012). RA: Rheumatoid Arthritis; WPAI: GH: Work Productivity Activity Impairment Questionnaire; QQ Method: Quantity and Quality Method; HLQ: Health and Labour Questionnaire

The total cost of absenteeism and presenteeism as a result of working with RA was greater than the cost of absenteeism and presenteeism in the control group. The proportion of the total cost of presenteeism compared with absenteeism was substantially greater in both the RA and control groups. Extensive variation of the cost estimates associated with each self-reported method a finding consistent previously studies (Zhang et al. 2010).

Ankylosing Spondylitis

A recent study conducted in 2015 estimated the cost of AS in the UK using a linked routine and patient reported survey data (Cooksey et al. 2015). The study used a sample of 228 working patients with AS. The Work Limitations Questionnaire (WLQ) (Lerner et al. 2001) and the WPAI were used to measure lost productivity associated with absenteeism and presenteeism. Working patients with AS absent for 3.5% of their work time and reported a 21.6% increase in their levels of presenteeism. Using the HCA and applying average daily wages rates collected from the Office of National Statistics (ONS), the average cost of absenteeism and presenteeism was estimated at £869 and £7,241 per year, per working patient, respectively. The results of this study were in line with a previous UK study which found the majority of the total costs of AS were attributable to absenteeism, unemployment, and presenteeism (Rafia et al. 2012).

Psoriatic Arthritis

A systematic review was conducted to determine the clinical, economic, and humanistic burden of illness (Lee, Mendelsohn, and Sarnes 2010). The review identified seven studies that evaluated the impact of lost productivity (absenteeism only). The proportion of the total cost of PsA caused by absenteeism varied between 52% and 72%. Whilst the review was firm in its conclusion that lost productivity caused by PsA was substantial, the lack of data meant that an accurate estimation the total cost of absenteeism caused by PsA could not be provided.

In summary, whilst there was substantial variation in the estimated costs of presenteeism across studies, evidence consistently shows presenteeism caused by RA, AS or PsA is responsible for a significant proportion of the total cost of lost productivity.

1.5 Research Questions

The overall aim of this thesis was formulated into two broad objectives: 1) to identify and appraise current methods used to identify, measure, and value presenteeism; and 2) to develop methods that predict presenteeism. Five research questions were established to address the overall aim of this thesis.

1. What are the current issues of presenteeism in the context of economic evaluation of WPIs for RA, AS and PsA conducted from the context of NHS England?
2. To what extent are self-reported measures of presenteeism, suitable for RA, AS or PsA, underpinned by economic theory?
3. If, and how, is the impact of productivity included in economic evaluations of WPIs designed for MSD[1]?
4. Are existing measures of health-related quality of life (HRQoL), such as the EuroQol Five Dimension (EQ5D) (EuroQol 1990) able to capture the impact of presenteeism caused by RA, AS and PsA?
5. Whether existing measures of HRQoL can be used to predict presenteeism caused by RA or AS?

[1] MSDs were the focus of research objective three because of the limited number of economic evaluations of WPIs suitable for RA, AS or PsA.

1.6 Overview of Methods

Three distinct reviews were used to address objectives one to three. Objective one was addressed using a narrative review of the current issues of presenteeism and methods to quantify it in the context of economic evaluations. Objectives two and three used systematic review methods to: 1) assess whether self-reported measures of presenteeism are underpinned by economic theory; and 2) explore if, and how, presenteeism is included in empirical economic evaluations. As a collection, the three reviews provided detail about the theoretical and empirical state of the methods used to quantify presenteeism and also informed the methods used in the two subsequent empirical studies of this thesis. The first empirical study used qualitative methods to explore how RA, AS and PsA affect employees ability to work and whether HRQoL measures were able to capture the impact on presenteeism. The second empirical study applied econometric regression methods to quantify the relationship between HRQoL and presenteeism of employees with RA or AS. The two empirical chapters provide complimentary evidence to support the extent to which a relationship between HRQoL and presenteeism exists.

1.6.1 Qualitative Methods

Health economics typically applies the quantitative methods used in mainstream economics; however, the use of qualitative methods to aid understanding of a research area is increasingly being used in health economics (Coast 1999). This thesis employed qualitative methods to understand whether generic measures of health status, such as the EQ5D (EuroQol 1990), have the ability to capture those aspects of RA, AS, and PsA that increase levels of presenteeism. Semi-structured telephone interviews were conducted and the data was analysed using thematic analysis methods (see chapter five, section 5.3).

1.6.2 Quantitative Methods

Econometric methods are frequently used in health economics to estimate economic relationships, test economic theories, and evaluate government policies. This thesis applied econometric methods to develop a prediction model for presenteeism using HRQoL measures and self-report measures of presenteeism. The quantitative chapter presents the best performing prediction model for presenteeism selected via the assessment of multiple measures of error that describe how accurate the model predicts presenteeism.

1.7 Thesis Outline

This chapter has introduced the key concept of presenteeism, the case study, and the methods used to address the five objectives of this thesis. Table 1.4 outlines the structure of the thesis, the topics and the methods used in each chapter.

Chapter two presents a narrative review which describes relevant concepts, theories and methods used to quantify the impact of productivity. The chapter also describes the current limitations of the methods used to identify, measure, and value presenteeism.

Chapter three presents a systematic review that focussed on investigating the theoretical underpinnings of self-reported instruments used to identify and measure presenteeism. All published systematic reviews of the methods of presenteeism were identified and a summary of the current methods used to quantify presenteeism was provided. Subject to pre-specified criteria, one of the identified systematic reviews was selected and updated to identify all published self-report instruments of presenteeism used to identify and measure presenteeism in the context of RA, AS and PsA.

Chapter four presents a systematic review that focused on investigating if, and how, presenteeism has been incorporated in economic evaluations of WPIs. This chapter presents the extent to which absenteeism and presenteeism are included in economic evaluations of WPIs for individuals with MSDs and reports the methods used to identify, measure, and value productivity.

Chapter five presents a qualitative study to identify how RA, AS and PsA affect an employee's ability to work and whether measures of health status and capability (hereafter termed 'HRQoL') are able to capture the impact on presenteeism. Semi-structured telephone interviews were conducted with a sample of employed individuals with RA, AS or PsA. The study gained an in-depth understanding of how each condition affected an individual's ability to work effectively. The results of this chapter provided face validity for the existing measures of HRQoL and their ability to capture presenteeism caused by RA, AS or PsA.

Chapter six presents the development of several prediction models that use HRQoL measures to predict presenteeism caused by RA or AS. The chapter also reports the performance of each prediction model by assessing multiple specification tests to measure error. The results from the specification tests informed the selection of the best

performing prediction model for presenteeism for a population with RA or AS. The results were generated from a sample of employed individuals working with RA or AS who completed a bespoke online survey designed to collect HRQoL and presenteeism data.

Chapter seven collates the results from each study and discusses the implications in terms of quantifying presenteeism suitable for use in economic evaluations of WPIs in the context of the UK.

Table 1.5 Structure of thesis and key methods

Chapter	Aim	Methods	Title
Two	To identify and describe the current challenges of quantifying the impact of presenteeism	Narrative review	What are the current issues of presenteeism in the context of health economic evaluation?
Three	To assess the extent to which self-reported methods used to identify and measure presenteeism are based on economic theory	Systematic review	Economic theory and self-reported measures of presenteeism in musculoskeletal disorders
Four	To identify if, and how, economic evaluations of WPIs that manage MSDs identify, measure, value, and incorporate the impact on productivity	Systematic review	If, and how, is the impact of productivity included in economic evaluations of workplace interventions for musculoskeletal disorders?
Five	To explore the extent to which selected measures of health status and capability are able to capture the impact RA, AS, and PsA has on presenteeism	Qualitative interviews	Exploring the impact of health status and capability of people with inflammatory arthritis on presenteeism in the workplace: A qualitative study
Six	To predict presenteeism, caused by RA and AS, using health status and capability measures, suitable for use in economic evaluations of WPIs	Econometric (mapping) methods	To develop a prediction model for presenteeism using existing measures of health status and capability
Seven	To discuss the results and implications for future research	Assimilation of findings from reviews and empirical studies	Discussion and Conclusion

RA: Rheumatoid Arthritis: AS; Ankylosing Spondylitis; WPIs: Workplace Interventions; MSDs: Musculoskeletal Disorders.

Chapter 2 What are the current issues of presenteeism in the context of health economic evaluation?

2.0 Introduction

The aim of chapter two was to summarise the issues related to productivity in economic evaluations conducted within the context of National Health Service England (NHS - England). The two objectives of this chapter were to describe if, and how, to incorporate productivity in economic evaluations of healthcare interventions and to discuss the main challenges of existing methods used to identify, measure, and value presenteeism. To address these objectives a narrative review was conducted. This review was not a systematic review, as defined by Centre for Reviews and Dissemination (CRD 2009), but does provide a summary overview of the relevant supporting literature.

The chapter is presented in four sections. Section 2.1 defines economic evaluation and describes the methods that can be applied. Section 2.2 describes the normative and positive issues of productivity and how it relates to WPIs. Section 2.3 describes the current methods used to identify, measure, and value presenteeism and section 2.4 presents the conclusion.

2.1 What is Economic Evaluation?

Economic evaluation has been defined as the "comparative analysis of alternative courses of action in terms of both their costs and consequences" (Drummond et al. 2015). Economic evaluations provide decision-makers with evidence to inform the best allocation of scare healthcare resources across a healthcare service (Sanders et al. 2016). Typically, economic evaluations are most useful when implemented in the context of non-marketable goods. Healthcare is an example of a market that provides non-marketable goods as it does not generally provide explicit information about the cost of products that would ordinarily be generated by market forces (Morris et al. 2012). In a resource constrained world economic evaluations are essential to help avoid making inefficient and sub-optimal decisions which may have potentially large opportunity costs. Opportunity cost describes the benefits forgone that could have been realised if an investment in alternative product, intervention or policy had been made (Palmer and Raftery 1999).

2.1.1 Overview of the role of Economic Evaluation

The decision making process for the effective use of healthcare and how to spend the healthcare budget is complex. In the England, the budget for all public spending is set annually by the Chancellor of the Exchequer (government). A portion of the budget is earmarked for healthcare spending and in 2015 total healthcare expenditure, in England, was £185 billion, accounting for 9.9% of gross domestic product (GDP). Curative and rehabilitative care receive 63.7% of the £185 billion, long-term care, including residential, nursing and domiciliary home care, and palliative service, accounted for 15.3% of government expenditure on healthcare, and preventative care, including public health services, screening programmes, and health check-ups accounted for 5.1% of the total spend on healthcare (ONS, 2015). The money directed to curative and rehabilitative care goes to commissioning groups including clinical commissioning groups (CCGs) and NHS England (King's Fund 2014). CCGs and NHS England decide how to spend the budget across both public and private health services with the goal of maximising health. Investment decisions made by CCGs and NHS England are informed using evidence provided by specialist support units and organisations. One such organisation that assists in the making of these decisions is the National Institute for Health and Care Excellence (NICE). NICE provides evidence based guidelines to support decision making for commissioners, practitioners, and managers across the healthcare sector. NICE collates information using a systematic process of Health Technology Appraisal (HTA). HTAs assess both the clinical and cost-effective evidence of medical technologies to understand if, and how well, a treatment works and whether it is considered to be cost-effective (NICE 2014). The evidence from HTAs are assessed by a committee using a deliberative process. NICE has also published a set of guidelines that inform healthcare researchers of how to conduct economic evaluations in the context of the NHS (NICE 2014). At present, NICE recommend adopting a health and social services perspective when conducting economic evaluations within the context of NHS England (see section 2.2.5). The choice of perspective is based on the premise that all spending comes from the set healthcare budget. In practice, only those health-related costs and benefits that fall on the healthcare sector are considered relevant when assessing the cost-effectiveness of healthcare interventions; productivity is a factor that falls beyond the remit of the healthcare sector and is therefore excluded from economic evaluations of healthcare interventions.

The adoption of a narrow healthcare sector perspective is still a subject of debate and the limitations associated have been highlighted in the context of assessing complex interventions such as public health interventions. Since 2006, NICE broadened its remit to include the evaluation of public health interventions (NICE 2006). Public health interventions are those that are targeted towards helping people maintain and improve their own level of health and well-being (WHO Europe 2017). Anti-smoking campaigns and the 'eat 5 (fruit and vegetables) a day' are two examples of public health interventions. NICE (2014) also state that they want to record benefits of an intervention that have an impact on other people, principally family and informal carers. The statement signals NICE position to move towards widening the perspective of economic evaluations of healthcare interventions

In March 2013, the responsibility of public health in England was devolved to local governments (NICE 2012b). Local governments are now responsible for their community's overall welfare meaning that public health has to compete with other public sectors, such as education, to secure funding. Public health interventions are known to have a positive impact on outcomes beyond health; for example, adults who eat a healthier diet, as a result from the 'eat 5 a day' campaign, may have reduced sick-leave. In these circumstances, NICE recommends the adoption of a broader perspective, namely the public sector, including the NHS and personal social services (PSS), or the local government, to reflect the context in which public health sits and to allow for the evaluation of outcomes beyond health (NICE 2012b). The original EQ5D-3L assigned want to record benefits of an intervention that have an impact on other people, principally family and informal carers. The statement signals NICE's position to move towards widening the perspective of economic evaluations of healthcare interventions.

Weatherly, Faria, and ven den Berg (2017) argued why it is important to take into account the impact of informal care, provided predominantly by unpaid family members, on investment decisions of healthcare interventions. The issues regarding the measurement and valuation of informal care and incorporating its impact into economic evaluations are similar to those issues regarding the quantification of presenteeism in economic evaluations. Weatherly, Faria, and ven den Berg (2017) state that informal care givers provide a service that may cause a shift in resources use away from public services and towards informal care givers. It may be useful to include contributions made by caregivers when assessing interventions that are aimed at the care recipient but have an impact on the caregiver in order to assess the cost-effectiveness of targeting

services with different levels of access. Outcomes that are derived from one sector (namely the informal care sector) and benefit another (public health services) are known as 'spill-over effects'. Spill-over effects can be directly applied to WPIs that are funded by employers. Employers pay for the WPI which is expected to increase productivity and prevent conditions from being made worse whilst working. It may also be useful to allow a consideration for the contributions made by employers to improve health and assess the cost-effectiveness of targeting services in that context. The arguments put forward by Weatherly and colleagues motivate the need to adopt an extra-welfarist perspective.

WPIs are not a healthcare intervention per se; however, they do aim to improve health outcomes of employees at the workplace. WPIs are frequently funded by the employer therefore the relevant perspective of an economic evaluation of WPIs is arguably the employer. There exists no national decision making body, such as NICE for the healthcare sector, that decides whether WPIs should be invested in. Individual employers must, therefore, make the decision to invest in WPIs. Economic evaluations of WPIs may be useful to inform employers of the costs and benefits of WPIs. Productivity is a key factor that is important to employers; however current methods are not adequate to inform employers of the potential productive benefits of WPIs. Hence, there is a need to develop robust methods that adequately quantify the impact of presenteeism.

In the UK, employers are legally bound to make sure their employees work in an environment that protects their 'health, safety and welfare (HSW Act. 1974). WPIs may be employed to promote health and safety at work by reducing the risk of injury, for example heavy lifting without adequate equipment. Such WPIs prevent ill-health from occurring or being exacerbated at the workplace which is also linked to presenteeism. Economic evaluations of WPIs that include presenteeism may provide information to employers regarding the need to install interventions that also promote safety and avoid excessive injury at the workplace.

In the United States (US), employers frequently fund their employees' health insurances and therefore have an intrinsic interest in their workforce's health. Methods that provide employers with an estimation of presenteeism could be used towards building a business case for investment in WPIs that manage health conditions and promote productivity. The US context is, whilst applicable, is not the focus for this thesis,

however a discussion about the methods developed for presenteeism will be presented to highlight the potential usefulness of methods in this context.

2.1.2 Types of Economic Evaluation

Economic evaluation broadly consists of three methods including cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost utility analysis (CUA). Drummond et al. (2015) suggest a fourth method, cost-minimisation analysis (CMA), in which costs alone are compared; however, this method is no longer recognised as a useful method (Briggs and O'Brien 2001; Dakin et al. 2014). Briggs and O'Brien argue that CMA is most often an inappropriate method because it is very difficult to understand given uncertainty that two interventions actually produce the same outcome. Briggs and O'Brien state that the costs and outcomes should be jointly estimated with rigours quantification of the uncertainty that surrounds the incremental cost-effectiveness ratio results. A fifth method, cost-consequence analysis (CCA), was introduced by Coast (2004). CCA presents decision-makers with a list of disaggregated consequences (outcomes or patient related benefits) costs. CCA, compared to CEA and CUA, is able to take into account consequences of an intervention beyond health status, for example, the impact on productivity. CCA is currently recommended by NICE to assess the cost-effectiveness of public health interventions because of its ability to help understand the costs and outcomes that have an impact beyond health (NICE 2012b). The method has been criticised for its complicated presentation of the costs and outcomes of a public health intervention (Brazier and McCabe 2017). Brazier and McCabe (2017) explain that multiple pieces of information generated by CCA cannot be easily used to inform a single investment decision. CCA has also been criticised for its inability to make an assessment of the degree of uncertainty associated with a decision (Claxton, Sculpher, and Ades 2004).

2.1.3 Cost-Benefit Analysis

Cost-benefit analysis (CBA) can be used to inform decision-makers about allocative efficiency; what activities are worth pursuing (Drummond et al. 2015). CBA is rooted in welfarist economics (see section 2.2) where outcomes are based on individual utility. CBA generally adopts a societal perspective (see section 2.2.5) and compares all of the relevant costs and benefits no matter who incurs them meaning that both health-related and non-health related outcomes are considered (Morris et al. 2012). The costs and benefits in a CBA are valued in monetary units which allow interventions across public sectors to be compared (Drummond et al. 2015).

Monetary costs and monetary benefits are converted into present values using a discount rate. Discount rates are applied under the assumption that individuals would prefer to gain health benefits now rather than later but offset costs. In England, the discount rate applied to both costs and benefits is 3.5% (HM Treasury, 2011). A direct comparison of the discounted incremental future costs and benefits are compared. The difference is known as the net social benefit (Drummond et al. 2015). The simplest decision rule applied to CBA is if the net social benefit is greater than the net social cost then, based on efficiency, the intervention should be funded (Feldstein 1964). The same decision can also be presented as a ratio of benefits to costs. If the estimated ratio is greater than one then the intervention or programme should be undertaken.

CBA: benefits valuation

Stated preference methods are a technique that are frequently used to value benefits, including health and non-health related, in CBA (Bayoumi 2004). Stated preference methods aim to value goods and services that typically have characteristics that enhance individuals' utility that are not directly valued in a marketplace. Methods used to value goods or services in the context of a hypothetical market are known as contingent valuation studies. Contingent valuation (CV) methods, such as willingness-to-pay (WTP) techniques, are used to reveal the maximum amount an individual is willing to pay for a good or service (Gafni 1997). An individual's WTP can be elicited either directly, by asking the respondent to state a value, or indirectly, by describing a hypothetical choice and observing their response. CV studies may also be used to assign monetary values to all types of benefits, non-health and health, that may arise from a healthcare intervention or programme.

CBA is generally not used as a method for economic evaluations conducted in England due to methodological and ethical challenges (Smith and Sach 2009). CV has been argued to have advantages when valuing health benefits suitable for economic evaluations; however, CV methods have failed to achieve their theoretical advantages (Smith and Sach 2009). For example, Olsen and Smith (2001) argue that CV may offer an improved set of methods to value outcomes because: 1) the method is rooted in welfare economics; 2) allows non-health benefits to be valued too; and 3) monetary valuation is needed to make investment decisions across sectors. First, Smith and Sach (2009) report that CV methods rooted in welfare economics is unimportant and should instead focus on being consistent with social value judgements. For example, NICE has

rejected the inclusion of CV methods indicating that society does not share the values of CV for valuing health benefits. Second, CV methods have rarely been used to value non-health benefits. Third, CV methods have largely been used to value one programme only, failing to take their potential advantage of informing investment decision across sectors. NHS England and NICE have opted to adopt an extra-welfarist point of view, and a restrictive one at that, which aims to maximise health gains subject to a fixed budget.

2.1.4 Cost-Utility Analysis and Cost-Effectiveness Analysis

NHS England and NICE have opted to adopt an extra-welfarist point of view (see section 2.2.2) with the aim of maximising health gains subject to a fixed budget. Cost-effectiveness analysis (CEA) and cost-utility analysis (CUA) are two methods that can be used within an extra-welfarist approach (section 2.2). CUA and CEA are:

(i) Theoretically underpinned by the normative framework of extra-welfarism;
(ii) Set-up to answer questions of technical efficiency (to produce the maximum output with minimum amount of inputs) to inform decision-makers of the maximum amount of health gains achievable for a given amount of resources (inputs) (Drummond et al. 2015);
(iii) Most useful when the decision-maker has to make choices about the allocation of resources within a fixed budget;
(iv) Designed to value benefits using non-monetary units;
(v) Not able to compare the costs and benefits between different public sectors.

CEA and CUA: benefits valuation

CEA and CUA are largely similar methods; however, they differ in the way benefits are valued. CEA measures the outcomes of an intervention or programme using natural units that are relevant to the intervention under analysis, for example life-years gained/saved, the number of deaths averted, and the number of cases detected. The disadvantage of CEA is that it is limited in its ability to compare the costs and benefits across different types of healthcare interventions aimed at deriving different outcomes (Luce and Simpson 1995). In order to make any kind of judgement about a healthcare intervention it is required that both interventions be valued using the same natural units. CEA is therefore limited in its ability to make trade-offs explicit between different outcomes for healthcare interventions that generate multiple outcomes (Birch and Gafni

1992). CUA provides an alternative approach that improves the comparability of outcomes across healthcare interventions.

CUA values all health-related outcomes of an intervention or policy using a measure called quality adjusted life-years (QALYs) (see chapter six, section 6.3.1). QALYs capture changes in both health-related quality (morbidity) and quantity (mortality) of life. The quantity of health is measured using a count of the number of life-years. The quality of health, a more challenging construct to value due to the lack of objective measures, is valued through the use of subjective patient reported outcome measures (PROMs) which may be generic or disease-specific. Generic PROMs, for example the European Quality of Life Five Dimension Five Level (EQ5D-5L) (EuroQol Group 1990a), attempt to capture the impact on health status as a whole, whereas disease-specific instruments, such as Rheumatoid Arthritis-specific quality of life instrument (RA-Qol) (de Jong et al. 1997) capture the impact of living a specific condition.

The original EQ5D-3L assigned three degrees of severity for each dimension including: no problems; some problems; and major problems. No problems are coded as one, some problems are coded as two, and major problems are coded as three; therefore, a patient who suffers no problems walking around (1), has no problems with self-care (1), has some problems performing usual activities (2), experiences severe pain (3), and is moderately anxious or depressed is assigned a health state 11232. In total there are 245 health states; 243 are defined by the EQ5D-3L with an additional two for ‘unconscious’ and ‘dead’.

The EQ5D-3L has a set of preference weights assigned to each level of severity within each domain. Each health state is weighted using an interval scale where one represents perfect health and zero represents death. Health state values below zero represent health states considered to be worse than death. The idea is that more preferable health states, for example 1111 versus 22212, are given greater weighting making them more favourable in an analysis. The weights used for EQ5D-3L valuations of health states were derived by Dolan and Gudex (1995) and Dolan et al. (1996) using the time trade-off (TTO) method. The TTO method, developed by Torrance, Thomas, and Sackett (1972), requires respondents to sacrifice time, in terms of length of life, in order to gain quality of life in terms of reduced morbidity. The set of preference weights for the United Kingdom (UK) were estimated from a random sample of 3,235 adults from the UK population (Dolan et al. 1996).

Evidence has suggested that the EQ5D-3L frequently exhibits large numbers of respondents reporting full-health (Longworth and Rowen 2013) because it is unable to distinguish between health states that are between 'no problems' and 'some problems'(Devlin and Krabbe 2013). The results of the EuroQol's efforts were to expand the EQ5D-3L to include two further levels of severity: 'slight problems' and 'severe problems'. The EQ5D-5L was developed to by the EuroQol Group taskforce in 2005 to improve sensitivity and reduce 'ceiling effects' (Herdman et al. 2011).

Further research to extend the QALY is also being carried out by a group of health economists based at The University of Sheffield (ScHARR 2017). The EQ5D-3L (or 5L) captures changes in health only. There is a growing concern among NICE and academic health economists that the current QALY does not capture important non-health related benefits. The project 'Extending the QALY' is focussed on developing a measure of HRQoL that captures wider non-health benefits, such as social and emotional wellbeing, that can be used to capture a measure of QoL across both the health and social care sectors.

Decision rules applied to CUA and CEA are more complex than those used in CBA. The difficulty arises when an intervention generates greater health benefits for a greater cost. A common technique used in CEA and CUA is to calculate the incremental cost effectiveness ratio (ICER) of two competing alternatives. The ICER represents the additional cost compared with the additional benefits generated by a healthcare intervention when compared with the next best alternative. The ICER is calculated as the change in cost divided by the change in effect; see equation 1 where C represents costs, E represents effectiveness, and Δ represent change:

$$ICER = \frac{\Delta C}{\Delta E} \qquad (1)$$

To assess whether or not the intervention represents an efficient use of resources the ICER is compared with a cost-effectiveness threshold. The cost-effectiveness threshold represents the opportunity cost of programmes displaced elsewhere within the healthcare system; however it is also often defined as the willingness-to-pay for an extra QALY (McCabe, Claxton, and Culyer 2008). If the ICER is greater than the cost-effectiveness threshold then the healthcare intervention should be rejected. However, in practice, decisions are often not as clear cut (Claxton et al. 2015). An example of where this type of decision making is in use currently is in NHS England. NICE has

implemented a cost-effectiveness threshold that ranges of £20,000 to £30,000 per QALY gained (NICE 2014).

2.1.5 Economic evaluation, productivity and the employer

The five types of economic evaluation discussed above quantify outcomes using different units. Some methods allow for absenteeism and presenteeism to be accounted for as an outcome which may be useful to the employer.

Table 2.1 lists the five types of economic evaluation, the typical measure of outcome used and potential metrics that can be used to quantify productivity. All types of economic evaluation may include productivity as a cost in the analyses depending on the perspective adopted (see section 2.2.5).

Table 2.1 Economic evaluation methods, productivity and the employer

Economic Evaluation method	**Defined outcome**	**Absenteeism Metric**	**Presenteeism Metric**
Cost Minimisation Analysis (CMA)	Outcomes are assumed to be equal	Cost time absent from work e.g. days, hours (£)	Cost time not doing work whilst at work e.g. days, hours (£)
Cost Effectiveness Analysis (CEA)	Lost productivity selected as the main outcome	Time absent from work e.g. days, hours	Time not working e.g. days, hours Work performance measured as a %
Cost Consequence Analysis (CCA)	Numerous outcomes selected	Time absent from work e.g. days, hours	Time not working e.g. days, hours Work performance measured as a %
Cost Utility Analysis (CUA)	Quality Adjusted Life Year (QALY)	Cost time absent from work e.g. days, hours (£)	Cost time not doing work whilst at work e.g. days, hours (£)
Cost Benefit Analysis (CBA)	All outcomes included and valued in monetary units	Monetary value of time absent from work e.g. days, hours (£)	Monetary value of time not working e.g. days, hours (£) Monetary value of work performance measured as a % (£)

As mentioned in section 2.1.2, CMA assumes all outcomes are equal; therefore, lost productivity could only be included as a cost in a CMA analysis. A CEA uses one

primary outcome to assess the performance of an intervention. CEA allows for the selection of non-health outcomes therefore productivity (absenteeism and/or presenteeism) may be selected. Absenteeism is typically measured as the amount of time absent from work, measured in days or hours, and presenteeism is typically measured as time spent not working, in days or hours, or work performance, measured as a percentage. Workplace Interventions (WPIs) are designed to improve an individual's ability to manage their condition at work which may result in an improvement in their ability to work therefore increasing productivity. A CEA may provide results that describe how the intervention impacts an individual's ability to work, however it cannot be used to assess the impact the intervention has on health. Cost consequence analysis (CCA) may be used to inform employers of the health and productivity benefits of an intervention by measuring each separately and assigning a cost. The disaggregated presentation of CCA results make it difficult for an employer to make an informed decision.

Cost Utility Analysis (CUA) measures outcomes using QALYs (see section 2.1.4). An economic evaluation, conducted from the perspective of the employer, may allow productivity to be included as a cost. However, including productivity as a cost, most frequently valued using wages, has the ability to influence cost-effective decisions in favour of those who earn the highest wages (Olsen and Richardson 1999) (see chapter 6, section 6.1).

CBA allows all for the inclusion of all costs and benefits, both health and non-health related. Productivity, a non-health related outcome, may easily be incorporated as a cost or benefit in a CBA. The disadvantage of CBA methods is relates to the valuation of outcomes and costs; all must be valued in monetary terms and this has implications when applied to productivity. There are two methods that may be used to value presenteeism in monetary units: 1) human capital approach; and 2) stated preferences.

Productivity is most often valued using the human capital approach (HCA) (see section 2.3.5) where the amount of presenteeism, measured in time units or as a percentage decrease in work performance, is multiplied by the wage rate. Pauly et al. (2002) and Koopmanschap et al. (1999) argue that the HCA over-estimates the value of presenteeism. Pauly et al. (2002 and 2008) suggest that there is a need to take into account the impact of compensation and multiplier effects (see section 2.3.4) when

estimating the impact on presenteeism. Stated preference methods, as mentioned in section 2.1.3, could be used to value employers WTP for a reduction in presenteeism. A potential difficulty with using WTP methods would be to disconnect employers' values of productivity away from wages being paid. WTP methods were considered as a potential way of valuing presenteeism; however it was discarded because of the strong likelihood that employers would think about productivity in terms of wages.

2.2 Understanding the Evaluative Framework

When discussing economic evaluations and productivity both normative and positive issues must be considered. Normative issues are those that address what ought to be done; should productivity be included in economic evaluations? Positive issues concern the practicalities of how to conduct an economic evaluation; how to include productivity in economic evaluations? The next section discusses the normative arguments from which value judgements emerge that underpin the inclusion or exclusion of productivity in economic evaluations of healthcare interventions.

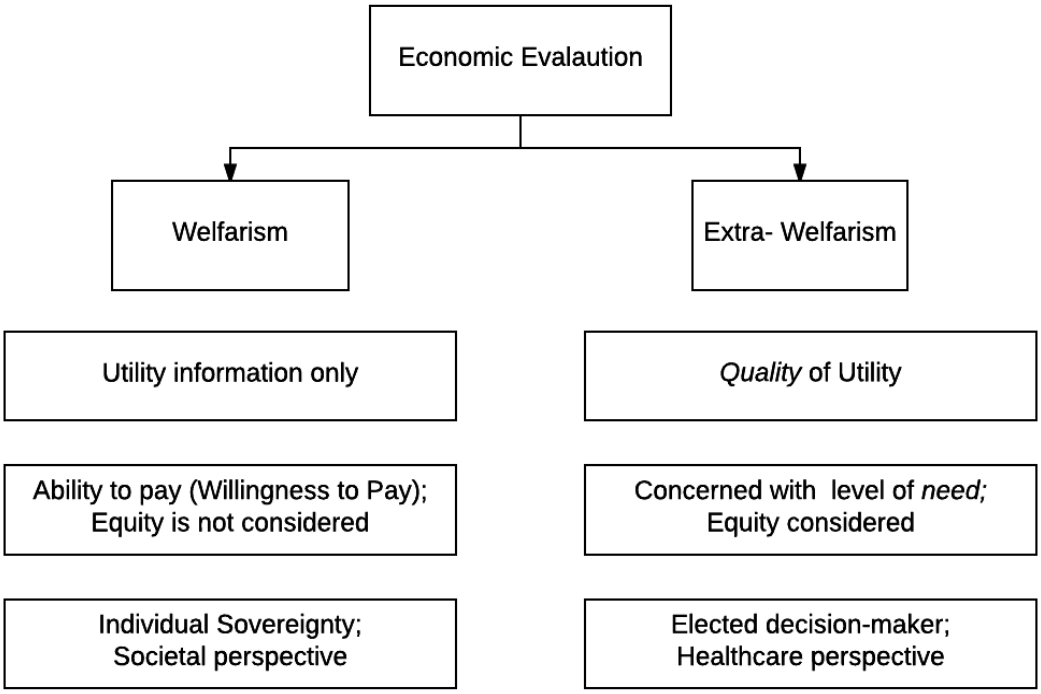
2.2.1 Normative Issues of Economic Evaluations and Productivity

In the context of England's NHS, the inclusion of the impact of productivity in economic evaluations of healthcare interventions remains a contentious issue (Drummond et al. 2015). This section now describes the two main normative frameworks used in health economics: welfarism and extra-welfarism, and to discuss the motivations provided by each that support or reject the incorporation of productivity in economic evaluations.

2.2.2 Welfarism and Extra-Welfarism

Welfarism and extra-welfarism are two normative frameworks that have emerged in the analysis of health and healthcare (Brouwer et al. 2008). Figure 2.1 illustrates the broad differences between welfarist and extra-welfarist schools of thought.

Figure 2.1 Welfarism vs. Extra-Welfarism



Welfarists assume that individuals are rational in their decision making and only consume those goods and services that maximise their own personal utility. Individuals are also assumed to be the best judges of their own consumption of goods and services that maximise their own personal utility (individual sovereignty). Individual utility is solely derived from outcomes that come as a direct result of consuming goods and services. The processes that led to the outcome are not valued. As a consequence, social welfare decisions are based on demand and willingness or ability to pay (Brouwer et al. 2008).

Extra-welfarists have criticised welfarism for basing social decisions on individual utility only (Brouwer et al. 2008). One of the main criticisms is that welfarist theory does not allow for the inclusion of outcomes other than individual utility (Sen 1980). Sen (1980) argued that a broader perspective should be adopted to incorporate the quality of utility. He proposed that the ability to function (functioning) and the ability to make a choice (capability) are both important factors that also, alongside utility, need to be taken into account when valuing social welfare. Birch and Donaldson (2003) argued that welfarism is not constrained to consider only utility attached to commodities and that non-goods characteristics (processes, capabilities) do matter but only in so far as these characteristics have an effect on individual utility.

Extra-welfarism concerns itself with *need* rather than ability to pay. Demand caused by need is profoundly different from demand arising from want. Need implies health care is a necessity for an individual with a health problem, whereas demand, arising from wants, describes an individual's ability to pay (Morris et al. 2012). Need has been interpreted by economists as *the capacity to benefit* meaning that a sick individual who needs healthcare must be able to benefit from the service. The 'need' argument is related to the issue of equity. Extra-welfarism considers how equitable the distribution of benefits and losses are across members of society and is done by attaching equity weights to utilities or health (Brouwer et al. 2008) (see chapter 6, section 6.3.1). An example of basing decisions according to need rather than ability to pay is the NHS provided for the United Kingdom's (UK) population. The NHS was established on the principle that good healthcare should be available to all regardless of the individual's ability to pay (wealth) (NHS 2015).

Extra-welfarism also differs from welfarism through the election of a 'decision-maker'. Sugden and Williams (1978) proposed individual sovereignty be replaced by a decision-maker who determines the values needed for public decision-making. The incorporation of a decision-maker may have practical implications regarding the choice of perspective when carrying out an analysis. Welfarism adopts a societal perspective consistent with welfarist views of social welfare; in contrast, extra-welfarism allows alternative perspectives to be specified, for example the healthcare sector. In the context of the NHS in England, a healthcare sector perspective is adopted to reflect the context in which resource allocation decisions are made using a fixed healthcare budget (see section 2.2.5). Alternative relevant perspectives for economic evaluations could also be specified, such as the perspective of the employer, who funds WPI and requires information on the relative costs and benefits to decide whether an intervention is a good use of their finite budget for improving the health and productivity.

Extra-welfarism is, however, not without its criticisms in both a theoretical and practical sense. Theoretically, it has been argued that extra-welfarism does not sufficiently incorporate considerations of the true opportunity costs of health (Stephen Birch and Donaldson 2003). Birch and Donaldson (2003) argue that extra-welfarism ignores important trade-offs between individual utility derived from health and the utility generated by alternative sectors, such as education. This means that the importance of health as a maximand, the goal set by extra-welfarists, may not be consistent with aims of individuals that make up society (Stephen Birch and Donaldson 2003). Inconsistent

aims set by the state (Government) may lead to the creation of poorly designed policies where individuals continue to pursue their own best interests (Stephen Birch and Donaldson 2003).

2.2.3 Should productivity be included?

Welfarism recommends that all costs and benefits, including productivity, of an investment decision be taken into account. Some health economists argue that productivity lost due to ill-health is an important factor that must be taken into account for two reasons. First, lost productivity has an influence over the scarcity of resources affecting the overall level of societal wealth (Koopmanschap et al. 1995). Second, the additional monetary benefits accrued from higher productivity through taxation could be used to fund healthcare or other public sectors (Koopmanschap and Rutten 1994).

Others have strongly objected to the inclusion of productivity for reasons associated with equity (Williams 1992; Olsen and Richardson 1999). Including productivity as a cost in an economic evaluation is potentially problematic because it may influence and produce cost-effective results that favour of the 'most productive' (those who earn higher wages) individuals in society. Decisions that systematically favour one group over another will lead to inequitable allocations of health across society.

2.2.4 Positive issues of Economic Evaluation and Productivity

Methods of economic evaluation of healthcare aim to identify the relevant costs and consequences of healthcare interventions, programmes, and policies, the results of which are used to inform the optimal allocation of resources (Dakin et al. 2014). The perspective adopted in an economic evaluation and the methods selected to evaluate the costs and benefits are determined by the theoretical underpinnings of welfarism and extra-welfarism.

The following two sections discuss how the two normative frameworks, welfarism and extra-welfarism, influence the choice of the perspective adopted and the methodological framework used to assess the costs and benefits (CEA/CUA or CBA). Welfarism and extra-welfarism inform decisions related to productivity in economic evaluations of healthcare interventions.

2.2.5 Study Perspective and Costs

The study perspective chosen for a healthcare economic evaluation should be relevant to the question being asked and to who is asking (Morris et al. 2012). There are a number

of alternative perspectives that can be adopted including a societal, healthcare sector, provider agency (such as an insurance company), employer, or patient. The study perspective determines whether or not a cost is relevant for inclusion into the analysis. There are, broadly, three categories of costs: direct; indirect; and intangible (Drummond et al. 2015). Direct costs refer to those resources incurred by the healthcare sector which include fixed costs, such as rent, heating, and lighting, and variable costs, such as time worked by medical professionals, drugs and laboratory tests. Indirect costs are those that are incurred by the patient or other members of society. The majority of indirect costs are attributable to lost productivity caused by morbidity and mortality (Cooper 2000). Intangible costs are those that represent the non-quantifiable impact of the deterioration in the quality of life experienced by the patient, their family and other members of society (Hodgson 1983). Table 2.1 illustrates the types of direct and indirect costs that may be included using a societal, health sector, employer, or patient perspective.

Table 2.2 Study perspectives and some illustrative relevant costs in the context of NHS England

Perspective	Direct Costs	Indirect Costs
Society	Medical professionals' time. Drugs Medical Devices Laboratory Tests Hospital Administration Utilities (electricity, heating)	Lost productivity Extra hired care Travel costs Sick pay
Healthcare Sector	Medical professionals' time. Drugs Medical Devices Laboratory Tests Hospital Administration Utilities (electricity, heating)	Not included
Employers	Sick pay Cost of productivity enhancing technologies or programmes	Lost productivity
Patient	Medical professionals' time Travel costs	Lost productivity Informal care (home based)

The choice of study perspective is partly determined by the normative framework on which the economic evaluation is based. For economic evaluations conducted from the welfarist point of view the choice of perspective is relatively simple. Welfarism aims to maximise social welfare, therefore, a societal perspective is adopted (Morris et al. 2012). In contrast, the choice of perspective that can be used in extra-welfarist evaluations depends on the authority that has been elected to act as the decision-maker. For example, NHS England's spending decisions are made by decision-makers within the healthcare sector and therefore a healthcare sector perspective is adopted. Other healthcare sectors across the world have also produced their own set of guidelines for conducting and reporting economic evaluations of healthcare interventions. Many other healthcare systems adopt a wider perspective (societal) allowing for the inclusion of productivity in the assessment of the cost-effectiveness of healthcare interventions (ISPOR 2017).

2.2.6 Economic evaluations, WPIs and the Employers' perspective

There are, however, situations where it may be inappropriate to adopt a societal or healthcare sector perspective. Workplace interventions (WPIs) (chapter one, section 1.3) are a set of interventions that are frequently funded for by employers with the aim of improving levels of productivity (Hemp 2004). In such circumstances, the relevant costs and benefits that are relevant for an analysis are those that that fall on the employer. Relevant costs to the employer may include direct costs, such as health insurance or sick

pay, and indirect costs, such as lost productivity (see table 2.1). Of course, the inclusion of some types of costs depends partly on the country that company is based in; for example, employers based in the United States of America may pay for their employees' health insurance a cost that is less likely to be relevant to companies based in the UK.

NICE public health guidelines recommend conducting economic evaluations of WPIs using CCA or CBA methods (NICEb 2012) .These recommendations are problematic for two reasons. First, CCA is not in line with either normative frameworks of welfarism or extra-welfarism and is difficult to use as a method to inform decisions (Brazier and McCabe 2017). Second, CBA converts all costs and outcomes into monetary units which typically involve valuing lost productivity using wages. Valuing productivity using wages has the potential to discriminate and influence decisions away from those who do not earn high income levels (young, retired, females) (Olsen and Richardson 1999).

This thesis focusses on developing a method that quantifies presenteeism using the extra-welfarist normative framework. Extra-welfarism assumes that there is an elected decision maker; in this case the elected decision maker is the employer because it is often the case that employers fund WPIs. Extra-welfarism is also concerned with equity such that individuals, regardless of their WTP, are able to access healthcare treatment. Therefore, in order to be consistent with extra-welfarist ideals, the method for the quantification of presenteeism must be developed based on the employers perspective and consider equity.

2.2.7 Productivity and double-counting

Double-counting occurs when a factor is incorporated into an economic evaluation as a cost *and* benefit. CBA and CEA are two approaches that allow productivity to be counted simply and explicitly as a cost *or* benefit. In contrast, the potential for double-counting productivity in CUA that adopt a societal or employer perspective is more complex.

In 1996, the Washington (US) panel suggested productivity be captured and valued as an effect in the denominator of the incremental cost effectiveness ratio (ICER) (Gold et al. 1996). The equation used to calculate the ICER is given by equation one:

$$ICER = \frac{(C2-C1)}{(E2-E1)} \qquad (1)$$

C1 is the cost of usual care or 'do nothing' and E1 is the effect of usual care or 'do nothing'. C2 and E2 are the costs and effects of the new intervention, respectively. The US panel made the assumption that respondents consider the effect productivity has on their income level when valuing health states of PROMS, such as the EQ5D. The US panel argued that including individual (patient) productivity as cost would lead to double-counting of productivity in the ICER. In 2012, the US panel revised their recommendations and now suggest, based on evidence, that the effects on productivity are unlikely to be captured by PROMs (Sanders et al. 2016).

2.3 Methods of valuing the impact of presenteeism

Current methods used to value the impact of productivity are varied and derive substantially different estimates (Schultz, Chen, and Edington 2009). This section describes the three components that need to be addressed when quantifying presenteeism:

1) How to identify presenteeism
2) How to measure presenteeism
3) How to value presenteeism

At present, existing methods are inadequate to satisfactorily address these three components (Soeren Mattke et al. 2007; Brooks et al. 2010).

2.3.1 Identification and Measurement of Presenteeism

The identification and measurement of presenteeism is complex because the concept is essentially invisible and it is difficult to count due to the lack of available objective measures. Economies, such as the UK economy, predominately consist of service based industries that do not produce tangible outputs; a common example of this is 'knowledge' produced by Universities. Some attempts have been made to identify and measure presenteeism using objective measures, such as logging the amount of calls made in a specific time period (Burton et al. 2006), but these measures are rare (Mattke et al. 2007). The lack of objective measures has led to the development of numerous subjective approaches to identify and measure presenteeism (see chapter one, section 1.6) all of which differ in their conceptualisation of presenteeism, the length of the recall period applied, and the inclusion and consideration of contextual factors. The next section describes how the methods used in self-reported measures of presenteeism differ

in terms of how presenteeism is conceptualised, the length of the recall period applied, and whether contextual factors are considered. Methods used to value presenteeism as a cost or monetary benefit is also described.

2.3.2 Conceptualisation of Presenteeism

The conceptualisation of presenteeism refers to the way in which respondents are asked to estimate their level of lost productivity (absenteeism and presenteeism) whilst at work as a result of their health-condition. Some instruments ask employees to compare their level of productivity to themselves at different time points whilst others ask respondents to compare their productivity levels with colleagues working in a similar job (Mattke et al. 2007) (see chapter one, section 1.6). The various conceptualisations of presenteeism mean that respondents have different reference points from which to judge their level of presenteeism consequently increasing the amount of heterogeneity of reported levels of presenteeism across self-report measures of presenteeism (Brooks et al. 2010).

2.3.3 Recall Period

The recall period is the length of time a participant is asked to think back on and provide answers about their level of productivity. The recall period needs to be short enough to minimise the risk of recall bias (inaccurate recollections of the past), but also be long enough to capture possible fluctuating periods of presenteeism, especially when dealing with chronic conditions (Leggett et al. 2016). Lengths of recall periods used include one day, one week, two weeks, one month, and one year (Mattke et al. 2007). Each recall period is likely to lead to over or underestimations of the volume of presenteeism (Brooks et al. 2010).

2.3.4 Contextual Factors: Job and Industry Type

A common criticism of presenteeism instruments is that very few take into account specific job characteristics, such as heavy lifting, and how these may affect levels or the probability of presenteeism occurring (Verstappen et al. 2012). A study estimated and compared the level of presenteeism in a population of patients with RA and osteoarthritis (OA) using four self-reported measures of presenteeism: the Health and Labour Questionnaire (HLQ); the Work Limitations Questionnaire (WLQ); the Health and Work Performance Questionnaire (HPQ); and the Work Productivity and Activity Impairment Questionnaire (WPAI) (see chapter three, section 3.4.3). None of the four measures used to measure presenteeism identified a significant relationship between lost

productivity and occupation type (Zhang et al. 2010). The finding suggests measures of presenteeism were not sensitive enough to differentiate the impact of lost productivity associated with different job types. In response to this finding, a new self-reported measure for presenteeism was developed by Zhang et al., (2012), the valuation of lost productivity (VOLP), which asks respondents to provide information about their job and the workplace. The VOLP has demonstrated valid measurements in some aspects of quantifying the impact of presenteeism including response rates and correlations between the VOLP and clinical outcomes (Zhang et al. 2012). However, more research is needed to assess the validity of the VOLP as a measure that can be used to capture the level of lost productivity associated with various occupation characteristics.

Some evidence has suggested that the cost of lost productivity depends on the: 1) relative ease of replacing the employee (substitutability); 2) extent to which the employee works as part of a team; and 3) how time sensitive the output is (Nicholson et al. 2006; Mattke et al. 2007). A common criticism regarding presenteeism instruments is their inability to take account of multiplier effects or compensation mechanisms that emerge from team based work.

Multiplier effects are production losses that may occur in team-based work where each member plays a unique and integral role to the productive capacity of the team. In these circumstances, low productivity levels from one employee may also cause a decline in the productivity levels of other members of the team.

Compensation mechanisms, on the other hand, reduce the level of lost productivity that would have otherwise occurred. An example of a compensation mechanism is individuals working overtime or colleagues completing extra work of the sick employee. Jacob-Tacken et al. (2005) argued that productivity costs are often overestimated because short-term absence can be compensated for leading to zero net lost productivity. At present, very few methods used to estimate the impact of presenteeism have been adjusted for multiplier and compensation effects; a possible reason for this may be down to the reliance of collecting information and data from employers (Tang 2015).

2.3.5 Valuation tools: Cost and Monetary Benefit

Converting the volume of lost productivity into a cost or monetary benefit is a complicated task for which no agreed standardised method is used. The following section briefly describes the two most common methods used to value lost productivity.

The first method developed to value lost productivity as a cost to society was the human capital approach (HCA) (Becker 1964). The method was developed to initially estimate the cost of lost productivity from absenteeism and was later extended to presenteeism (Mattke et al. 2007). The HCA adopts a patient perspective and values the amount of lost productivity by multiplying the amount of time an individual is unproductive during a specified time period by the wage (van den Hout 2010). The advantage of the HCA is that it is simple to apply.

Many methodologists argue that the HCA overestimates the cost of lost productivity because it does not take into account compensation mechanisms (Koopmanschap et al. 1995; Jacob-Tacken et al. 2005; Pauly et al. 2008) or multiplier effects (Nicholson et al. 2006). In 1992, Koopmanschap et al., proposed an alternative method, the friction cost approach (FCA), to estimate the cost of lost productivity. The FCA was developed to calculate the cost of absenteeism by totalling the hours not worked up to the moment the replacement employee is hired. Once the sick employee has been replaced the FCA assumes initial production levels are restored. The FCA make a strong assumption implying all employees are perfect substitutes; however, in reality, perfectly substitutable employees are unlikely to exist due to the extra time needed for a replacement to understand systems or complete specific training that the previous employee may have had. The FCA method has also been criticised because it is not underpinned by economic theory and does not give a value to leisure time (Krol, Brouwer, and Rutten 2013). It has also been recognised that the FCA is not very useful for the valuation of presenteeism since the worker is not absent from work and does not needed to be replaced (Lensberg et al. 2013).

2.3.6 Valuation Tools: Impact on consequences

The majority of methods developed to quantify the impact of presenteeism have focussed on valuing it as a cost or monetary benefit (Kigozi et al., 2017; Brooks et al., 2010; Mattke et al., 2007). It is possible to include and value productivity as a non-monetary outcome; however, limited research exists regarding the inclusion and valuation of productivity as a non-monetary outcome.

2.4 Conclusion

In conclusion, it is possible to argue for the inclusion of productivity in economic evaluations of WPIs that are funded by the employer. However, current methods used to quantify presenteeism are markedly heterogeneous producing estimates that vary extensively making it difficult to inform employers of the costs and/or benefits of WPIs.

At present, there is no universally accepted 'gold standard' method that can be used to adequately identify, measure, and value presenteeism as a cost or consequence. Several systematic reviews have investigated the methods used in self-reported measures of presenteeism to identify, measure and value lost productivity. However, limited evidence has explored the theoretical underpinnings of self-reported instruments. An understanding of how and why each measure was developed may allow for the identification of a link between each self-reported measure of presenteeism.

This chapter has briefly described key concepts of health economics and the issues of whether to include or exclude productivity in economic evaluations of healthcare interventions. This chapter also provided information about the key limitations of current methods developed and used to estimate the impact of presenteeism. The challenges associated with estimating the impact of presenteeism, identified by this literature review, inform the subsequent systematic reviews (Chapters 3 and 4). Chapter three presents a systematic review that investigates whether existing self-reported measures of presenteeism, suitable for use in inflammatory arthritic conditions (see chapter 1) are underpinned by economic theory.

Chapter 3 Economic Theory underpinning self-reported measures of presenteeism used for individuals working with Rheumatoid Arthritis, Ankylosing Spondylitis and Psoriatic Arthritis

3.0 Introduction

This chapter reports a study that aimed to identify whether existing self-reported instruments and methods used and quantify presenteeism are underpinned by economic theory. The study had four objectives: 1) to summarise the instruments and methods used to identify and quantify presenteeism; 2) to describe whether the identified instruments and methods aimed to value presenteeism as a cost or consequence ((dis)benefit); 3) to identify and summarise the extent to which economic theory has been used in the development of identified self-reported instruments of presenteeism; and 4) to describe the economic theories, if any, that have been used to underpin the design of self-reported instruments of presenteeism. Anecdotal evidence (see chapter two) suggested that a number of systematic reviews have been published summarising the available self-reported instruments and methods used to identify and quantify presenteeism; however these were published before 2015 and have not assimilated the evidence of the extent to which economic theory has been used to underpin these methods.

This chapter reports a systematic review in the following order. Section 3.1 explains the importance of economic theory and why it is useful to aid the quantification of a subjective construct such as presenteeism. Section 3.2 states the aims and objectives of this review. Section 3.3 describes the methods used to identify relevant studies. Section 3.4 describes the results and section 3.5 presents a discussion followed by a conclusion in section 3.6.

3.1 Why is economic theory important?

A theory is defined as a formal statement of the rules on which a subject of a study is based or of ideas that are suggested to explain a fact or event or, more generally, an opinion or explanation (Cambridge dictionary 2017). In its broadest sense, economic theory provides a set of assumptions that explain how economies interact. According to Hal Varian (1989), economic theory is useful in a theoretical, practical, and empirical sense. Table 3.1 lists the reasons given by Varian (1989) to explain why economic theory is useful.

Table 3.1 Uses of economic theory

Reasons for Use	Description
Substitute for data	Useful when data are not available
Identifies important parameters and how they might be measured	Indicates relevant parameters and therefore the tools that can be used to measure them
Keeps track of the costs and benefits	Considers the opportunity cost
Relates seemingly disparate problems	The relationship between two distinct problems, for example efficiency and equilibrium.
Generates useful insights, even if the theory is wrong	Aids the understanding of complex economic relationships
Provides a method for solving problems	Provides methods that can be used to examine economic relationships
Antidote to introspection	Provides an alternative explanation beyond an individual's experiences
Verifies relationships that seem obvious but are not true	Theoretical analysis may show that in certain circumstances 'obvious' relationships are not true
Quantification and calculation of economic relations	Theory can be used to compute answers to problems

Source: adapted from Varian (1989)

In the context of understanding which economic theories have been used to underpin self-reported measures of presenteeism, the most relevant of the reasons suggested by Varian is that it serves as a foundation from which to build further concepts and methods that may explain and quantify economic relationships.

Numerous self-reported instruments used to identify and quantify presenteeism have been developed over recent years (Brooks et al. 2010). Available self-report instruments have used different methods to quantify presenteeism and, as a result, generate substantially different estimates of the impact of presenteeism (Mattke et al. 2007; Brooks et al. 2010). In contrast with methods used to value presenteeism, such as the human capital approach (HCA) (chapter 2, section 2.3.5); it is currently unknown whether presenteeism instruments were developed based on economic theory or theory, such as psychology theory, in general.

Potentially, identifying those self-reported instruments that have been developed using economic theory may allow researchers to more clearly identify aspects of presenteeism that have not yet been accounted for. For example, Pauly et al., (2008) identified the

need for taking into account multiplier and compensation effects when measuring the impact of presenteeism. The need to account for multiplier and compensation effects, when measuring absenteeism, was identified through economic theory of the production function (Pauly et al. 2002). Pauly et al., went on and applied economic theory of the production function to presenteeism and developed 'presenteeism multipliers' taking into account: 1) substitutability of labour; 2) the degree and dependence of team work and 3) time sensitivity of output (chapter 2, section 2.3.4). Anchoring the development of the 'presenteeism multipliers' clarified the importance of taking into account compensation and multiplier effects and, therefore, how the multipliers fit into the wider context of developing methods to measure presenteeism.

It is important to understand the theoretical underpinnings of an approach to quantify the impact of a subjective construct, such as presenteeism, to enable the development of a robust approach to its identification and measurement. Economic theories provide a common framework from which to develop and link methods that can be used to identify and measure the impact of presenteeism.

3.2 Aims and Objectives

The primary aim of this study was to assess the extent to which self-reported methods used to identify and measure presenteeism, in the context of rheumatoid arthritis (RA), ankylosing spondylitis (AS) and psoriatic arthritis (PsA), are based on economic theory. The secondary aim was to report key gaps in the empirical literature relating to: 1) the quantification of the volume of presenteeism; 2) the identification of the relevant unit costs; and 3) if, and how, to value the dis-benefits of presenteeism. The objectives of the study were to:

1) identify and summarise the extent to which economic theory has been used in the development of self-report presenteeism instruments;
2) describe any identified economic theories used;
3) summarise the historical development of presenteeism instruments;
4) summarise the methods used to quantify presenteeism;
5) Identify whether presenteeism was valued as a cost or consequence by the identified methods.

3.3 Methods

A systematic review was conducted in line with advice and guidelines published by the Centre for Reviews and Dissemination (CRD) (CRD 2009). Systematic reviews are used to robustly summarise all published evidence on a specific topic (Sauerland and Seiler 2005). Systematic reviews may be used to: summarise the volume of relevant clinical evidence and highlight research gaps; identify existing evidence on the cost-effectiveness of healthcare interventions; and identify parameters and their values for decision-analytic modelling (Anderson 2010; Gomersall et al. 2015). The evidence collated by systematic reviews is valuable to enhance decision-makers understanding of the context in which an intervention is cost-effective compared with alternative interventions (Gomersall et al. 2015).

Systematic reviews differ from literature reviews in two ways: 1) systematic reviews employ systematic methods to identify relevant studies (Gomersall et al. 2015); and 2) systematic reviews are conducted using a pre-specified protocol that provides details on the precise question that will be answered, the methods that will be used to identify relevant studies to answer the question, how the quality of each study is assessed, and how the evidence is summarised (Khan et al. 2003). The protocol for this systematic review is presented in Appendix 3.1.

3.3.1 The Search Strategy

This systematic review was conducted in three distinct stages.

- Stage one involved identifying all existing published systematic reviews of methods used to identify and measure presenteeism (called ‘presenteeism instruments’ hereafter).
- Stage two involved selecting the most relevant and appropriate systematic review, found in stage one, to update and identify relevant presenteeism instruments for inclusion into this review. In stage two, presenteeism instruments were identified by screening the identified systematic reviews and updating the search initially conducted by the chosen systematic review to identify newly developed presenteeism instruments. Selection of the most appropriate systematic review to screen and update was informed by the assessment of each review against pre-specified criteria.
- Stage three involved screening the systematic reviews identified in stage one for relevant presenteeism instruments that may have been missed by the selected

systematic review and the references of all included self-report measures of presenteeism studies.

3.3.1 Stage one: Search Strategy

The first stage of this systematic review involved identifying published systematic reviews that focussed on methods used to quantify presenteeism via the electronic database PubMed. PubMed comprise over 26 million citations of the biomedical literature from MEDLINE, life science journals, and online books (National Centre for Biotechnology et al. 2017). It was decided, given the coverage of PubMed, searching this electronic resource alone would be sufficient to identify relevant systematic reviews of presenteeism methods. PubMed (start date 1946 to 7th September 2015) was searched for relevant manuscripts via the OVID SP interface. A search strategy (Appendix 3.2) was created using relevant index and free-text terms including: 'systematic review', 'presenteeism', 'productivity loss or change', 'health-related', 'productivity', 'self-reported instrument(s)', 'work function(ing)', 'method(s)', 'measure(s)' and 'health'. The search was run on 7th September 2015.

3.3.1.1 Stage one: Selection Process

The systematic reviews identified by the electronic search in PubMed were transferred into the referencing software package Reference Manager 12. Three reviewers (CJ, SV and SG) independently screened all titles and abstracts and included or excluded each identified study based on pre-defined inclusion criteria listed in Table 3.2.

Table 3.2 Inclusion criteria

	Inclusion Criteria	**Description**
Study Type	Systematic Review	Systematic reviews that have used systematic methods to identify relevant studies
Focus	Assess methods used to quantify health-related presenteeism	Assess methods used to quantify productivity change associated with other factors, e.g. shirking
Publication Type	English language	Foreign languages

3.3.1.2 Stage one: Data Extraction

One reviewer (CJ) extracted the data from each included systematic review using a tailored data collection form. The following information was extracted from each systematic review: name of the author; year the review was published; the country it

was published in; the title; the aim; the focus of the presenteeism measure; generic or disease-specific; and the name of the productivity loss instrument.

3.3.1.3 Stage one: Critical Appraisal

The reporting standards of the identified systematic reviews were assessed using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) (Moher et al. 2009) (Appendix 3.3). However, because of the type of information reported by the included systematic reviews, not all of the points that make up the PRISMA checklist were relevant. The following points from the PRISMA checklist were considered relevant to critique the systematic reviews included in this study: title; abstract; rationale; objectives; eligibility criteria; information sources; search strategy; study selection; data collection process; data items; synthesis of results; study selection; study characteristics; summary of evidence; limitations; conclusions; and funding.

3.3.2 Stage Two

The aim of the second stage was to update one of the systematic reviews, found in stage one, to identify presenteeism instruments suitable for use in RA, AS or PsA. The identification of presenteeism instruments was completed in two parts: 1) to update the search conducted by the chosen systematic review to identify newly published presenteeism instrument studies; and 2) to identify and select relevant presenteeism instruments included in the chosen systematic review.

3.3.2.1 Stage Two: Selecting the relevant Systematic Review

Selecting the most appropriate systematic review to screen and update was decided using some pre-defined criteria listed in Table 3.3. The first criterion stated that the systematic review needed to be of high reporting quality as measured by PRISMA (Moher et al. 2009). The second criterion required the systematic review focussed on identifying generic or self-reported presenteeism instruments suitable for use in RA, AS and PsA, from a variety of perspectives including societal, employer, and individual. The third criterion judged the appropriateness of the review by the number of presenteeism instruments identified. The fourth criterion considered the date on which the systematic review was published; logically, the later the systematic review the more up-to-date were the results. The final criterion requires the systematic review to give access to their full search strategy. The search strategy must be detailed enough in order to enable replication of the search. The search strategy may have been reported in one of several ways: 1) in the main text of the systematic review; 2) in an appendix or supplementary document; or 3) as an emailed request to the corresponding author.

Table 3.3 Criterion for selection of most appropriate systematic review

Criteria	Description
High PRISMA Score	Scored at least 13 out of 17
Focus	Identify generic or disease-specific methods for use in RA, AS or PsA to identify and measure presenteeism Narrow perspectives adopted
Number of Instruments Identified	A review that has identified a large number of instruments that quantify productivity
Date published	A review that has been recently published compared to the others
Search Strategy Included	Search strategy included in the text, appendix, supplementary documents, or requested and emailed

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis; RA: Rheumatoid Arthritis; AS: Ankylosing spondylitis

3.3.2.2 Stage two: Updating the Search Strategy

The first part of stage two's search strategy involved updating the electronic search strategies, designed by Ospina et al., (2015) (Appendix 3.4). The searches were ran using the same eight databases used by Ospina et al., including: Medline (1946 to September week 3 2015), Embase (1980 to week 40 2015), Cochrane Central Register of Controlled Trials (CENTRAL) (August 2015), PsychINFO (1806 to September week 4 2015), Web of Science (1900 to 6th November, 2015), CINAHL (1937 to 6th November 2015), Business Source Complete (1886 to 6th November, 2015), and ABI inform (1970 to 6th November, 2015). The last search performed by Ospina et al., was on 11th October 2012; therefore, updated searches were constrained to run from the 1st January 2012 and up to the 6th November 2015. The search strategies were made up of specific names of presenteeism instruments, such as the 'Endicott Work Productivity Scale' and more generic terms such as 'productivity' and 'work performance'. The second part of stage two's search was to collect all twenty-three presenteeism instruments included in the review by Ospina et al., (2015).

3.3.2.3 Stage two: Selection Process

All titles and abstracts, identified by the updated electronic search, were double screened by two independent reviewers (CJ, KP) using pre-specified inclusion and exclusion criteria, summarised in Table 3.4. The studies considered relevant for inclusion were those that focussed on developing methods used to identify and measure presenteeism. The first inclusion criterion required the study to have described and discussed the development of the instrument used to quantify presenteeism. Studies that applied existing presenteeism instruments in studies, for example in economic evaluations, were not included. Also, studies that discussed the validity and reliability of the developed presenteeism instrument were not included. It was, however, possible that a single study may have described the development of their presenteeism instrument and reported its performance in terms of its validity and reliability. The studies that discussed both the development and psychometric properties of the presenteeism instrument were considered relevant for inclusion into this review. The second criterion stated that the presenteeism instrument must measure health-related presenteeism as opposed to other forms of productivity loss associated with, for example, shirking. Presenteeism instruments suitable for use in RA, AS or PsA were considered relevant; these included instruments designed for generic health conditions as well as measures designed for inflammatory arthritis only. The third criterion required only those studies

that described the development of the original method. Many presenteeism instruments have been adapted and refined for use in different countries, for example the Working Limitations Questionnaire was adapted for use in Japan (WLQ-J) (Ida et al., 2012). Adapted versions of presenteeism instruments were excluded from this review. The final criterion allows for the inclusion of studies that are in English or have a readily available English translation.

Table 3.4 Inclusion criteria for stage two

	Inclusion Criteria	**Exclusion Criteria**
Study Type	Development of method that quantifies presenteeism	Studies that apply the developed method, for example in economic evaluations Studies that test methods of presenteeism in term of their psychometric properties *and* do not discuss the development of the instrument
Focus	Methods developed for assessing health-related presenteeism Methods developed for assessing generic health or inflammatory arthritis Original development of presenteeism methods	Methods developed for assessing other forms of productivity loss e.g. shirking Methods developed that focus on alternative disease area; for example mental health Adaptations of methods for use in other countries e.g. Work Limitations Questionnaire – Japanese (WLQ-J)
Publication Type	English language	Foreign languages without easily accessible English translations

3.4.3 Stage three: Identifying presenteeism instruments in Systematic Reviews

The final stage of the search for presenteeism instruments was to screen all presenteeism instruments included in the systematic reviews identified in stage one. The studies identified in stages two and three were pooled together to identify the total number of presenteeism instruments suitable for use in the context of RA, AS or PsA. The references of each included study were also screened for potential relevant studies of presenteeism instruments.

3.4.3.1 Stage three: Data Extraction

Data were extracted by one reviewer (CJ) using a tailored data collection form. The data extracted comprised: the name of the author; the year the study was published; the

country in which the study was published; the name of the presenteeism instrument; the aim of the presenteeism instrument; the structure of the instrument; the recall period; whether the instrument could be used to convert the volume of lost productivity into a monetary amount; whether psychometric properties were assessed in the same study; and whether, economic theory, or other theories, were used to underpin the development of the presenteeism instruments identified.

3.4.3.2 Stage three: Critical Appraisal

At present, no formal guidelines exist to assess the reporting quality of the development of presenteeism instruments. Instead, based on knowledge from previous systematic reviews, each identified study reporting the development of a presenteeism instrument was assessed for the following: 1) author, year, country of publication and name of presenteeism instrument; 2) the aim of the presenteeism instrument; 3) the motivation used to develop the instrument; 4) the perspective of the instrument; 5) how the instrument identified presenteeism; 6) how the instrument measured presenteeism; 7) the methods used to value presenteeism, if appropriate; and 8) the interpretation of results generated by the instrument . The critical appraisal did not seek to infer information from the presenteeism instrument itself, but rather what was reported in the study.

3.4 Results

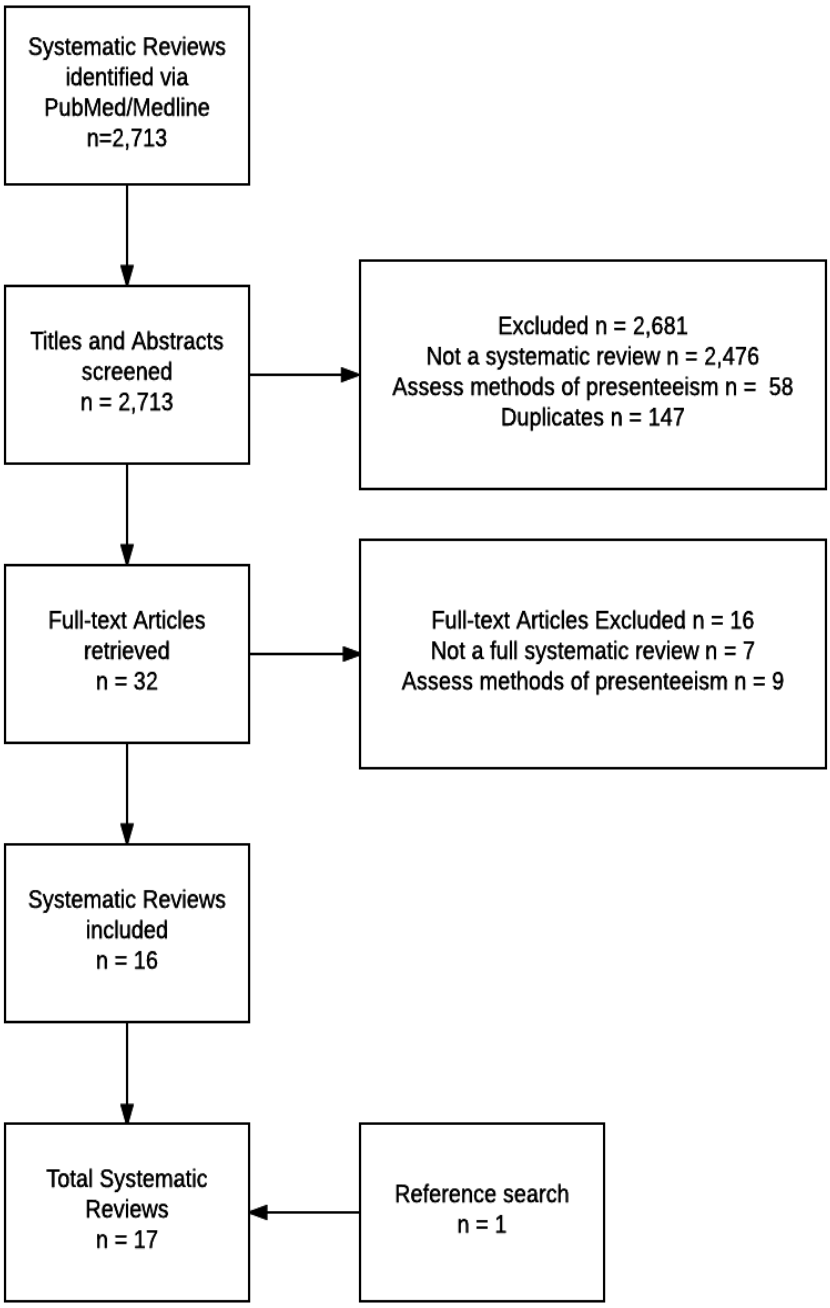
This review identified 24 studies that discussed the development of a presenteeism instrument suitable for use in RA, AS or PsA. The studies were found from a three stage systematic review. The number of studies identified from each stage is now described.

3.4.1 Stage one: Identified Systematic Reviews

In total, 2,713 studies were identified by the electronic search of PubMed. All titles and abstracts were double-screened and 2,681 studies were excluded: 2,476 studies were not systematic reviews; 58 did not assess methods of presenteeism, and 147 were duplicates. Thirty-two full-text articles were retrieved and assessed for inclusion. All 32 systematic reviews were double-screened (CJ and KP) after which a further 16 studies were excluded; seven were excluded because they did not use systematic review methods and nine were excluded because the focus was not on assessing methods of presenteeism. A total of 16 systematic reviews were accepted for inclusion into the final review. The reference lists of all 16 systematic reviews were screened and one

further systematic review was identified and included in the final study. Figure 3.1 illustrates the study selection process and reasons for exclusion at each stage.

Figure 3.1 Selection process stage one



n: number of studies

3.4.2 Stages two and three: Identified Presenteeism Instrument Studies

Figure 3.2 presents a flow diagram of the selection of studies identified in stages two and three.

Selection of the systematic review to update

All 17 systematic reviews, identified from stage one of the search strategy, were quality assessed using PRISMA (Appendix 3.3). Six systematic reviews scored a minimum of 13 of the 17 criteria from the PRISMA checklist (Abma et al. 2012; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Uegaki et al. 2011; Williams et al. 2007). The focus of the six identified systematic reviews was then judged in terms of their relevance with respect to the overall aims of this review. Uegaki et al's., (2011) systematic review focussed on methods used to estimate lost productivity from the perspective of the employer. The constraint on the perspective used was too narrow for this review and the study was excluded as a result. Two systematic reviews; Abma et al. (2012) and Nieuwenhuijsen, Franche, and van Dijk (2010); were excluded because they focussed on identifying instruments only suitable for use in mental health disorders. The review by Williams et al. (2007) was not included because it included six self-reported measures of presenteeism; a number considerably lower compared with reviews by Ospina et al., (2015) and Noben et al., (2014).

The two remaining systematic reviews had high PRISMA scores (14 and 15 out of 17) and focussed on identifying measures of presenteeism (Noben et al. 2014; Ospina et al. 2015). The decision to update Noben et al., (2014) or Ospina et al., (2015) was based on: 1) the number of presenteeism instruments identified; the date the review was published and the successful retrieval of all search strategies used in each electronic database searched. Ospina et al., (2015) identified 23 and Noben et al., (2014) identified 15 presenteeism instruments, respectively. Ospina et al., (2015) published a year later than Noben et al., (2014). The authors of both systematic reviews provided full search strategies used to identify relevant studies. Therefore, based on the number of presenteeism instruments identified and the year of publication the Ospina et al., (2015) systematic review was selected.

The systematic review by Ospina et al., (2015) was used to identify presenteeism instruments suitable for RA, AS or PsA by updating the electronic search and collating

all of the presenteeism instruments included in the original review. The aim of the updated electronic search was to identify newly published presenteeism instruments.

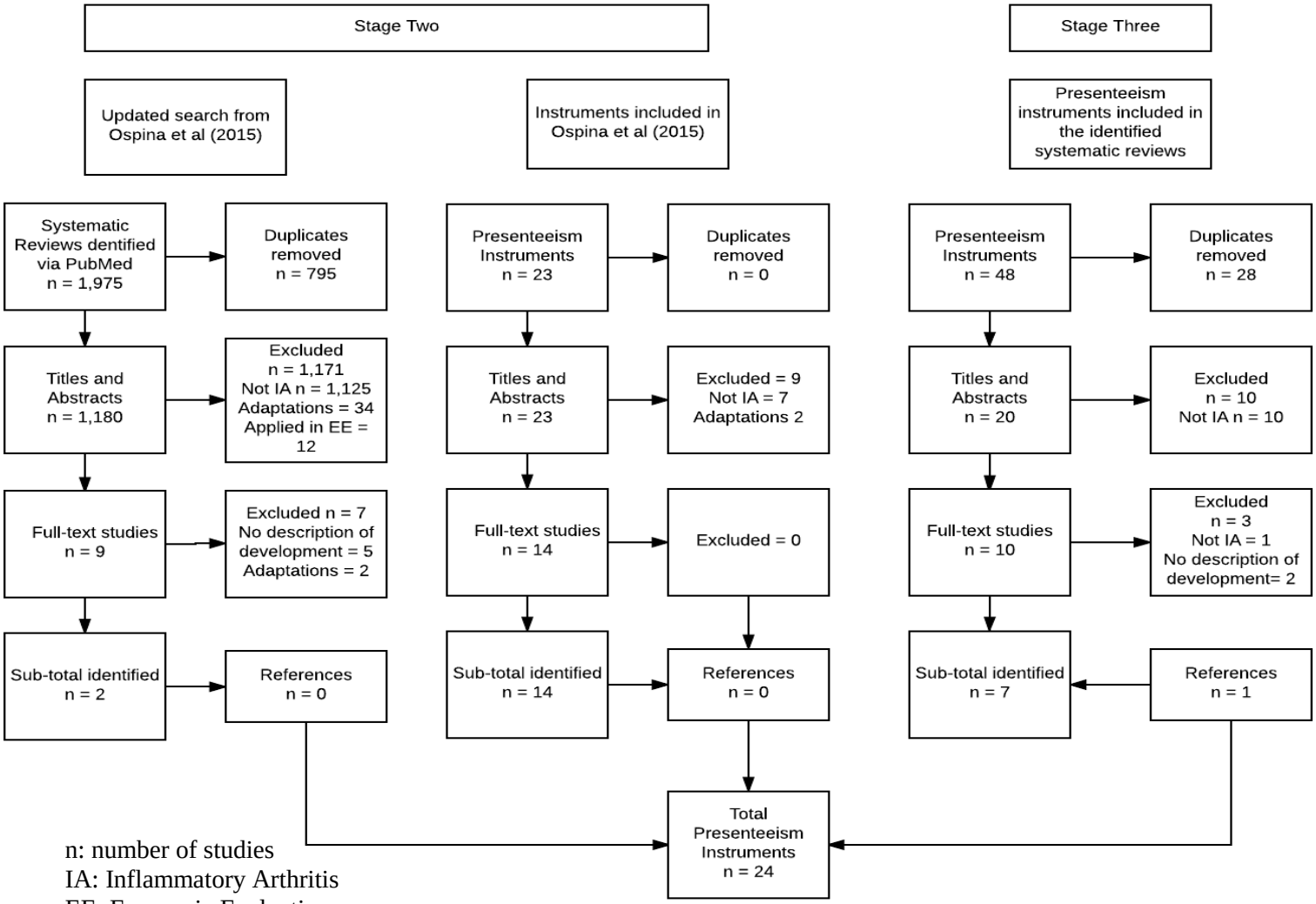
Selection Process: Stage two

The updated electronic search strategy identified 1,975 studies. All titles and abstracts were exported into Reference Manager 12 and 795 duplicates were removed. The remaining 1,180 titles and abstracts were screened by two independent reviewers (CJ and KP) and 1,171 studies were excluded: 1,125 studies were not suitable for use in RA, AS or PsA; 34 studies were adaptations of existing presenteeism instruments; and 12 studies were economic evaluations that used presenteeism instruments. Nine full-text studies were retrieved and a further seven were excluded because: 1) there was no description of the development of the instrument (n = 5) and; 2) the measures were adaptations of the original (n=2). Two presenteeism instrument studies were identified and included in the systematic review.

Twenty-three presenteeism instruments were included in Ospina et al., (2015) and all were screened for inclusion by two independent reviewers (CJ, KP). Nine studies were excluded; seven because the presenteeism measure developed focussed on other disease areas and two because the presenteeism instrument was an adaptation of the original. Fourteen full-text articles were retrieved and no further studies were excluded.

Stage three involved screening the systematic reviews, included in stage one, to identify presenteeism instruments suitable for use in RA, AS or PsA. Forty-eight presenteeism instruments were included across the 17 systematic reviews. Twenty-eight presenteeism instruments already identified in stage two were removed as duplicates. Twenty titles and abstracts of the presenteeism instruments were screened and ten were excluded because they focussed on disease areas other than RA, AS or PsA. Ten full-text studies were retrieved and a further three studies were excluded: one because the instrument was not applicable in RA, AS or PsA; and two because there were no descriptions of the development of the instruments. Seven presenteeism instruments were identified and included in the final review. Finally, the references of each identified study were screened and one further presenteeism instrument was identified and included. In total, 24 presenteeism instruments suitable for use in RA, AS or PsA, were identified and included in the final review.

Figure 3.2 Selection process stage two and three



n: number of studies
IA: Inflammatory Arthritis
EE: Economic Evaluation

3.4.3 General Description of Systematic Review of Methods of Presenteeism

The search identified 17 systematic reviews that critically appraised the methods used to identify, measure, and value presenteeism. Five reviews were published in the USA, four in the Netherlands, three each in the United Kingdom and Canada, and one each in Switzerland and Australia. Five systematic reviews focussed their discussion on the limitations and challenges associated with the methods used to identify, measure, and value presenteeism (Brooks et al. 2010; Filipovic et al. 2011; Mattke et al. 2007; Uegaki et al. 2011; Verstappen et al. 2012). Ten systematic reviews focussed on evaluating the psychometric properties of self-reported productivity loss instruments in terms of their validity, reliability, and responsiveness (Nieuwenhuijsen, Franche, and van Dijk 2010; Lofland, Pizzi, and Frick 2004; Roy, Desmeules, and MacDermid 2011; Ospina et al. 2015; Williams et al. 2007; Noben et al. 2014; Abma et al. 2012; Tang 2015; Prasad et al. 2004; Brown et al. 2014). Two systematic reviews (Loeppke et al. 2003; Despiégel et al. 2012) discussed the methods and psychometric properties of each productivity loss instrument. A summary of each systematic review can be found in Table 3.5.

Table 3.5 Summary of systematic reviews of methods of presenteeism and psychometric properties of self-reported measures of presenteeism

Author (Year) Country	Title	Aim	Disease-Specific Generic	Productivity loss instruments included
Systematic Reviews; Methods				
Mattke et al. (2007) USA	A Review of Methods to Measure Health-Related Productivity Loss	To inform the debate by describing the existing types of instruments and monetary conversion methods and by highlighting limitations and trends in current research	Generic	APAWHI, ALWQ, EWPS, HLQ, HPQ, HWQ, HRPQ-D, MDAQ, MWPLQ, OT, SPS, WHI, WLQ, WPAI, WPAI-AR, WPSI, WPI
Brooks et al. (2010) USA	Presenteeism: Critical Issues	To review the methods that relate to the measurement, conversion and translation of presenteeism	Generic	ALWQ, EWPS, HAQ, HLQ, HPQ, WLQ, WPAI, WHI, HRPQ-D, HWQ, SPS, MIDAS, MWPLQ, OT, WIS,WPSI
Filipovic et al. (2011) United Kingdom	Quantifying the Economic Burden of Productivity Loss in Rheumatoid Arthritis	To examine the methods used to assess the impact of RA on employment and work productivity, and provide an overview of the issues surrounding work productivity loss in the arthritis population (absenteeism and presenteeism)	Generic	HPQ, HWQ, SPS, WLQ, WPAI, WPSI, WIS, HLQ
Uegaki et al (2011) The Netherlands	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	To map out methods used to estimate indirect costs of health-related productivity in economic evaluations of occupational health interventions from a company's perspective	Generic	ALWQ, EWPS, HLQ, HPQ, HRPQ-D, HWQ, LEAPS, MIDAS, MWPLQ, OT, QQ(PRODISQ), SPS-6 and 13, VOLP, WPS, WPAI, WPS-RA, WRF

Verstappen et al (2012) United Kingdom	Methodological issues when measuring paid productivity loss in patients with arthritis using biologic therapies: an overview of the literature	To explore and discuss the variation in methods used in published literature assessing employment/productivity loss in patients with inflammatory rheumatic diseases treated with biologic csDMARDs	Disease Specific; Inflammatory Rheumatic Diseases	HLQ, WAI, WPS-RA, WPAI, WIS-RA/AS
Systematic Reviews; Psychometric Properties				
Lofland et al. (2004) USA	A Review of Health-Related Workplace Productivity Loss Instruments.	To identify health-related workplace productivity loss survey instruments with emphasis on those that captures a metric suitable for direct translation into a monetary unit	Generic	OT, WPAI, HLQ, EWPS, ALWQ, MWPLQ, WPI, HWQ,UHI, WLQ, SPS-6
Prasad et al. (2004) USA	A review of Self-Reported Instruments Measuring Health-Related Work Productivity	To provide a comprehensive assessment of productivity instruments in the current literature that attempt to measure health-related work outcomes	Generic	WPAI, WLQ, HPQ, HWQ, EWPS, HLQ, MIDAS, MWPLQ
Williams et al (2007) Canada	Psychometric Evaluation of Health-Related Work Outcome Measures for Musculoskeletal Disorders: A Systematic Review	To present a comprehensive assessment of the quality of psychometric properties of self-reported health- related work outcome measures for patients with musculoskeletal disorders	Disease Specific; Musculoskeletal Disorders	FACS, ORQ, WRFQ, WLQ-25 and 16, WIS.
Nieuwenhuijsen et al. (2010) The Netherlands	Work functioning measurement; Tools for occupational Mental Health Research	To provide recommendations for instrument development and validation and to examine the properties of available work functioning measures that are potentially applicable to workers with common mental disorders	Disease Specific; Mental Health	Disability profile/LSRDS, EWPS, HLQ, HPQ, HRPQ-D, HWQ, LIFE, NIMH-PQ, OFS, ORQ, OT, QQ (PRODISQ). SAS-SR, SPS, WALS, WHI, WLQ, WL-

				26/WRF, WPAI-GH, WPI, WPSI
Roy, Desmeules, and MacDermid (2011) Canada	Psychometric Properties of Presenteeism Scales for Musculoskeletal Disorders: A Systematic Review	To conduct a systematic review of the quality and content of the psychometric evidence for presenteeism scales for workers with MSDs	Disease Specific; Musculoskeletal Disorders	EWPS, SPS-6 and 13, WALS, RA-WIS, WLQ, WRFQ
Abma et al (2012) The Netherlands	Evaluation of the measurement properties of self-reported health-related work-functioning instruments among workers with common mental disorders	To appraise critically and compare the measurement properties of the identified self-reported health-related work functioning instruments in populations with common mental disorders	Disease Specific; Mental Disorders	EWPS, WLQ, SPS, WPS and LEAPS.
Brown et al. (2014) Australia	Measuring Presenteeism: Which Questionnaire to Use in Physical Activity Research?	To evaluate the characteristics, constructs measured, and measurement properties	Generic	WAI, EWPS, HLQ, HWQ, WLQ, WPSI, SPS, HPQ
Noben et al (2014) The Netherlands	Quality Appraisal of generic self-reported instruments measuring health-related productivity changes; a systematic review	To critically appraise and compare the measurement properties of generic, self-reported instruments measuring health-related productivity changes	Generic	WLQ, EWPS, WPAI, WPSI, AMA-guides, HPQ, SPS, PRODISQ, HRPQ-D, HLQ, QQ, WBA-P, HWQ, WHI, VOLP
Ospina et al. (2015) Canada	A Systematic Review of Measurement Properties of Instruments Assessing Presenteeism	To assess and compare the measurement properties (validity, reliability, and responsiveness) and the quality of the evidence of presenteeism instruments	Generic	ALWQ, EWPS, HLQ, HPQ, HRPQ-D, HWQ, LEAPS, MIDAS, MWPLQ, QQ, SPS-6 and 13,

				VOLP, WHI, WPAI (CD, CG, GERD, GH, IBS, SpA, RA), WPS, WPSI
Tang (2015) Switzerland	Estimating Productivity Costs in Health Economics Evaluations: A Review of Instruments and Psychometric Evidence	To provide a review of instruments with potential for estimating health-related productivity costs	Generic	HLQ, HPQ, HRPQ-D, PRODISQ, QQ, SPS-13, VOLP, WHI, WLQ, WPAI, WPSI
Systematic Reviews; Methods and Psychometric Properties				
Loeppke et al. (2003) USA	Health -Related Workplace Productivity Measurement: General and Migraine-Specific Recommendations from the American College of Environmental Medicine Expert Panel	To evaluate lost workplace productivity measurement tools to examine the characteristics of each instrument to determine the advantages and disadvantages of each tool in the measurement of lost workplace productivity of patients with migraine headache within the context of specific employment settings	Generic and Disease Specific; Migraine	EHCA, HPQ, SPS-6, WLQ, WPAI.
Despiegel et al. (2012) France and the United Kingdom	The Use and Performance of Productivity Scales to Evaluate Presenteeism in Mood Disorders.	To review presenteeism instruments with the aim to provide recommendation on the use of productivity scales in the area of mood disorders	Disease Specific; Mood Disorders	HLQ, HPQ, HWQ, WHI, WLQ, WPAI, LEAPS, EWPS, SDS, SPS.

RA: Rheumatoid Arthritis; csDMARDS: Disease ModifyingAnti-Rheumatic Drugs.

3.4.4 Critical Appraisal of Systematic Reviews

The reporting standards of each identified systematic review varied. The PRISMA scores for each systematic review are presented in Appendix 3.5. Five of the 17 systematic reviews stated the review was systematic in the title (Noben et al. 2014; Ospina et al. 2015; Roy, Desmeules, and MacDermid 2011; Uegaki et al. 2011; Williams et al. 2007). All studies included a structured abstract; however, information commonly missed was the background and data sources. All 17 systematic reviews provided a rationale for the overall aim of the review; however only six studies defined separate objectives of their review (Abma et al. 2012; Filipovic et al. 2011; Loeppke et al. 2003; Nieuwenhuijsen, Franche, and van Dijk 2010; Tang 2015; Williams et al. 2007). Eleven reviews explicitly stated the characteristics of the studies that were to be included (Abma et al. 2012; Brown et al. 2014; Despiégel et al. 2012; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Prasad et al. 2004; Ospina et al. 2015; Tang 2015; Uegaki et al. 2011; Verstappen et al. 2012; Williams et al. 2007).

The methods used to conduct the systematic reviews were poorly reported. The databases used to carry out the systematic searches were reported in all reviews; however, five did not report the dates the search was conducted between (Abma et al. 2012; Brooks et al. 2010; Despiégel et al. 2012; Mattke et al. 2007; Nieuwenhuijsen, Franche, and van Dijk 2010). Five systematic reviews reported a full search strategy within the review or directed the reader to additional files or appendices that contained the full search strategy (Abma et al. 2012; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Uegaki et al. 2011); five reported keywords used to make up their search strategy; however the full strategy was not made available (Lofland, Pizzi, and Frick 2004; Prasad et al. 2004; Roy, Desmeules, and MacDermid 2011; Tang 2015; Verstappen et al. 2012), and seven did not provide a search strategy (Brooks et al. 2010; Brown et al. 2014; Despiégel et al. 2012; Filipovic et al. 2011; Loeppke et al. 2003; Mattke et al. 2007; Williams et al. 2007).

Nine studies reported the process by which studies were selected (Abma et al. 2012; Brown et al. 2014; Despiégel et al. 2012; Filipovic et al. 2011; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Uegaki et al. 2011; Verstappen et al. 2012; Williams et al. 2007) and ten defined the inclusion and exclusion criteria used to assess the eligibility of studies (Abma et al. 2012; Brown et al. 2014; Despiégel et al. 2012; Loeppke et al. 2003; Noben et al. 2014; Ospina et al. 2015; Roy, Desmeules, and MacDermid 2011; Uegaki et al. 2011; Verstappen et al. 2012;

Williams et al. 2007). The data items that were collected from each study were explicitly stated in less than half of the studies (n=8) (Despiégel et al. 2012; Loeppke et al. 2003; Lofland, Pizzi, and Frick 2004; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Prasad et al. 2004; Verstappen et al. 2012) and only six reviews explained how they planned to synthesise the evidence base (Abma et al. 2012; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Roy, Desmeules, and MacDermid 2011; Williams et al. 2007).

Six systematic reviews described the number of studies screened, how many were included in the final review and how many were excluded with reasons. The same six reviews also included a detailed flow diagram illustrating how the number of studies were identified (Abma et al. 2012; Brown et al. 2014; Noben et al. 2014; Ospina et al. 2015; Roy, Desmeules, and MacDermid 2011; Uegaki et al. 2011). The other eleven systematic reviews either reported the number of studies identified and qualitatively explained how the studies were excluded (without the use of a flow diagram) or reported the final number of studies identified and included only (Brooks et al. 2010; Despiégel et al. 2012; Filipovic et al. 2011; Loeppke et al. 2003; Lofland, Pizzi, and Frick 2004; Mattke et al. 2007; Nieuwenhuijsen, Franche, and van Dijk 2010; Prasad et al. 2004; Tang 2015; Verstappen et al. 2012; Williams et al. 2007). All 17 systematic reviews presented characteristics for which data were extracted with citations.

All reviews extensively discussed the main findings of their review and considered the relevance of the results to key groups. Fourteen reviews extensively discussed the limitations of the review and the results (Brooks et al. 2010; Despiégel et al. 2012; Filipovic et al. 2011; Lofland, Pizzi, and Frick 2004; Mattke et al. 2007; Nieuwenhuijsen, Franche, and van Dijk 2010; Noben et al. 2014; Ospina et al. 2015; Prasad et al. 2004; Roy, Desmeules, and MacDermid 2011; Tang 2015; Uegaki et al. 2011; Verstappen et al. 2012; Williams et al. 2007), two mentioned potential limitations of the results (Abma et al. 2012; Brown et al. 2014), and one study did not discuss the limitations of the results or the review (Loeppke et al. 2003).

3.4.5 Results: Systematic Reviews of Presenteeism Methods

Seven systematic reviews critiqued the methods used to identify, measure, and value presenteeism. A summary of those results are now presented.

3.4.5.1 Presenteeism defined by the instruments

The majority (n=14) of the 17 systematic reviews provided a definition of presenteeism. Nine reviews defined presenteeism as the 'reduction in performance (or decrease in productivity) at work caused by a health-related condition' (Brooks et al. 2010; Brown et al. 2014; Despiégel et al. 2012; Loeppke et al. 2003; Prasad et al. 2004; Roy, Desmeules, and MacDermid 2011; Uegaki et al. 2011; Tang 2015; Verstappen et al. 2012); four defined presenteeism as the 'reduction in productivity whilst in paid work' (Filipovic et al. 2011; Lofland, Pizzi, and Frick 2004; Mattke et al. 2007; Ospina et al. 2015); and one review did not explicitly define presenteeism and instead referred to the concept as productivity loss at work (Noben et al. 2014). Three reviews did not provide a definition of presenteeism (Nieuwenhuijsen, Franche, and van Dijk 2010; Williams et al. 2007; Abma et al. 2012).

3.4.5.2 Conceptualisation of Presenteeism

The conceptualisation of presenteeism refers to the way in which respondents were asked to estimate their level of lost productivity whilst at work due to a health-condition. Mattke et al., (2007) and Filipovic et al., (2011) identified three conceptualisations of presenteeism used in self-reported instruments to identify and measure the volume of presenteeism. Respondents were asked to: 1) assess their *perceived impairment* by stating how much their health condition affects their ability to perform mental and physical tasks at work; 2) estimate their productivity loss by *comparing* their level of work, in terms of performance and efficiency, against a work colleague who works in a similar job; and 3) estimate their amount of lost productivity (presenteeism) in terms of time.

3.4.5.3 Measuring Presenteeism

Four systematic reviews reported large differences in the length of the recall period (see chapter 2, section 2.3.3) used by different presenteeism instruments ranging from one day to one year (Brooks et al. 2010; Brown et al. 2014; Verstappen et al. 2012; Despiégel et al. 2012). It is argued that using a recall period that is too short may be limiting the instrument's ability to capture fluctuations that occur in some diseases. For example, an individual with RA may experience increased levels of disease activity for

a period of one week at the beginning of the month. The individual may be asked four weeks later to rate their level of productivity using a recall period of two weeks. The potential productivity loss associated with the flare would not be captured and may bias the results. However, a recall period that is too long may compromise the respondent's ability to accurately remember their levels of presenteeism.

The systematic review by Verstappen et al., (2012) addressed the issue of confounding when attempting to lost measure productivity in patients with RA, AS or PsA who use biological therapies (see chapter one, section 1.4.4). The authors explained how observational studies are more likely to suffer from confounding due to selection bias of the comparison group. Patients in the control group who receive Disease Modifying Anti-Rheumatic Drugs (csDMARDs) were less likely to have severe disease activity and co-morbidities compared to those on biologics. Therefore, the presence of confounding is likely to bias estimates of lost productivity.

Verstappen et al,. (2012) also examined the extent to which contextual factors (see chapter two, section 2.3.4) relating to a patient's job type were used to refine measures of presenteeism. Many studies did not consider including contextual factors in their evaluations, which is surprising because it is understood job characteristics impact on levels of presenteeism (Verstappen et al. 2012).

3.4.6 Results: Systematic Reviews of Psychometric Properties

Twelve systematic reviews discussed and critiqued the psychometric properties of existing measures of presenteeism. Three out of twelve systematic reviews assessed the psychometric properties of self-reported productivity instruments using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist (Abma et al. 2012; Noben et al. 2014; Ospina et al. 2015). The COSMIN checklist was developed to evaluate the methodological quality of studies by examining their psychometric properties based on the information reported in studies (Mokkink et al. 2010). Table 3.6 lists the key psychometric properties, and the associated definitions, that are assessed by the COSMIN checklist.

Table 3.6 COSMIN psychometric properties

Property	Definition
Validity	The extent to which the instrument measures the construct it is designed to measure
Content validity	The degree to which the content (questions) reflects the construct to be measured
Construct validity	The degree to which the instrument produce scores that are consistent with the hypothesis
Criterion validity	The degree to which the new instrument scores agree with the 'gold standard'
Reliability	The degree to which the instrument is free from measurement error
Internal Consistency	The degree of interrelatedness among items
Measurement error	The proportion of the variance that measures the 'true' differences between patients
	The random error of a patient's score that is not attributable to the 'true' differences between patients
Responsiveness	The ability of the instrument to detect change in the construct measured over time

Source: (Mokkink et al. 2010)

Overall, the three systematic reviews that assessed the psychometric properties using the COSMIN checklist found evidence for the majority of productivity loss instruments were limited and consistently, in terms of their validity, reliability and responsiveness, rated the as poor to fair (Abma et al. 2012; Noben et al. 2014; Ospina et al. 2015). Ospina et al., (2015) and Noben et al., (2014) reported construct validity and reliability as the most frequently evaluated psychometric properties. Criterion validity could not be established due to the absence of a 'gold standard' from which to compare results. Noben et al., (2014) also commented on the limited generalisability of the results produced by lost productivity questionnaires attributing it to selective and inadequately sized study samples.

3.4.6.1 Psychometric properties not assessed using a standardised checklist

Eight systematic reviews (Brown et al. 2014; Despiégel et al. 2012; Lofland, Pizzi, and Frick 2004; Roy, Desmeules, and MacDermid 2011; Tang 2015; Williams et al. 2007; Prasad et al. 2004; Loeppke et al. 2003) examined the psychometric properties of self-reported productivity loss instruments without using a standardised checklist. Nevertheless, the eight reviews reported here did assess the majority of the key

measurement properties listed in Table 3.6. The systematic reviews found mixed evidence to support the validity, reliability, and responsiveness of existing productivity loss instruments. Despiégel et al., (2012) and Brown et al., (2014) found the vast majority of productivity loss instruments included in their reviews established construct validity as either convergent or discriminant. However, the reliability, responsiveness, and other forms of validity, including criterion validity, were supported by weak evidence. Roy, Desmeules, and MacDermid (2011) stated that productivity instruments had limited evidence that supported their reliability and responsiveness. Prasad et al., (2004) highlighted that it was difficult to assess criterion validity because of the lack of a 'gold standard' measure from which to compare results. In a recent review, Tang (2015) argued that the validity of all self-reported presenteeism instruments was limited because most studies do not include: 1) a meaningful comparator; and 2) a specific *a priori* hypothesis that states the expected relationship between the instrument and its comparators.

3.4.7 Results from the Presenteeism Instrument Studies

The next section reports the results regarding the use of economic theory to underpin the design of presenteeism instruments suitable for use in RA, AS or PsA.

3.4.7.1 General Description of Presenteeism Instrument Studies

A summary of the results can be found in Table 3.7. Seventeen of the 24 presenteeism instrument studies were conducted in the USA (Jette et al. 1986; Osterhaus, Gutterman, and Plachetka 1992; Reilly, Zbrozek, and Dukes 1993; Endicott and Nee 1997; Burton et al. 1999; Abma et al. 2012; Lerner et al. 2001; Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Goetzel, Ozminkowski, and Long 2003; Kessler et al. 2003; Kumar et al. 2003; Shikiar et al. 2004; Stewart et al. 2004; Turpin et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011), four in the Netherlands (van Roijen et al. 1996; Brouwer, Koopmanschap, and Rutten 1999; Boezeman, Sluiter, and Nieuwenhuijsen 2015; Bouwmans et al. 2015), and one each in the United Kingdom (Gilworth et al. 2003), Canada (Zhang et al. 2012), and Finland (Ilmarinen, Tuomi, and Klockars 1997).

Twenty-three instruments were generic and designed for use across all types of health conditions and one was specifically designed for use in RA (Osterhaus, Purcaru, and Richard 2009). The majority of the identified presenteeism instruments also measured absenteeism (n=17); five do not measure absenteeism (Brouwer, Koopmanschap, and

Rutten 1999; Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Gilworth et al. 2003; Boezeman, Sluiter, and Nieuwenhuijsen 2015); and two are not clear (Amick et al. 2000; Prochaska et al. 2011).

The recall period used by presenteeism instruments were as follows; nine used one month, 30 days or four weeks (Jette et al. 1986; Osterhaus, Gutterman, and Plachetka 1992; Amick et al. 2000; Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Turpin et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011; Bouwmans et al. 2015); five used one week or seven days (Reilly, Zbrozek, and Dukes 1993; Endicott and Nee 1997; Burton et al. 1999; Shikiar et al. 2004); three used two weeks (van Roijen et al. 1996; Lerner et al. 2001; Stewart et al. 2004); and one used the last working day (Brouwer, Koopmanschap, and Rutten 1999). Four presenteeism instruments use various recall periods; Goetzel, Ozminkowski, and Long (2003) used two weeks, three months, and 12 months; Kessler et al. (2003) used one week and 4 weeks; Kumar et al. (2003) used one day and one week; and Ilmarinen, Tuomi, and Klockars (1997) used one year, two years, and a lifetime.

All presenteeism instruments studies, except three (Osterhaus, Gutterman, and Plachetka 1992; Burton et al. 1999; Amick et al. 2000), also tested and reported the psychometric properties of the instrument developed.

Table 3.7 Summary of presenteeism instruments

Author Year Country	Name of Presenteeism Instrument	Absenteeism measured	Structure of Instrument	Recall Period	Monetise Productivity loss	Psycho-metric Properties	Economic Theory
Jette et al 1986 USA	The Functional Status Questionnaire/ Work Performance Scale (WPS)	Yes	Four domain scale; 1) Physical Functioning 2) Psychological Functioning 3) Social/role Functioning 4) Six Single Item Questions Respondents answer a set of statements and assign a grade, for example 1 = usually did not complete and 4 = no difficulty.	1 month	No	Yes	No
Osterhaus et al 1992 USA	Osterhaus Technique (OT)	Yes	12 item questionnaire 1) Presenteeism 2) Absenteeism 3) Healthcare use 4) Demographics Respondents are asked to rate their work performance in terms of %	30 days	Yes	No	No
Reilly et al 1993 USA	Work Productivity and Activity Impairment Instrument (WPAI)	Yes	Questionnaire asks for; 1) Number of days and hours missed from work 2) Days and hours worked 3) Number of days work was difficult 4) Extent to which poor health was attributable to work loss 5) Parallel set of questions (1 to 4) about regular activities of unpaid work	7 days	Yes	Yes	No

			Respondent asked to rate their own working performance. Overall work productivity calculated as a %.				
van Roijen et al 1996 The Netherlands	Health and Labour Questionnaire (HLQ)	Yes	4 Modules; 1) Absenteesim (paid work) 2) Reduced productivity at work (paid) 3) Unpaid work 4) Impediments to paid and unpaid labour Absenteeism: Respondents asked to state whether they performed paid work or not during past two weeks. If large sample mean number of work days lost can be calculated by multiplying by 26 Reduced Productivity: 7 items. Respondents asked to rate 1 = never, to 4 = always. Score equals the sum of all items. Unpaid Production: Respondents asked how many hours spent doing unpaid work. These hours are compared to those of the general population or control group. The mean annual hours of unpaid production calculated by multiplying by 26 Impediments to paid and unpaid labour: Respondents asked to state	2 weeks	Yes	Yes	No

			level of impediment experienced whilst performing the job: 0 = no impediment and 3 = a lot of impediment				
Endicott et al 1997 USA	Endicott Work Productivity Scale (EWPS)	Yes	25 item questionnaire. Domains; 1) Employment status including self-employed 2) Absenteeism 3) Presenteeism Respondents are asked to rate their performance using a 5 point scale; 0 = never and 4 = almost always	1 week	Not Reported	Yes	No
Brouwer, Koopmanschap and Rutten 1999 The Netherlands	The Quality and Quantity Method (QQ)	No	Three Domains: 1) Quantity of work 2) Quality of work 3) Total Efficiency Respondents asked to rate the quantity and quality of work completed using a scale from 1 to 10.	Last working day	Yes	Yes	No
Burton et al 1999 USA	Work Productivity Index (WPI)	Yes	Two Domains; 1) Time away from job 2) Time lost to maintain in the productivity standard Threshold of productivity established: employee score over 0.5 indicates meeting standard productivity = 0 hours lost. Scores less than 0.5 evaluated as a proportion of 0.5 and subtracted from 100%.	1 week	Not reported	No	No

Amick et al 2000 USA	Work Role Functioning Questionnaire (WRFQ)	Not Clear	5 Modules; 1) Work scheduling 2) Physical demands of jobs 3) Mental demands of jobs 4) Social demands 5) Output demands Respondents asked to state the amount of time they have had difficulty meeting the demands of their work. 100% = All the time, 50% half of the time and 0% none	4 weeks	Not Reported	No	No
Lerner et al 2001 USA	The Work Limitations Questionnaire (WLQ-25)	Yes	25 item questionnaire grouped in 4 modules; 1) Time management 2) Physical demands 3) Mental/Interpersonal demands 4) Output demands Respondents asked to rate the level of difficulty or ability to perform 25 specific job demands.	2 weeks	Not Reported	Yes	No
Pelletier and Koopman 2001 USA (see reference; Lynch, Mercer, and Riedel 2001)	Stanford/American Health Association Presenteeism Scale (SAHAPS)	No	42 items grouped in 2 modules; 1) Demographics 2) Presenteeism Respondents are asked to compare their usual performance and rate their performance using 5 point Likert scales and 0-100% scales	1 month	Yes	Yes	No

Koopman et al 2002 USA	Stanford Presenteeism Scale – 6 (SPS-6)	No	Six item questionnaire each designed to capture specific aspects related to presenteeism; 1) Cognitive 2) Emotional 3) Behavioural Respondents are asked to rate their health by stating the degree to which they agree on a scale of 1 to 5. 1 = Strongly disagree, 5 = strongly agree.	1 month	Not Reported	Yes	No
Gilworth et al 2003 UK	Work Instability Scale (WIS)	No	23 item questionnaire capturing information regarding; 1) Health 2) Work situation 3) Physical work factors 4) Hobbies The overall score indicates low, medium or at high risk of work disability. 0 = no problems at work, 4 = Majority of job is unsuitable and individual is unlikely to cope.	Not Stated	Not Reported	Yes	No
Goetzel et al 2003 USA	Work Productivity Short Inventory (WPSI)	Yes	Questionnaire asks for; 1) Demographics 2) Employment status 3) Absenteeism 4) Presenteeism 5) Productivity loss associated with being a caregiver The questionnaire asks about productivity loss associated with 15 health conditions. The respondent is asked to state number of days with the	Various; 12 months 3 months 2 weeks	Yes	Yes	No

			health condition, number of hours unproductive because of health condition and number of days missed from work. The same questions apply to those who are caregivers.				
Kessler et al 2003 USA	The World Health Organisation Work Performance Questionnaire (HPQ)	Yes	Three Domains: 1) Work performance 2) Absenteeism 3) Job-related accidents Respondents asked to rate their overall work performance using a 0 to 10 scale. 0= worst possible work performance, 10 = best possible work performance	Various; 1 week, 4 weeks	Yes	Yes	No
Kumar et al 2003 USA	Health-related Productivity Questionnaire – Diary (HRPQ-D)	Yes	Questionnaire asks for; 1) Premature retirement or reduction to part-time work 2) Absenteeism 3) Presenteeism Questionnaire designed in a diary format. Respondents are required to answer all questions for every day of the week. Some responses require the respondent to estimate their unproductive time other require respondent to choose from list of pre-specified %	Various; 1 day, 1 week	Not Reported	Yes	No
Shikar et al 2004 USA	Health and Work Questionnaire (HWQ)	Yes	24 item questionnaire assessing; 1) Work Quality 2) Work Quantity 3) Impatience 4) Concentration/Focus	1 week	Not Reported	Yes	No

			5) Work Satisfaction 6) Non-Work Satisfaction Respondents are asked to rate their work performance using a 1 to 10 scale: 1=worst and 10=best				
Stewart et al 2004 USA	Work Health Interview (WHI)	Yes	Telephone interview. 6 modules; 1) Informed consent 2) Employment status 3) Health conditions 4) Tasks and activities performed at work 5) Lost Productive Time (LPT): absenteeism and presenteeism 6) Demographics Respondents asked to choose the response that most applies; e.g. all of the time, some of the time, half of the time, none of the time. These responses are then converted into %	2 weeks	Yes	Yes	No
Turpin et al 2004 USA	Stanford Presenteeism Scale – 13 (SPS-13)	Yes	13 item questionnaire. Respondent states their primary health condition and is asked to base all answers given this health state. Presenteeism measured using the Work Impairment Score which is the sum of responses to 10 Likert type questions. The final result is presented as a percentage of lost productivity.	4 weeks	Not reported	Yes	No
Ilmarinen et al 2007	Work Ability Index	Yes	7 item questionnaire capturing information regarding;	Various; 1 year	Not Reported	Yes	No

Finland	(WAI)		1) Presenteeism 2) Health Conditions 3) Absenteeism 4) Mental Health Respondents rate their ability to work using various scales (1-10, 1-4 etc). The index is calculated by summing the ratings given by the respondent.	2 years Lifetime			
Osterhaus, Purcaru and Richard 2009 USA	Work Productivity Survey for Rheumatoid Arthritis or Ankylosing Spondylitis (WPS-RA; WPS-AS)	Yes	3 item questionnaire; 1) Employment status and occupation type 2) Absenteeism and Presenteeism related to paid work 3) Absenteeism and Presenteeism related to unpaid work Respondent is asked to rate the extent to which arthritis has interfered with their ability to work using a scale of 0–10. 0=no interference to 10=complete interference	1 month	Not Reported	Yes	No
Prochaska et al 2011 USA	Well-being Assessment for Productivity (WBA-P)	Not Clear	12 items assess reduced functioning related to personal and work well-being domains; 1) Personal: Health, caring for others, financial, personal issues, depressed/stressed 2) Work: Lack of resources, issues with co-workers, not enough time, issues with supervisors, lack of training, technical issues	4 weeks	Not Reported	Yes	No

			Respondents asked to choose response they most associate: not at all, some, a lot. A single number is estimated comprising of 11 items. The score ranges from 0 (not at all) to 100 (a lot for all 11 reasons).				
Zhang et al 2012 Canada	The Valuation of Lost Productivity (VOLP)	Yes	Five modules; 1) Employment status 2) Absenteeism 3) Presenteeism 4) Unpaid work activity loss 5) Job and workplace characteristics Questionnaire also identifies information: 1) Team dynamics 2) Substitutability of work 3) Time sensitivity of output 4) Compensation 5) Availability of substitutes Respondents are asked to state the amount of extra time they have taken to do job tasks. Also, they are asked to rate their work performance on a scale from 0 to 10.	7 days	Yes	Yes	Yes
Boezeman et al 2015 The Netherlands	Composite Work Functioning Approach	No	Questions based on WLQ and the Tilburg Psychological Contract Questionnaire. 1) Capacity to work 2) Quantity of work 3) Quality of work performance Respondents are asked to rate their	Not stated	Not Reported	Yes	No

			ability to work using a 0 to 4 scale. Weights attached to different aspects of work functioning to reflect relative importance				
Bouwmans et al 2015 The Netherlands	iMTA Productivity Cost Questionnaire (iPCQ)	Yes	Questions based on the short form HLQ and the PRODISQ. 18 items questionnaire. 1) Demographics 2) Absenteeism 3) Presenteeism 4) Unpaid work Respondents are asked to rate their work performance using a 10-point rating scale.	4 weeks	Yes	Yes	No

3.4.7.2 Presenteeism Instruments and Economic Theory

Of the 24 studies that reported the development of presenteeism instruments, only one (Zhang et al. 2012) reported and discussed the use of economic theory to underpin the design of the presenteeism instrument, the Valuation of Lost Productivity (VOLP) see Table 3.7. Zhang et al., (2012) stated that the concept of productivity is based on the theory of the production function where output is a function of inputs including labour, capital and technology. Based on economic theory of the production function, lost productivity due to ill-health was defined as the lost output associated with reduced labour input. Therefore, based on economic theory of production, the aim of the VOLP was to lost value productivity as the output loss resulting from time input loss; different from other measures of presenteeism where the focus was to measure the individual's labour input (time not spent working) (Zhang et al. 2012).

Pauly et al. (2002) first attempted to develop a method that aimed to value lost productivity in terms of output. Pauly et al., (2002) argued that the wage rate is not an accurate reflection of the true value of productivity at the margin because of various other factors such as team production, availability of perfect substitutes, and time sensitivity of outputs (chapter two, section 2.3.4). In order to take account of these three factors, Pauly et al., (2002) estimated a series of multipliers to be used when calculating the value of lost productivity. Pauly et al., generated multipliers using estimations made by managers based on the assumption that managers consider work output, compared to employees who focus on work input, when asked to estimate lost productivity. However, the limitations of the multiplier method include intensive data collection and the possibility that managers may not understand the true effect of productivity loss on their company's output. Also, estimated multipliers may not be generalisable across labour forces based in different countries (estimated using a sample of workers in the USA). The limitations of the Pauly's multiplier method motivated the need to find an alternative way of deriving multipliers (Zhang et al. 2012). Existing presenteeism instruments do not capture information about time input loss and information about team production, availability of perfect substitutes, and time sensitivity of outputs, needed to estimate multipliers (Zhang et al. 2012). The VOLP was designed to capture this information so that it could be used to value lost productivity in terms of output loss associated with reduced labour input caused by ill-health.

The remaining 23 studies did not report underpinning their instruments using other theories from relevant disciplines, such as psychology.

3.4.7.3 Structure of Presenteeism Instruments

The earliest identified measure of presenteeism was designed by Jette et al in 1986; the Work Performance Scale (WPS). The WPS asks the respondent to rate their ability to function physically, mentally and socially. The measure is simple and originally designed for clinical use. In 1993, Reilly et al., (1993) developed the Work Productivity Activity Index (WPAI). The WPAI asks the respondent to state the number of days missed from work and the number of days they found work difficult. The instrument also asks about productivity loss when doing unpaid work. By 2003, presenteeism instruments were being designed to collect additional information regarding the contextual factors of an individual's occupation. For example, the WIS, developed by Gilworth et al (2003), asks questions about the respondent's work situation and physical work factors.

In 2004, Stewart et al., (2004) developed the Work Health Interview (WHI). The WHI is a telephone interview designed to collect information that can be used to estimate the cost of productivity loss. The interview introduces questions about the type of work tasks individuals are expected to complete as part of their job. The WHI is one of the first instruments that explicitly take into account the affect job characteristics may have on levels of presenteeism. In 2012, Zhang and colleagues developed the VOLP, a presenteeism instrument that explicitly takes into account how factors such as team dynamics, availability of perfect substitutes, and time sensitivity of outputs either compensate or multiply levels of productivity loss caused by presenteeism. Since 2015, presenteeism instruments, including the Composite Work Functioning Approach (Boezeman, Sluiter, and Nieuwenhuijsen 2015) and the iMTA Productivity Cost Questionnaire (iPCQ) (Bouwmans et al. 2015), have been developed using questions from pre-existing measures including the WLQ, the HLQ and the PRODISQ rather than developing entirely new presenteeism instruments.

3.4.7.4 Critical Appraisal of Presenteeism Instrument studies

Table 3.8 summarises the reporting standards of studies that developed presenteeism instruments. All studies reported a clear aim for the development of the presenteeism instrument. However, not all studies reported the context for which they envisaged the instrument being used. Five instruments were explicitly reported for use in economic

evaluations (Osterhaus, Gutterman, and Plachetka 1992; Reilly, Zbrozek, and Dukes 1993; van Roijen et al. 1996; Brouwer, Koopmanschap, and Rutten 1999; Lerner et al. 2001; Zhang et al. 2012; Bouwmans et al. 2015), four in clinical settings (including trials) (Lynch, Mercer, and Riedel 2001; Jette et al. 1986; Endicott and Nee 1997; Gilworth et al. 2003), and three in a workplace setting (Goetzel, Ozminkowski, and Long 2003; Shikiar et al. 2004; Boezeman, Sluiter, and Nieuwenhuijsen 2015). Four studies reported their presenteeism instruments could be used in multiple settings; three in a workplace and clinical setting (Kessler et al. 2003; Ilmarinen 2007; Kumar et al. 2003) and one in a clinical setting and economic evaluation (Reilly, Zbrozek, and Dukes 1993). Seven studies did not reported the context for which the presenteeism instrument should be used (Burton et al. 1999; Amick et al. 2000; Koopman et al. 2002; Stewart et al. 2004; Turpin et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011).

The perspective adopted by the presenteeism instrument was reported for only five instruments. Three stated that their presenteeism instrument adopted a societal perspective (van Roijen et al. 1996; Zhang et al. 2012; Bouwmans et al. 2015), one an employee and employer perspective (Boezeman, Sluiter, and Nieuwenhuijsen 2015), and one a patient perspective (Jette et al. 1986).

Developers of presenteeism instruments identified presenteeism in the following ways: nine studies identified presenteeism as work performance (Osterhaus, Gutterman, and Plachetka 1992; Reilly, Zbrozek, and Dukes 1993; Brouwer, Koopmanschap, and Rutten 1999; Stewart et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011; Zhang et al. 2012; Bouwmans et al. 2015; Turpin et al. 2004); four as an amount of work time extra or lost (van Roijen et al. 1996; Burton et al. 1999; Goetzel, Ozminkowski, and Long 2003; Endicott and Nee 1997), one as an output in terms of quantity and quality (Amick et al. 2000), and one as work performance compared to those in a similar job (Jette et al. 1986). Five studies identified presenteeism using multiple concepts: two studies identified presenteeism as work performance and output in terms of quantity and quality (Lerner et al. 2001; Shikiar et al. 2004); one as work performance and amount of work time lost (Kumar et al. 2003); and one as work performance, output (quantity and quality); and comparison of work performance with colleagues in a similar job (Kessler et al. 2003). Four studies did not report how presenteeism was identified (Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Gilworth et al. 2003; Ilmarinen 2007).

Presenteeism was also measured in different ways. Seven studies measured presenteeism as an amount of time (van Roijen et al. 1996; Endicott and Nee 1997; Burton et al. 1999; Amick et al. 2000; Goetzel, Ozminkowski, and Long 2003; Stewart et al. 2004; Prochaska et al. 2011) and five as the percentage effectiveness (Turpin et al. 2004; Osterhaus, Gutterman, and Plachetka 1992; Bouwmans et al. 2015; Boezeman, Sluiter, and Nieuwenhuijsen 2015; Shikiar et al. 2004; Bouwmans et al. 2015). Many instruments reported measuring presenteeism using two methods; six used percentage of work performance and an amount of time (Osterhaus, Purcaru, and Richard 2009; Reilly, Zbrozek, and Dukes 1993; Brouwer, Koopmanschap, and Rutten 1999; Kumar et al. 2003; Kessler et al. 2003; Zhang et al. 2012), and one used percentage of work performance and percentage amount of time (Lerner et al. 2001). Five studies did not report how the instrument measures presenteeism (Jette et al. 1986; Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Gilworth et al. 2003; Ilmarinen 2007).

Four presenteeism instruments were designed for clinical use and therefore did not need to convert presenteeism into cost using monetary units (Jette et al. 1986; Endicott and Nee 1997; Gilworth et al. 2003; Ilmarinen 2007). Of the remaining 21 presenteeism instruments, four reported the methods that can be used to value presenteeism; three explained that it is possible to use the HCA to value presenteeism in monetary units (Osterhaus, Gutterman, and Plachetka 1992; Goetzel, Ozminkowski, and Long 2003; Bouwmans et al. 2015) and one described the valuation of presenteeism using multipliers and the HCA (Zhang et al. 2012). Five studies commented that their instruments could be used to convert the measured volume of presenteeism into a monetary amount; however the methods to do so were not specified (Reilly, Zbrozek, and Dukes 1993; van Roijen et al. 1996; Amick et al. 2000; Stewart et al. 2004; Kessler et al. 2003). The other eleven studies did not discuss the possibility or methods of valuing presenteeism in monetary terms (Brouwer, Koopmanschap, and Rutten 1999; Burton et al. 1999; Lerner et al. 2001; Lynch, Mercer, and Riedel 2001; Koopman et al. 2002; Kumar et al. 2003; Shikiar et al. 2004; Turpin et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011; Boezeman, Sluiter, and Nieuwenhuijsen 2015).

The majority of presenteeism instrument studies (n=21) reported how the results of their instrument should be interpreted (Jette et al. 1986; Osterhaus, Gutterman, and Plachetka 1992; Reilly, Zbrozek, and Dukes 1993; van Roijen et al. 1996; Brouwer, Koopmanschap, and Rutten 1999; Burton et al. 1999; Amick et al. 2000; Lerner et al. 2001; Koopman et al. 2002; Gilworth et al. 2003; Goetzel, Ozminkowski, and Long

2003; Kessler et al. 2003; Kumar et al. 2003; Shikiar et al. 2004; Stewart et al. 2004; Turpin et al. 2004; Osterhaus, Purcaru, and Richard 2009; Prochaska et al. 2011; Zhang et al. 2012; Boezeman, Sluiter, and Nieuwenhuijsen 2015; Bouwmans et al. 2015). Three studies did not provide clear explanations regarding the interpretation of the results of their instrument (Endicott and Nee 1997; Lynch, Mercer, and Riedel 2001; Ilmarinen 2007).

Table 3.8 Summary of the presenteeism instruments

Author, Year, Country, Name of Presenteeism Instrument	**Motivation for development: Aim and Use**	**Perspective**	**Identification of presenteeism**	**Measurement of presenteeism**	**Valuation of presenteeism** **Methods reported**	**Interpretation of instrument** **Includes a copy of the questionnaire**
Jette et al 1986 USA Functional Status Questionnaire	Aim: To measure four dimensions of functional disability needed for clinical use Use: Clinical	Patient	Work performance compared to colleagues in similar job	Amount of time	Not applicable	Interpretation: Yes Questionnaire Included: Yes
Osterhaus et al 1992 USA Osterhaus Technique (OT)	Aim: To identify costs associated with migraine in terms of healthcare resource utilisation and labour costs, using reduced productivity and missed work days Use: Economic Evaluations	Not reported	Work performance	Percentage productivity when working with symptoms Number of days worked with migraine symptoms	Yes Methods: HCA	Interpretation: Yes Questionnaire Included: Yes

Reilly et al 1993 USA Work Productivity and Activity Impairment Instrument (WPAI)	Aim: To measure the effect of general health and symptom severity on work productivity and regular activities Use: Clinical trials and economic evaluations	Not reported	Work performance	Percentage productivity at work Number of hours actually spent working	Yes Methods: Possible to convert into monetary units	Interpretation: Yes Questionnaire Included: No
van Roijen et al 1996 The Netherlands Health and Labour Questionnaire (HLQ)	Aim: To collect quantitative data on the relationship between illness and treatment and work performance Use: Economic evaluations	Societal	Amount of extra time worked	Number of additional hours worked	No Methods: Possible to convert into monetary units	Interpretation: Yes Questionnaire Included: No, available from authors
Endicott et al 1997 USA Endicott Work Productivity Scale (EWPS)	Aim: To assess the degree to which a medical condition affects work functioning Use: Clinical	Not reported	Amount of work time lost	Number of hours lost	Not applicable	Interpretation: Not clear Questionnaire Included: Example provided
Kopec and Esdaile 1998 Canada Occupational Role Performance	Aim: Develop a back pain instrument to measure individual's ability to perform	Not reported	Not reported	Not reported	Not reported	Interpretation: Not clear Questionnaire Included: Yes

(ORQ)	job Use: Not reported					
Brouwer, Koopmanschap and Rutten 1999 The Netherlands The Quality and Quantity Method (QQ)	Aim: To measure the quantity and quality of work performed Use: Economic evaluation	Not reported	Work performance in term of quantity Work performance in terms of quality	Work performance on a scale from 0 to 10 Number of hours lost	Not reported	Interpretation: Yes Questionnaire Included: Yes
Burton et al 1999 USA Work Productivity Index (WPI)	Aim: Make a more complete estimate of the decrease in worker productivity that is associated with health problems by using an objective measure of productivity Use: Not reported	Not reported	Amount of time lost due to lower productivity	Number of lost hours	Not reported	Interpretation: Yes Questionnaire included: No
Amick et al 2000 USA Work Role Functioning Questionnaire (WRFQ)	Aim: The value of successful return to work via functional ability Use: Not reported	Not reported	Completing work on time (quantity) Completing work to a high level of quality	Percentage amount of time	No Methods: Percentage time allows for immediate economic translation	Interpretation: Yes Questionnaire included: Yes

Lerner et al 2001 USA The Work Limitations Questionnaire (WLQ)	Aim: To measure on-the-job impact of chronic conditions and treatment Use: Economic evaluation	Not reported	Work performance Amount of output produced Quality of output produced	Percentage effectiveness Percentage amount of time	Not reported	Interpretation: Yes Questionnaire Included: No, available on request
Pelletier and Koopman* 2001 USA (see reference; Lynch, Mercer, and Riedel 2001) Stanford/American Health Association Presenteeism Scale (SAHAPS)	Aim: Assess outcomes for working populations with chronic conditions often experienced on the job Use: Clinical trials	Not clear	Not reported	Not reported	Not reported	Interpretation: No Questionnaire Included: Yes
Koopman et al 2002 USA Stanford Presenteeism Scale – 6 (SPS-6)	Aim: To assess the relationship between presenteeism, health problems, and productivity in working populations. Use: Not reported	Not reported	Not reported	Not reported	Not reported	Interpretation: Yes Questionnaire Included: Yes
Gilworth et al 2003 UK	Aim: Indicate level of risk of work disability	Not reported	Not reported	Not reported	Not applicable	Interpretation: Yes Questionnaire Included: No, example

Work Instability Scale (WIS)	Use: Clinical					provided
Goetzel et al 2003 USA Work Productivity Short Inventory (WPSI)	Aim: Assess the financial impact of worker absence and on-the-job productivity losses Use: Workplace	Not reported	Amount of work time lost	Number of hours lost	Yes Methods: HCA	Interpretation: Yes Questionnaire Included: Yes
Kessler et al 2003 USA The World Health Organisation Work Performance Questionnaire (HPQ)	Aim: To estimate workplace costs of health problems in terms of reduced job performance, sickness absence, and work-related accidents or injuries Use: Workplace and clinical	Not reported	Quantity of work Quality of work Work performance Comparison of work performance with colleagues in similar jobs	Amount of time Amount of time Likert scale 0 to 10 Likert scale 0 to 10 and qualitatively rated e.g. a lot worse than others	Yes Methods: Not reported	Interpretation: Yes Questionnaire Included: No, available upon request
Kumar et al 2003 USA Health-related Productivity Questionnaire – Diary (HRPQ-D)	Aim: To create a generic instrument for measuring work productivity that would be applicable to different disease states Use: Clinical	Not reported	Work performance during work hours Amount of work time lost	Effectiveness measured as a decimal place between 0 and 1 (similar to percentages) Number of work hours lost	Not reported	Interpretation: Yes Questionnaire Included: Yes

	trials and survey data collections					
Shikiar et al 2004 USA Health and Work Questionnaire (HWQ)	Aim: To assess various aspects of productivity without relying on only self-reported estimations Use: Workplace	Not reported	Quantity of work Quality of work Work performance	Rate themselves from "my worst ever" to "my best ever"	Not reported	Interpretation: Yes Questionnaire Included: Yes
Stewart et al 2004 USA Work Health Interview (WHI)	Aim: To estimate the cost of illness of both absenteeism and presenteeism Use: Not reported	Not reported	Work performance	Number of hours lost	Yes Methods: Not reported	Interpretation: Yes Questionnaire Included: No
Turpin et al 2004 USA Stanford Presenteeism Scale – 13 (SPS-13)	Aim: To assess presenteeism on knowledge based and production jobs, and to provide information on the health condition most likely to affect productivity. Use: Not reported	Not reported	Work performance	Percentage of usual activity	Not reported	Interpretation: Yes Questionnaire Included: No
Ilmarinen et al 2007 Finland Work Ability Index	Aim: To assess work ability Use: Workplace and clinical	Not reported	Not reported	Not reported	Not applicable	Interpretation: No Questionnaire Included: No, example provided

(WAI)						
Osterhaus, Purcaru and Richard 2009 USA Work Productivity Survey for Rheumatoid Arthritis (WPS-RA)	Aim: To assess the impact of rheumatoid arthritis on productivity Use: Not reported	Not reported	Work performance	Interference level measured using a scale 0 = no interference and 10 = complete interference	Not reported	Interpretation: Yes Questionnaire Included: Yes, in an additional data file
Prochaska et al 2011 USA Well-being Assessment for Productivity (WBA-P)	Aim: To capture multiple domains of well-being and productivity Use: Not reported	Not reported	Work performance	Amount of time	Not reported	Interpretation: Yes Questionnaire Included: No
Zhang et al 2012 Canada The Valuation of Lost Productivity (VOLP)	Aim: To value productivity loss in terms of reduced output associated with reduced input Use: Economic evaluations	Societal perspective	Work performance	Number of hours Percentage productivity at work	Yes Methods: Multipliers combined with HCA	Interpretation: Yes Questionnaire Included: No, available upon request
Boezeman et al 2015 The Netherlands Composite Work Functioning Approach	Aim: To estimate weights that reflect the relative importance of the capacity to work, quantity and quality of work, and recovery	Employee and employer	Quantity of work Work performance	Level of work performance (more than usual, as much as usual, less than usual)	Not reported	Interpretation: Yes Questionnaire Included: No

	Use: Workplace					
Bouwmans et al 2015 The Netherlands iMTA Productivity Cost Questionnaire (iPCQ)	Aim: Measuring and valuing productivity losses Use: Economic evaluations	Societal perspective	Work performance	Level of work performance using 10 point Likert scale	Yes Methods: HCA	Interpretation: Yes Questionnaire Included: No

The methods used to identify, measure and value presenteeism used by self-report measures of presenteeism vary extensively. To aid the development and reporting of self-reported presenteeism instruments some initial guidelines are proposed in Table 3.9. Further empirical research is, however, needed to confirm whether the proposed guidelines are sufficient for use in practice (Moher et al. 2010).

Table 3.9 Proposed guidelines for the development and reporting of self-reported presenteeism instruments

Criteria	Description
Aim	The aim of the presenteeism instrument must be clearly specified
Use	To clearly define the; 1) Context for use, e.g. clinical trials or economic evaluations 2) Population, e.g. patient, employee 3) Disease, specify the disease or conditions for which the measure is developed for
Theory	What is the theoretical underpinning of this presenteeism instrument? Explain it. If there is no theoretical foundation, what is the justification for this?
Perspective	State the perspective from which the presenteeism instrument is designed
Recall period	State the recall period used and justify the length chosen
Absenteeism	Does the instrument also measure absenteeism?
Identification	How is presenteeism identified in terms of its conceptualisation? (see Mattke et al. 2007; Filipovic et al. 2011)
Measurement	How is presenteeism measured? E.g. work performance, amount of time. What units are used? E.g. percentage or number of hours
Valuation	Can the volume of presenteeism be valued in monetary units? If so, what methods can be used
Interpretation	Explain how the results from the instrument should be interpreted

3.5 Discussion

This section discusses the findings from the systematic reviews and individual presenteeism instrument studies, respectively.

3.5.1 Discussion of Systematic Reviews

This systematic review identified 17 systematic reviews that assessed the methods and psychometric properties of self-reported instruments of presenteeism. The five systematic reviews that discussed the methods of presenteeism highlighted variation between the techniques used to identify and measure presenteeism. An agreement regarding how it is 'best' to identify and measure presenteeism in self-reported instruments needs to be reached in order to reduce the amount of variation across estimates of presenteeism. The Outcome Measures in Rheumatology group (OMERACT) (see section 1.2) have collected evidence from which to reach a consensus regarding the standardisation of instruments that measure lost productivity (absenteeism and presenteeism) in the context of rheumatic conditions (Beaton et al. 2016). Four self-reported instruments to identify and measure productivity have been recommended for use by OMERACT with the aim of reducing the amount of variation in estimates of presenteeism in patients with rheumatic conditions. The four measures include the Workplace Activity Limitations Scale (WALS), the Work Limitations Questionnaire with modified physical demands scale (WLQ PDmod), the Work Ability Index (WAI), Arthritis-specific Work Productivity Survey (WPS), and the Work Productivity Activity Impairment Questionnaire (WPAI).

The evidence supporting the psychometric properties of self-reported instruments of presenteeism is poor. Far more research is needed to establish the validity, reliability, and responsiveness of all instruments. Criterion validity may only be established once a gold standard method of presenteeism quantification has been agreed upon; however, taking into account the results identified by the systematic reviews that reported methods of presenteeism, it seems unlikely that a gold standard method will emerge soon.

The 17 systematic reviews make three broad and similar recommendations for future research. First, no single self-reported productivity loss instrument can be recommended for use in studies that attempt to identify and measure the impact of presenteeism (Abma et al. 2012; Despiégel et al. 2012; Roy, Desmeules, and MacDermid 2011; Tang 2015; Ospina et al. 2015; Brooks et al. 2010). Second, future research should focus on

testing the validity and reliability of productivity loss instruments, with special attention towards criterion validity. Third, there is a need to establish best practice guidelines that set standards, assumptions, and limitations when working with and developing self-reported productivity loss instruments so that the comparability and transparency of results improves (Abma et al. 2012; Ospina et al. 2015; Williams et al. 2007; Brooks et al. 2010; Uegaki et al. 2011).

3.5.2 Discussion of Presenteeism Instruments

This systematic review has identified 24 self-report instruments of presenteeism. Only one of the presenteeism instruments, the Valuation of Lost Productivity questionnaire (VOLP), was explicitly underpinned by economic theory. The majority of self-reported instruments were found to be a-theoretical, which is problematic because the rationale, construct, and development of the instruments cannot be linked. This is the first systematic review that has investigated the theoretical underpinnings of presenteeism instruments.

Zhang et al. (2012) was the only study that described how economic theory of productivity was used to inform the design of the VOLP. Zhang et al. (2012) argued that in order to be consistent with the economic theory of production; productivity loss should be measured as the change in output. The VOLP values output loss as a result of working time (input) lost by collecting data on workplace and job characteristics (Zhang et al. 2012). Interestingly, the VOLP also takes into account multiplier and compensation effects accounting for: 1) substitutability of labour; 2) teamwork; and 3) time sensitivity of output. Zhang et al. (2012) developed a method for deriving multiplier effects using the methods and theory described by Pauly et al. (2002 and 2008). The link between the methods used by Pauly et al. and Zhang et al. was possible because of the common economic foundation from which both methods were developed and is an excellent example of how economic theory can aid the development of measures of presenteeism.

The absence of economic theory used to support instruments that identify and measure presenteeism may have contributed to the way in which researchers have approached the development of presenteeism instruments. The lack of a theoretical model for presenteeism means that researchers do not have a common framework from which to begin their research and develop their ideas. Therefore, as research into the measurement of presenteeism has grown the instruments developed have become more

differentiated and more complex. Presenteeism instruments developed in the 1980's and 1990's are relatively simple where the respondent is asked to provide information about their perceived level of absenteeism and presenteeism based on their health condition. In comparison, those presenteeism instruments developed in the late 2000's ask the respondent to consider a wide range of factors including job characteristics, team dynamics, time-sensitive output, job satisfaction, job security and relationships with colleagues, as well as the direct impact on presenteeism caused by their health condition. Ospina et al., and Noben et al., recommended the development of more self-reported presenteeism instruments is not needed and instead the literature should focus on improving instruments that already exist. To some extent, this is happening where the two latest presenteeism instruments, the Composite Work Functioning Approach and the iPCQ, use questions from pre-existing presenteeism instruments.

At present, the evidence base regarding the challenges that face self-reported measures of presenteeism are clear. It is recommended that researchers focus their attention on: 1) reaching a consensus regarding existing methods; 2) testing psychometric properties where there is limited evidence; and 3) exploring new methods, beyond self-report instruments, that may accurately identify and measure presenteeism.

3.5.1 Limitations

The studies that were used to identify whether or not presenteeism instruments were developed using economic theory did not provide extensive detail. Many studies provided limited information that described how the presenteeism instrument was developed. In an area where the quantification of a concept is subjective it should be encouraged that researchers publish information about the conceptualisation and development of their presenteeism instrument. Such information would help inform the correct application and interpretation of the instrument, especially in the absence of applying an economic framework from which to underpin the instrument.

This systematic review is the first to critically appraise the reporting standards of studies that reported the development of presenteeism instruments. It is clear that studies were primarily concerned with reporting the psychometric properties of their developed instrument; however this seems to come at the expense of reporting the development of the instrument itself. Key information regarding the context or setting for which the presenteeism instrument was developed, the perspective adopted by the instrument, and whether it is possible to convert the volume of presenteeism into monetary units and the

methods used to do so, were missing. Guidelines that inform the development and reporting of self-report presenteeism instruments would allow issues to be highlighted allowing for progress in the methods used to identify, measure, and value presenteeism using self-report presenteeism instruments. At present, it is not yet clear which instruments are the most appropriate for measuring presenteeism in the context of health conditions in general and IA specifically.

A limitation of the design of this systematic review regards the strict inclusion/exclusion criteria used to identify presenteeism instruments. Only those studies which reported the development of the presenteeism instrument were selected. It may be possible that this review has missed some relevant presenteeism instruments that could have been included in this study; however the aim here was to assess each instrument based on its development and not to identify every instrument suitable to measure presenteeism in the context of RA, AS and PsA.

3.6 Conclusion

This review has systematically identified all self-reported presenteeism instruments, providing a historical context, and whether presenteeism instruments are underpinned by economic theory. With the exception of the VOLP, none of the instruments were explicitly underpinned by economic theory. One key area for further research is to take account of the need to understand how to identify, measure, and value the impact of presenteeism, while underpinning these stages with relevant economic theory for each constituent part of this process. Economic theory would aid the correct interpretation and application of the self-report presenteeism instrument and valuation approach. It is also vital that further development of presenteeism instruments are informed by robust empirical studies that take account of the context in which the final instrument will be used. There is also a need to develop and publish reporting criteria for the development of self-reported presenteeism instruments to promote comparability of measures and the results they generate.

Chapter 4 If, and how, is the impact of productivity included in economic evaluations of workplace interventions for musculoskeletal disorders?

4.0 Introduction

This chapter reports a systematic review to identify if, and how, the impact of productivity was included in economic evaluations of workplace interventions aimed at managing musculoskeletal disorders (WPI-MSDs). The sections in this chapter are as follows. Section 4.1 describes WPI-MSDs, the current evidence base regarding the role of productivity in studies of WPIs, and the rationale for undertaking this review. Section 4.2 states the aim and objectives of the review. Section 4.3 describes the methods used to identify relevant studies for inclusion into the systematic review. Section 4.4 presents the results and sections 4.5 and 4.6 provide a discussion and conclusion, respectively.

4.1 Background

Musculoskeletal disorders (MSDs) describe a group of conditions that affect the body's joints, tendons, muscles, and ligaments. MSDs include inflammatory diseases (e.g. rheumatoid arthritis (RA), degenerative diseases (e.g. osteoarthritis), vasculitis (e.g. systemic lupus erythematosus), non-articular conditions (e.g. back pain), and bone diseases (e.g. osteoporosis) (Al Maini et al. 2015). The prevalence of MSDs increases with age and negatively impacts productivity and employment (Woolf and Pfleger 2003). In the UK, the total number of working days lost due to MSDs in 2015/16 was estimated to be 8.8 million days (Health and Safety Executive 2016). In a seminal report by Dame Carol Black it was reported that MSDs cost the UK economy in excess of £60 billion per year due to sick leave (Black 2008).

Many employees with MSDs are at increased risk of presenteeism, absenteeism, and job loss which can become costly to the employer in terms of output loss, loss of skills, recruitment and re-training (Black 2008). Employers recognise the importance of aiding the management of symptoms at the workplace to promote productivity and may, therefore, provide support by investing in workplace interventions for individuals with MSDs (WPI-MSDs) (Hemp 2004; Cancelliere et al. 2011). Typically, WPIs are funded by the employer from a fixed annual budget (see chapter 1, section 1.3 for more information about WPIs). This presents the employer with an investment decision. To avoid making the wrong investment decision it is important that employers are fully

informed of the expected costs and consequences of WPIs and the associated opportunity cost.

In 2007, the UK's Department of Health commissioned the National Institute for Health and Care Excellence (NICE) to systematically review the literature that assessed the cost and consequences of WPIs designed to prevent or reduce long-term sickness absence or promoted return to work after a period of sickness absence (Hillage et al. 2008). The aim of the review was to provide a summary of the costs and consequences of WPIs to inform and develop guidelines for good practice for primarily employers, but also primary care services, to manage long-term sickness and incapacity. Hillage and colleagues developed a specific search strategy to identify economic evaluations of WPIs that aimed to reduce the number of employees moving from short to long-term sickness and support return to work. The authors identified economic evaluations including cost-effectiveness analysis (CEA), cost-utility analysis (CUA), cost-benefit analysis (CBA), and also cost-analysis studies as defined by Drummond et al. (2015) (see chapter two, section 2.1). The search was run between 1990 and 2007. The review identified eleven economic evaluations of WPIs, ten of which focussed on interventions for people with MSDs and one for minor mental health disorders. Six studies conducted CBA, three CEA, and one conducted a CBA and CEA. All economic evaluations were trial based, ten were randomised control trials (RCTs) and one was a non-randomised trial. Hillage et al., (2008) found most interventions were developed to aid return to work and concluded that the majority of the studies (7 out of 10) found exercise-based interventions and multi-disciplinary treatment strategies were likely to be cost-effective in terms of achieving return to work. Substantial design heterogeneity across studies was identified and there was a lack of conclusive evidence about the relative cost-effectiveness of WPIs. With regards to productivity, limited information about its inclusion or exclusion, and the methods used to identify, measure, and value it, was reported by the authors of the 2008 review.

In a later systematic review, the aim was to investigate if, and how, lost productivity was included in economic evaluations of WPIs from the perspective of the employer (Uegaki et al. 2011). Uegaki et al., created a search strategy which ran from 1966 to April 2007. Economic evaluations were defined as those studies that performed a CEA, CUA, CBA, cost-analysis, or 'financial appraisals'. Financial appraisals were defined by the authors as 'the costs and consequences of two or more alternatives where the monetary consequences were limited to differences in healthcare utilisation and/or

productivity valued using market prices'. The authors included and defined 'financial appraisals' to make the distinction between cost-analysis studies where health-related productivity was considered as a cost not consequence. The review identified 34 economic evaluations of WPIs from the perspective of the employer. Thirty-two studies conducted a "financial appraisal" and two conducted CEA. Of the 34 studies included 17 focussed on MSDs, seven on influenza, two on mental health, two on migraines, and one on general health.

Uegaki et al. (2011) found estimates of absenteeism were included in all 34 studies, whereas presenteeism estimates were included six studies only. The latter was to be expected because presenteeism was a relatively new concept when the review was conducted. The review also summarised how each study measured and valued absenteeism and presenteeism. Absenteeism was measured using data on sick leave that was self-reported, reported by others, for example employers or colleagues, or recorded in administrative databases. Presenteeism, on the other hand, was measured using non-standardised or standardised questionnaires. All studies valued absenteeism and presenteeism using the Human Capital Approach (HCA) (chapter two, section 2.3.5) however, the time and price units used varied extensively across studies. Uegaki et al., (2011) concluded that, due to variation of the methods used to identify, measure, and value lost productivity, the results could not be used to inform employers decisions to invest in WPIs; a result that is consistent with many systematic reviews that have explored the methods used to estimate lost productivity (Jones et al. 2016; Brooks et al. 2010; Mattke et al. 2007).

Another, more recent, systematic review aimed to identify the methods used to value the impact presenteeism has on the total costs of interventions and health conditions (Kigozi et al. 2017). Kigozi et al., identified three economic evaluations and 25 cost-of illness (COI) studies that valued presenteeism as a cost. The review reported the methods used to identify, measure, and value presenteeism as a cost in each study. The review identified nine self-report instruments that were used to measure presenteeism and all studies, except one, valued the volume of presenteeism as a cost using the HCA, with one study using FCA. Overall, the review by Kigozi and colleagues concluded that the inclusion and estimation of presenteeism as a cost only was limited in current economic evaluations. The limitation of the review by Kigozi et al., is that it focussed on presenteeism valued as a cost only; previous evidence shows that presenteeism may be also valued as a monetary (cost) or consequence ((dis)benefit). By ignoring the

possibility that productivity may be captured as a consequence, analysis of how productivity is incorporated in economic evaluations was not able to be considered. Therefore, one of the aims of this review was to establish how productivity, absenteeism and presenteeism, is incorporated in economic evaluations as a cost or consequence. The review by Kigozi and colleagues also did not identify economic evaluations of WPIs, studies that could have been captured based on the search strategy they reported. The reviews by Hillage et al., and Uegaki et al., highlight that there exists a literature of economic evaluations of WPIs, an intervention that is highly likely to include productivity. A second aim of this review was to identify the literature that the Kigozi review did not include and provide an assessment of the treatment of productivity in economic evaluations of WPIs.

The two existing published systematic reviews of economic evaluations of WPIs (Hillage et al. 2008; Uegaki et al. 2011) provide a good summary and critique of the evidence base but only up to and including studies published in 2007. Importantly, to date, no-one has systematically summarised the methods used to identify, measure and value the two types of productivity loss, absenteeism and presenteeism, in published economic evaluations from all perspectives including societal, healthcare sector, employer, and patient/individual. It was necessary to update the systematic reviews to understand whether robust methods have been used to identify, measure, and value lost productivity in economic evaluations of WPIs.

4.2 Aims and Objectives

The aim of this study was to identify if, and how, the impact of lost productivity, in terms of absenteeism and presenteeism, has been identified, measured, and valued in economic evaluations of WPIs aimed at managing MSDs (WPIs-MSDs). The objectives of this study were to:

1) Identify the number of published economic evaluations of WPIs-MSDs that included estimates of the impact of productivity in their analysis;
2) Summarise whether absenteeism, presenteeism or both concepts were included in economic evaluations;
3) Report the methods used to identify, measure, and value the impact of productivity;
4) State whether estimates of the impact of productivity were presented as a cost or consequence (monetary benefit, or non-monetary benefit);

5) Describe how estimates of lost productivity, as a cost, monetary benefit, or non-monetary benefit, were incorporated into economic evaluations of WPIs-MSDs;
6) Comment on the reporting quality of the impact of productivity included in economic evaluations of WPIs-MSDs.

4.3 Methods

A systematic review was conducted in line with guidelines published by the Centre for Reviews and Dissemination (CRD) (CRD 2009) and the Preferred Reporting Items for Systematic Reviews (PRISMA) (Moher et al. 2009). The protocol for this systematic review is presented in Appendix 4.1.

The scope of this review was extended to identify WPIs suitable for MSDs. The decision to focus on WPI-MSDs was based on the findings of two previous systematic reviews of WPIs by Hillage et al., (2008) and Uegaki et al., (2011) where zero economic evaluations of WPIs individuals with RA, AS, and PsA were identified.

4.3.1 Search Strategy

Hillage et al., (2008) designed a search strategy to identify economic evaluations of WPIs from all relevant perspectives (societal, healthcare sector, employer), whereas Uegaki et al., (2011) created a search strategy was created to identify economic evaluations of WPIs from the perspective of the employer only. For the purposes of this review, it was decided that limiting the search to identify economic evaluations of WPI-MSDs from the perspective of the employer would be too narrow and may fail to capture important information from economic evaluations conducted from alternative perspectives. Kigozi et al., (2017) developed a search strategy to identify all economic evaluations and COI studies that incorporated presenteeism as a cost only. To meet the aim of this systematic review, economic evaluations that included absenteeism and presenteeism as a cost or benefit needed to be identified; it was decided that Kigozi et al's., search strategy would be too limited to meet this aim, therefore, the most appropriate search strategy to update was the one designed by Hillage et al., (2008).

The strategy used to identify relevant economic evaluations of WPIs comprised three stages: 1) screening the studies included in the review by Hillage et al., (2008); 2) updating the electronic database search ran by (Hillage et al. 2008); and 3) screening the references of each relevant study identified for inclusion in this systematic review.

Stage one

Stage one of the search involved screening the studies included and excluded by the systematic review by Hillage et al., (2008). In total, Hillage et al's., (2008) review identified 100 economic evaluations of WPIs, of which 89 were excluded. The reasons for exclusion of the identified economic evaluations were reported within the review and are summarised in table 4.1.

Table 4.1 Exclusion criteria

Exclusion Criteria	Details
Location	Studies conducted in non-OECD (Organisation for Economic Co-operation and Development)
Populations	Self-employed Pregnant women on sick-leave due to pregnancy or maternity leave Unemployed individuals
Interventions, programmes, policies and strategies	Prevention Pregnancy Non-sickness related Delivered outside workplace or primary care (Hillage et al. 2008) Private health insurance schemes or statutory or occupational sick pay Pharmacological or therapeutic interventions Clinical diagnosis
Study Type	Descriptive studies (non-evaluative) Dissertation/Theses Non-English language

Source: Hillage et al. (2008)

Stage Two

The electronic search created by Hillage et al., (2008) was combined with the National Health Service Economic Evaluation Database (NHS EED) economic evaluation search filter. Hillage et al. (2008) ran their search from 1990 to 2007; the search ran by this review started from 1st January 2007 until 17th September 2016. The search strategy by Hillage et al. (2008) was originally performed in three databases: Health Economic Evaluation Database (HEED); 2) National Health Service Economic Evaluation Database (NHS EED); and 3) Econlit. However, from the 31st December 2014 and 31st March 2015, HEED and NHS EED databases, respectively, ceased being updated. As a consequence, Medline and Embase, two databases that HEED and NHS EED used to search and collect economic evaluations, were used to update the search. The electronic search strategy was, therefore, performed using the following databases: Medline (1946

to 2015 September week 2); Embase (1974 to 2015 16th September); and Econlit (1886 to August 2015). The full search strategy used to identify studies in all three databases can be found in Appendix 4.2.

Stage three

The final stage of the search involved screening the reference lists of each identified study included in stages one and two.

4.3.2 Study Selection Process

All identified titles and abstracts were double screened by four independent reviewers (CJ and KP, SV and BG). The inclusion criteria were based around the 'PICO' (population; intervention; comparator; outcome) framework; each criterion used in this review are defined and described in Table 4.2.

Table 4.2 Inclusion criteria

Inclusion Criteria		**Definition**
Population	Musculoskeletal disorders (MSDs)	Defined by Al Maini et al., (2015), see appendix 4.3. Workplace injuries were excluded
	Adult employees	Participants who are 16 years old and above with MSDs and work full or part-time. The sample may be identified as patients, however it must be clear that the patients in the study are employed
Intervention	Workplace Intervention	Non-pharmacological programmes or strategies focussed on improving work-related productivity (e.g. return to work, prevent sickness absence, reduce presenteeism) Interventions that focused on MSD injury caused by work were not included
Comparator	Alternative intervention or no intervention	Interventions that are used in current practice (usual care) or do nothing.
Outcomes	Economic Evaluation	Full cost-benefit analysis, cost-utility analysis and cost-effectiveness analysis as defined by Drummond et al., (2015). Cost minimisation, cost-comparison studies and financial appraisals were excluded
Publication Language	English	Available in English or with a readily available translation

4.3.3 Data Extraction

One reviewer (CJ) extracted the data from each study using a tailored data extraction form. There are no published criteria to appraise productivity; therefore, the following data was extracted from each published study: the name of the author, the year the study was published, and the country the study was published in; the type of analysis conducted (CEA/CUA or CBA); the perspective of the study adopted; methods to identify and measure the impact of lost productivity; the methods used to value the impact of lost productivity; how the impact of lost productivity was presented (as a cost or consequence; and how estimates of productivity were incorporated into the economic evaluation.

4.3.4 Critical Appraisal

Uegaki et al., (2011) examined the reporting standards of economic evaluations of WPIs using the Consensus on Health Economic Criteria (CHEC-list) (Evers et al. 2005). The CHEC-list assesses the methodological quality of economic evaluations in systematic reviews by asking 19 yes/no questions about the inclusion of key factors reported in the study (see Appendix 4.4). Uegaki et al., found that fewer than half of the studies (44%) reported 50% of the quality criteria defined by the CHEC-list. The authors concluded that employers using information provided by studies with poor quality reporting standards may lead to inappropriate decisions being made on the investment of WPIs. The economic evaluations identified in this updated review were assessed for the quality of their reporting standards using a tailored version of the Consolidated Health Economics Evaluation Reporting Standards (CHEERS) (Husereau et al. 2013). The CHEERs checklist (Appendix 4.5) was developed to consolidate and update existing guidelines that can be used to assess the reporting quality of health economic evaluations (Husereau et al. 2013). The CHEERs checklist was selected over the CHEC-list to assess the quality of the reporting standards of identified economic evaluations of WPIs-MSDs included in this review for two reasons: 1) the CHEERs checklist assesses 24 important aspects that should be reported in an economic evaluation; and 2) CHEERs is a more up-to-date list than the CHEC-list.

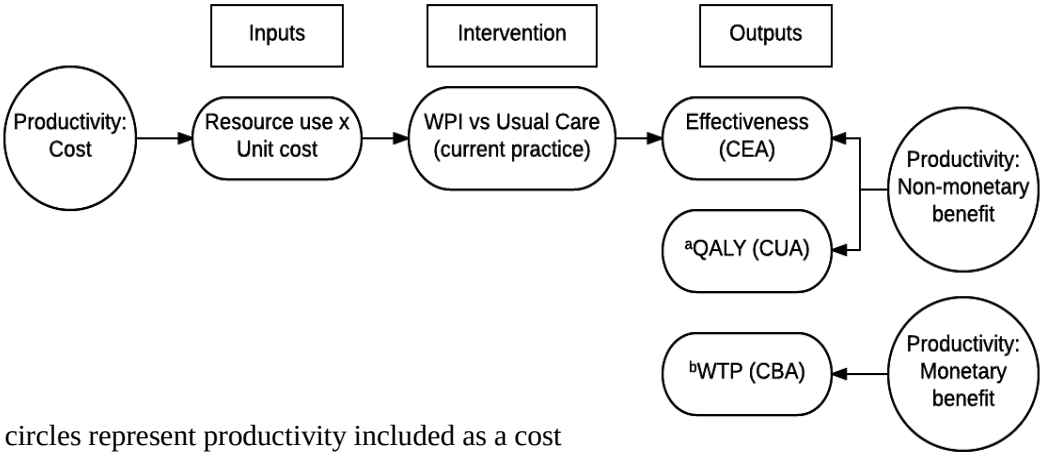
The CHEERS checklist states that it may be used to evaluate the reporting standards of all types of economic evaluation (CEA, CUA and CBA). However, in an editorial by Sanghera, Frew, and Roberts (2015) it was reported that the CHEERS checklist was

comprehensive for CUA and CEA, but significantly limited in its ability to assess CBA. The CHEERS checklist lacks guidelines needed to standardise the reporting of outcomes in CBA. The authors acknowledge that there are no standard reporting guidelines for CBA and, therefore, the methods used to measure and value the benefits should be 'explicit and unambiguously' reported (Sanghera, Frew, and Roberts 2015). In this review, CBA studies were critically appraised by: 1) assessing the reporting quality using relevant criteria included in the CHEERS checklist; and 2) providing a summary regarding the clarity of reporting methods used to measure and value the benefits.

4.3.5 Reporting Quality of Productivity in Economic Evaluations

There are three main types of economic evaluations methods that can be used to assess the cost and consequences of an intervention; CEA, CUA and CBA (see chapter two, section 2.1). Productivity may be included in these economic evaluations of WPIs-MSDs as a cost or consequence (Figure 4.1). If productivity is categorised as a cost then the value of the impact of lost productivity will be measured as a resource use and valued using a unit cost. On the other hand, if productivity is categorised as a consequence ((dis)benefit) it will be measured and valued according to the economic evaluation method used. In CEA or CUA productivity is valued as a non-monetary benefit and in CBA productivity is valued in monetary units. If productivity is included as a cost (or consequence) it should not also be included as a consequence (or cost) in order to avoid double-counting.

Figure 4.1 Including productivity in economic evaluations



The circles represent productivity included as a cost *or* as a monetary *or* non-monetary benefit

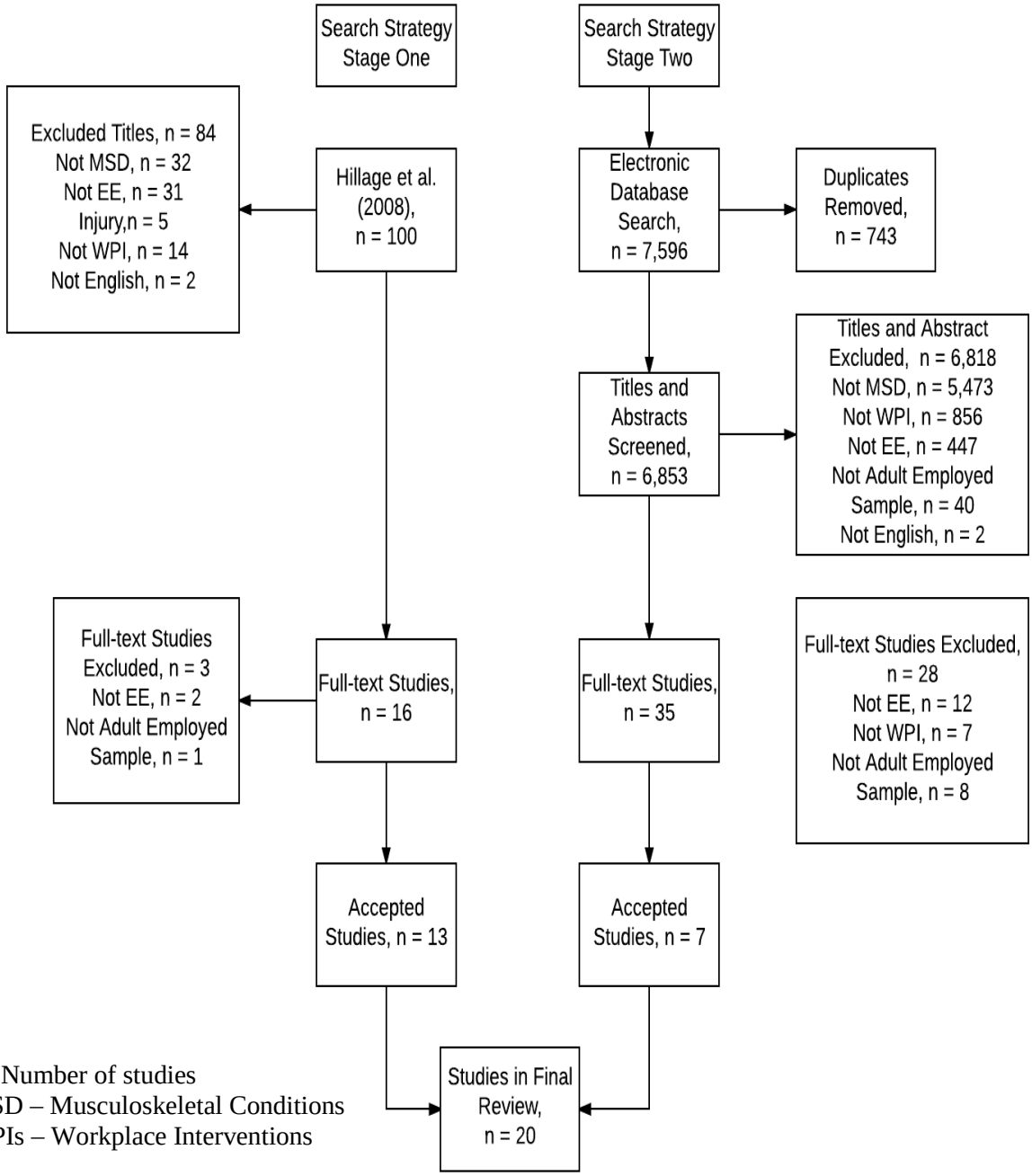
[a] Quality Adjusted Life Year (QALY)
[b] Willingness to Pay (WTP)

4.4 Results

In total, 20 economic evaluations of WPI- MSDs were included in the final systematic review. The results of the search are illustrated in Figure 4.2. During stage one; sixteen of the 100 economic evaluations identified in the Hillage et al., (2008) were identified as potentially relevant.

Stage two identified seven economic evaluations of WPIs. A total of 7,596 journal articles were identified by the updated electronic search: Medline (n=2,433); Embase (n=4,657); and Econlit (n=506). After the electronic removal of duplicates (n=743) using Reference Manager (Version 12), 6,853 titles and abstracts were identified for screening. Titles and abstracts were excluded because they: were not focussed on MSD (n=5,473); did not assess a WPI (n = 865); were not full economic evaluations (n=447); did not use an employed population (n=40); and were not in English or an English translation was not available (n=2). Thirty-five studies were identified as potentially relevant and after retrieving and screening the full-text articles (CJ, SV and KP) a further 28 studies were excluded: 12 were not full economic evaluations; seven did not assess a WPI; one was not MSD specific; and eight did not use an employed study sample. No further studies were identified via stage three (hand search of reference lists).

Figure 4.2 Selection of economic evaluations of WPIs



n - Number of studies
MSD – Musculoskeletal Conditions
WPIs – Workplace Interventions

4.4.1 General Description of Economic Evaluations of WPI

Table 4.4 summarises the studies included in this review using the key headings from the CHEERS reporting criteria. The target populations used in each study were described as: employees (n=10) (Bernaards et al. 2011; Bültmann et al. 2009; Driessen et al. 2012; Meijer et al. 2006; Niemistö et al. 2005; Phillips et al. 2012; Roelofs et al. 2010; Schweikert et al. 2006; Speklé et al. 2010; Tompa, Dolinschi, and Natale 2013); sick listed patients (n=9) (Grahn, Borgquist, and Ekdahl 2004; Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Hlobil et al. 2007; Jensen et al. 2013; Leon et al. 2009; Squires et al. 2012; Steenstra et al. 2006; Torstensen et al. 1998); and employed patients (n=1) (Lambeek et al. 2010). The WPIs evaluated were designed to help individuals' with: low back pain (n=10) (Bültmann et al. 2009; Hagen, Grasdal, and Eriksen 2003; Hlobil et al. 2007; Jensen et al. 2013; Lambeek et al. 2010; Niemistö et al. 2005; Roelofs et al. 2010; Schweikert et al. 2006; Steenstra et al. 2006; Torstensen et al. 1998); neck and back pain (n=1) (Driessen et al. 2012); neck and upper limb pain (n=2) (Bernaards et al. 2011; Speklé et al. 2010); upper limb problems (n=1) (Meijer et al. 2006); and general MSDs (n=6) (Grahn, Borgquist, and Ekdahl 2004; Haldorsen et al. 2002; Leon et al. 2009; Phillips et al. 2012; Squires et al. 2012; Tompa, Dolinschi, and Natale 2013).

Twelve studies defined the comparator as usual care and provided information about what usual care entailed (Grahn, Borgquist, and Ekdahl 2004; Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Lambeek et al. 2010; Meijer et al. 2006; Phillips et al. 2012; Roelofs et al. 2010; Schweikert et al. 2006; Speklé et al. 2010; Squires et al. 2012; Steenstra et al. 2006; Tompa, Dolinschi, and Natale 2013). Two studies also defined the comparator as 'usual care' but did not describe what was involved (Bernaards et al. 2011; Hlobil et al. 2007).Two studies used no intervention (do nothing) (Driessen et al. 2012; Bültmann et al. 2009) and three studies used alternative interventions as the comparator (Jensen et al. 2013; Niemistö et al. 2005; Torstensen et al. 1998).

Of a total of 20 economic evaluations of WPIs: eight were conducted in the Netherlands (Bernaards et al. 2011; Driessen et al. 2012; Hlobil et al. 2007; Lambeek et al. 2010; Meijer et al. 2006; Roelofs et al. 2010; Speklé et al. 2010; Steenstra et al. 2006); three in Norway (Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Torstensen et al. 1998), two each in the United Kingdom (Phillips et al. 2012; Squires et al. 2012); Denmark (Bültmann et al. 2009; Jensen et al. 2013); and one each in Canada (Tompa,

Dolinschi, and Natale 2013), Sweden (Grahn, Borgquist, and Ekdahl 2004), Finland (Niemistö et al. 2005), Spain (Leon et al. 2009) and Germany (Schweikert et al. 2006).

Of the studies that adopted one perspective the societal was most commonly used (n=9) (Bültmann et al. 2009; Grahn, Borgquist, and Ekdahl 2004; Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Lambeek et al. 2010; Niemistö et al. 2005; Roelofs et al. 2010; Schweikert et al. 2006; Steenstra et al. 2006). Three studies used an employer's perspective (Bernaards et al. 2011; Hlobil et al. 2007; Tompa, Dolinschi, and Natale 2013), and one the perspective of the healthcare sector (Phillips et al. 2012). Five studies carried out analyses that adopted more than one perspective. Two studies adopted a societal and employer perspective (Driessen et al. 2012; Speklé et al. 2010), one study a taxpayer and healthcare perspective (Jensen et al. 2013), and one study an individual and societal perspective (Meijer et al. 2006). Squires et al., (2012) adopted three perspectives: healthcare; societal; and employer. Two studies did not clarify the perspective they adopted (Leon et al. 2009; Torstensen et al.1998).

Nineteen studies were trial-based evaluations and one study was a model-based evaluation (Squires et al. 2012). Seventeen studies used data from randomised control trials (RCTs) and three used cohort data (Phillips et al. 2012; Tompa, Dolinschi, and Natale 2013; Grahn, Borgquist, and Ekdahl 2004). Phillips et al., (2012) and Tompa, Dolinschi, and Natale (2013) conducted pre and post intervention analyses and Grahn, Borgquist, and Ekdahl (2004) used a prospective follow up study design. The model-based evaluation used a Markov model to assess the cost-effectiveness of WPI-MSDs over a life time horizon (Squires et al. 2012). The model was populated using data from meta-analyses, systematic reviews, and national statistics. Eleven of the 20 studies carried out one-way sensitivity analyses (Bernaards et al. 2011; Bültmann et al. 2009; Grahn, Borgquist, and Ekdahl 2004; Hlobil et al. 2007; Lambeek et al. 2010; Phillips et al. 2012; Roelofs et al. 2010; Speklé et al. 2010; Squires et al. 2012; Steenstra et al. 2006; Tompa, Dolinschi, and Natale 2013).

Five studies reported they had conducted CEA (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Meijer et al. 2006; Niemistö et al. 2005; Speklé et al. 2010), six studies reported they had conducted a CBA and two studies reported they had performed a CUA (Phillips et al. 2012; Schweikert et al. 2006). Seven studies presented more than one analysis: three studies reported a CEA and CBA (Bültmann et al. 2009; Driessen et al. 2012; Jensen et al. 2013); three studies reported both CUA and CEA

(Roelofs et al. 2010; Squires et al. 2012; Steenstra et al. 2006): and one study reported CEA, CUA and CBA (Lambeek et al. 2010). Of those studies that reported a CEA, five actually reported cost-consequence analysis (CCA) where incremental cost effectiveness ratios (ICERs) for several outcomes were reported (Driessen et al. 2012; Meijer et al. 2006; Niemistö et al. 2005; Roelofs et al. 2010; Speklé et al. 2010). CCA is a form of health economic evaluation where direct and indirect costs are totalled and weighted against several outcomes which is different to CEA where a single outcome is used in the analysis (Coast 2004). In total, nine studies conducted CBA and the results were reported as: the net present value (NPV) or net social benefit (NSB) of the intervention (n=6) (Bültmann et al. 2009; Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Lambeek et al. 2010; Leon et al. 2009; Tompa, Dolinschi, and Natale 2013); one as net savings (Hlobil et al. 2007); and one as “no monetary benefit” (Driessen et al. 2012). The remaining two CBA studies did not report the results of their analysis clearly (Jensen et al. 2013; Torstensen et al. 1998).

Table 4.3 Summary of presenteeism instruments for musculoskeletal disorders

Author (Year) Country	Aim	Population	Musculoskeletal Disorders	Intervention and Comparator	Study Type, Perspective and Time Horizon	Type of EE	Results and Sensitivity Analysis/Type?
Bernaards et al (2011) The Netherlands	To evaluate the cost-effectiveness of work and lifestyle interventions	Dutch computer employees	Neck and upper limb problems	Intervention: 1) Workstyle intervention (WS) 2) Lifestyle Physical Activity (WSPA) Intervention Comparator: Usual care (not defined)	Study Type: CEA alongside an RCT Perspective: Employer Time Horizon: 6 and 12 months	CEA: Overall recovery and pain	CEA: Not cost-effective for overall recovery, results not reported S.A: Yes Type: One-way
Bültmann et al (2009) Denmark	To compare the effect of a coordinated tailored work rehabilitation intervention	Danish employees absent from work for 4 to 12 weeks	Low back pain	Intervention: Coordinated tailored work rehabilitation intervention (CTWR) Comparator: No intervention received by control group	Study Type: CBA and CEA alongside a RCT Perspective: Societal Time Horizon: 12 months	CEA: Averted absence day CBA: Average productivity loss	CEA: ICER = DKK183 ($33.70) per averted absence day (excluded productivity loss) CBA: Net-benefit = DKK 58,973 ($10,666) per person. S.A: Yes Type: One-way
Driessen et al (2012) The Netherlands	To present the results of an economic evaluation of participatory	Dutch employees aged between 18 and 65	Low back pain and neck pain	Intervention: Participatory ergonomics programme	Study Type: CEA and CBA alongside a cluster RCT	CEA: Prevalence of low back pain and neck pain	CEA: ICER = €23,749 per extra unit of adverse effect (low back pain prevalence)

	ergonomics (PE) programme to prevent low back and neck pain	years		Comparator: No intervention	Perspective: Societal and employer Time Horizon: 12 months	CBA: Lost productivity costs (benefit) using FCA	CBA: Cost of PE was €29 per worker and did not result in any monetary benefit S.A: No
Grahn et al (2004) Sweden	To compare the six-year outcome of a multidisciplinary rehabilitation program with continued care within primary care	Patients with long and/or repeated short periods of sick leave during the past year	Musculoskeletal disorders	Intervention: Multidisciplinary Rehabilitation Comparator: Usual Care: Primary Care	Study Type: CEA alongside a prospective controlled study Perspective: Societal Time Horizon: 6 years	CEA: Quality of Life measured by the Nottingham Health Profile (NHP)	CEA: ICER = -306 SEK per unit increase in NHP per patient S.A: Yes Type: One-way
Hagen et al (2003) Norway	Investigate the long-term effects of a Light Mobilisation Program	Norwegian employees between 18 and 60 years on sick list for 8 to 12 weeks	Low back pain	Intervention: Light Mobilisation Program Comparator: Usual care: treated within primary healthcare	Study Type: CBA alongside a RCT Perspective: Societal Time Horizon: 3 years	CBA: Number of days on sick leave converted into economic returns	CBA: Net present value = 6,574,349 NOK S.A: No
Haldorson et al (2002) Norway	To assess the cost-effectiveness of three treatments based on screening	Norwegian employees on sick-list for more than 8 weeks	Musculoskeletal Pain	Intervention: 1) Light multidisciplinary treatment (LMT) 2) Extensive multidisciplinary	Study Type: CBA alongside an RCT Perspective: Societal	CBA: Net present value of production for the society	CBA: Net present value LMT with screening = NOK 1,842,785 Net present value

	results for long-term sick-listed employees			treatment (EMT) Comparator: Usual Care: Primary care by general practitioner	Time Horizon: 14 months		LMT without screening= NOK 1,300,338 Net present value EMT with screening = NOK 5,489,492 Net present value EMT without screening = NOK 1,655,861 S.A: No
Hlobil et al (2007) The Netherlands	To compare the long-terms costs and benefits of a graded activity intervention for subacute work-related low back pain	Dutch employees on sick-list for a minimum of 4 weeks with	Low back pain	Intervention: Graded activity intervention (GA) Comparator: Usual Care (not defined)	Study Type: CBA alongside a RCT Perspective: Employer Time Horizon: 3 months and 3 year follow - up	CBA: Lost productivity days using HCA	CBA: Net savings of GA first 3 months = €999, per worker Cumulative net savings over 3 years = €1,661 S.A: Yes Type: One-way
Jensen et al (2013) Denmark	To assess the cost-effectiveness of a multidisciplinary versus a brief intervention programme	Danish employed patients aged between 16 and 60 and are on sick-leave from 4 – 12	Low back pain	Intervention: Multidisciplinary intervention Comparator: Brief intervention	Study Type: CBA (modified) and CEA alongside a RCT Perspective: Taxpayer and healthcare sector	CEA: Number of sick leave weeks CBA: Not clear	CEA: ICER = 2,631 kr (€353) per 1 week of sick leave CBA: Not clear S.A: No

		weeks			Study Type: 12 months		
Lambeek et al (2010) The Netherlands	To conduct an economic evaluation of an integrated care programme for patients with chronic low back pain	Patients employed in the Netherlands aged 18 to 65 years for and on sick leave	Low back pain	Intervention: Integrated care programme Comparator: Usual care: referred to occupational care and general practitioner	Study Type: CEA, CUA, and CBA alongside a RCT Perspective: Societal Time Horizon: 12 months	CEA: Return to work CUA: QALY using EQ5D CBA: Productivity costs (benefit) using HCA	CEA: ICER = -£3 per one day earlier return to work CUA: ICER = -£61,000 per QALY CBA: Net societal benefit = £5,744 S.A: Yes Type: One-way
Leon et al (2009) Spain	To evaluate an early cognitive-behavioural therapy programme designed to modify the transition between acute and chronic disability in daily practice	Patients with a musculoskeletal related temporary work disability episode lasting between 4 and 8 weeks	Musculo-skeletal pain	Intervention: Cognitive-behavioural programme Comparator: Usual care: continued rheumatological care program	Study Type: CEA and CBA alongside RCT Perspective: Not stated clearly Time Horizon: 12 months	CBA: Number of sick leave days using HCA	CBA: Net societal benefit = £172,607 S.A: No
Meijer et al (2006) The Netherlands	To determine the cost-effectiveness of the multidisciplinary treatment	Dutch employees, aged 18 to 65	Non-specific Musculoskeletal complaints in the upper extremity	Intervention: Outpatient training programme Comparator: Usual Care: Health care initiated by GP	Study Type: CEA alongside a RCT Perspective: Individual and societal	CEA: Half day per week return to work Decrease on VAS scale for	CEA: ICER = €39 per half-day per week of return to work ICER = €66 per unit decrease on

				including treatment at the workplace	Time Horizon: 12 months	pain Decrease on VAS for other complaints	VAS scale Unit decrease on VAS scale for other complaints: Effectiveness lower than usual care; no cost-effectiveness reported S.A: No
Niemisto et al (2005) Finland	To compare the cost-effectiveness of 2 treatments for patients with low back pain	Finnish employees aged between 24 and 46 years	Low back pain	Intervention: Combined manipulation, exercise and physician consultation Comparator: Physician consultation alone	Study Type: CEA alongside a RCT Perspective: Societal Time Horizon: 2 years	Functional Status (Disability) Pain	CEA: ICER = -$78 per unit improvement in self-reported disability ICER = $512 per unit improvement in pain intensity S.A: No
Phillips et al (2012) UK	To evaluate the cost-effectiveness of an NHS physiotherapy support for occupational health services	Employees in Wales who signed up to the Occupational Health Physiotherapy Project Pilot (OHPPP) service	Musculoskeletal disorders	Intervention: Occupational Health Physiotherapy Project Pilot (OHPPP) service Comparator: Participants acted as own controls, usual care and practice	Study Type: Cohort before and after study with a CUA Perspective: Healthcare sector Time Horizon: 12 months	CUA: QALYs using EQ5D	CUA: ICER = £1386 to £7769 per QALY S.A: Yes Type: One-way

Roelofs et al (2010) The Netherlands	To evaluate the cost-effectiveness of lumbar supports in addition	Dutch homecare workers (employed in a large care home organisation)	Low back pain	Intervention: Lumbar supports and usual care Comparator: Usual care: Short refresher course on health working methods provided by employer; healthcare management for low back pain	Study Type: CEA and CUA alongside a RCT Perspective: Societal Time Horizon: 12 months	CEA: Number of days per month with low back pain Number of calendar days on sick leave CUA: QALY using EQ5D	CEA: ICER = €109 per one less day with low back pain ICER = €51 per day less sick leave CUA: ICER = €25,500 per QALY S.A: Yes Type: One-way
Schweikert et al (2006) Germany	To investigate return to work and cost-effectiveness of the addition of cognitive-behavioural treatment for chronic low back pain	German workforce with low back pain < 6 months	Low back pain	Intervention: Cognitive behavioural treatment (CBT) Comparator: Usual care: Standardised 3 week inpatient rehabilitation program	Study Type: CUA alongside a RCT Perspective: Societal Time Horizon: 6 months	CUA: QALY using EQ5D	CUA: ICER = -€126,731 per QALY S.A: No
Spekle et al (2010) The Netherlands	Evaluate the cost effectiveness and cost-benefit of the RSO QuickScan intervention programme	Dutch computer employees	Shoulder and neck symptoms	Intervention: RSI QuickScan programme Comparator: Usual care: More general, limited advice	Study Type: CEA and CBA alongside a RCT Perspective: Societal and employer Time Horizon:	CEA: A reduction in risk Days of sick leave Prevalence of arm, shoulder	CEA: ICER = -€227.58 per one point reduction in risk ICER = €71.31per increase in sick leave of one day

					12 months	and neck symptoms Work posture and movement	ICER = €72.55 per one point reduction in arm, shoulder and neck symptoms ICER = €93.82 per one point improvement in work posture and movement S.A: Yes Type: One-way
Squires et al (2011) UK	To assess the cost-effectiveness of interventions to return employees to work following long-term sick leave	Patients on sick leave for 1week to 6 months	Musculo-skeletal disorders	Interventions: 1) Workplace intervention based on ergonomics (WE) 2) Physical activity and education (PAE) 3) Physical education and workplace visit (PEW) Comparator: Usual care: general practitioner visits and prescriptions for analgesics	Study Type: Markov model with a CEA and CUA Perspective: Healthcare sector, employer and societal Time Horizon: Lifetime	CEA: Sick leave avoided days CUA: QALYs using SF6D	CEA: ICER = 34 pence per day on sick leave avoided (WE) (Employer perspective) CUA: ICER = £2,800 per QALY (PAE) (Societal and Health perspective) S.A: Yes Type: One-way

Steenstra et al (2006) The Netherlands	To assess whether the workplace intervention, clinical intervention, or both are cost effective	Dutch employees aged between 18 and 65 years on sick leave for 2 to 6 weeks	Low back pain	Intervention: 1) Workplace intervention (WI) 2) Clinical intervention (CI) (graded activity programme) Comparator: Usual care: Dutch Occupational Physician guidelines for low back pain	Study Type: CEA and CUA alongside a RCT Perspective: Societal Time Horizon: 12 months	CEA: Return to work CUA: QALYs using the EQ5D	CEA: WI ICER = €19 per one day's less sick-leave CI: ICER = €29 per one day's less sick-leave WI and CI: ICER = €11 per one day's less sick-leave CUA: WI: ICER = -€1,483 per QALY CI: ICER = €5,447 per QALY WI and CI: ICER = €24,416 per QALY S.A: Yes Type: One-way
Tompa, Dolinschi and Natale (2013) Canada	To conduct an economic evaluation of participatory ergonomics programme	Canadian employees within a clothing manufacturer	Musculoskeletal health problems	Intervention: Participatory ergonomics intervention Comparator: Before the intervention was introduced	Study Type: Cohort before and after study with a CBA Perspective: Employer Time Horizon: 2 years pre-intervention and	CBA: Reduction in: First aid incidents Modified duty episodes Number of causal days absent Long-term sickness	CBA: The net present value $294,827 S.A: Yes Type: One-way

					12 months post-intervention	absences	
Torstensen et al (1998) Norway	To evaluate the efficiency of progressively graded medical exercise therapy, conventional physiotherapy, and self-exercise	Norwegian sick-listed patients between 8 and 52 weeks	Low back pain	Intervention: 1) Medical Exercise Therapy (MET) 2) Conventional Physiotherapy (CP) Comparator: Self-Exercise by Walking (SE)	Study Type: CBA alongside RCT Perspective: Not clear	CBA: Not Clear	CBA: MET costs $122,531 (NOK 906,732) less than SE CP costs $254,200 (NOK 1,882,560) less than SE S.A: No

4.4.2 Critical Appraisal of Economic Evaluations

Overall, the 20 studies included in this review consistently reported the following information clearly: the aim of the study; the population studied; intervention; and time horizon. Two studies did not explicitly report the comparator (Bernaards et al. 2011; Hlobil et al. 2007), and two studies did not report the perspective adopted by the study (Leon et al. 2009; Torstensen et al. 1998).

In eleven studies, CBA methods were reported as the method used to evaluate WPIs. Seven studies were clear and consistent in their reporting and use of CBA methods. Three studies reported they conducted a CBA but seem to have used methods more consistent with cost-analysis (Hlobil et al. 2007; Leon et al. 2009; Torstensen et al. 1998). Hlobil et al., (2007), and Torstensen et al., (1998) seem to have performed a cost-saving analysis where the cost difference between the interventions was reported and Leon et al., (2009) defined the net-benefit as the dollars saved minus dollars invested, rather than explicitly converting benefits into a monetary unit for CBA. The inconsistent reporting standards of CBA studies made it difficult to understand how the analysis was carried out and confused the interpretation of the results. It also limited the comparability of results across the studies.

4.4.3 Inclusion of Productivity Loss

All 20 studies included in this review incorporated lost productivity in terms of absenteeism. However, the rationale for the inclusion of absenteeism was not made explicit in any of the studies. Four studies included estimates of presenteeism (Driessen et al. 2012; Hlobil et al. 2007; Phillips, Phillips Nee Buck, et al. 2012; Tompa, Dolinschi, and Natale 2013) and, of these four studies, one explicitly explained the rationale behind incorporating estimates of presenteeism into the study (Hlobil et al. 2007). Table 4.5 and Table 4.6 summarise how productivity due to absenteeism and presenteeism, respectively, were identified, measured, valued and incorporated into economic evaluations of WPIs-MSDs.

4.4.3.1 Absenteeism

Of the 20 studies that estimated absenteeism: eleven valued absenteeism as a cost (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Hlobil et al. 2007; Jensen et al. 2013; Meijer et al. 2006; Niemistö et al. 2005; Roelofs et al. 2010; Schweikert et al. 2006; Squires et al. 2012; Steenstra et al. 2006; Torstensen et al. 1998); seven studies valued absenteeism as a monetary benefit (Bültmann et al. 2009; Driessen et al. 2012;

Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Leon et al. 2009; Speklé et al. 2010; Tompa, Dolinschi, and Natale 2013); and one as a non-monetary benefit (Phillips et al. 2012). One study valued absenteeism as a non-monetary benefit in a CEA, then as a cost in the CUA, and then as a monetary benefit in the CBA (Lambeek et al. 2010). Seventeen studies considered the impact of absenteeism from paid work only (Bernaards et al. 2011; Bültmann et al. 2009; Driessen et al. 2012; Grahn, Borgquist, and Ekdahl 2004; Hagen et al. 2003; Haldorsen et al. 2002; Hlobil et al. 2007; Jensen et al. 2013; Lambeek et al. 2010; Leon et al. 2009; Niemistö et al. 2005; Phillips et al. 2012; Schweikert et al. 2006; Speklé et al. 2010; Squires et al. 2012; Tompa, Dolinschi, and Natale 2013; Torstensen et al. 1998). Three studies, all conducted in the Netherlands, included absenteeism from paid and unpaid work (Meijer et al. 2006; Roelofs et al. 2010; Steenstra et al. 2006).

Table 4.4 Reported inclusion of absenteeism in economic evaluations of WPIs

Author (Year) Country	**Type of EE**	**Perspective**	**Identify and Measurement Method**	**Cost**[2]	**Benefit**[3] (Monetary or Non-Monetary)	**Valuation Method**	**Absenteeism Incorporated in: baseline or Sensitivity analysis (S.A)**
Bernaards et al (2011) The Netherlands	CEA	Employer	Number of absence days Data source: Company records	Yes	No	HCA: Mean income of Dutch population by age and gender Data source: Dutch Healthcare Insurance Board: Costing manual Elasticity rate of 0.8 (production losses estimated to be 80% in workers who are 100% absent) Elasticity of 0.3 (production losses of 30% associated with attending the intervention during work hours)	CEA: Total cost S.A: 1) Total costs using elasticity values of 0.3, 0.55, and 1.0 2) Analysis of those with complete cost and effect data
Bültmann et al (2009) Denmark	CEA CBA	Societal	Sickness absence hours Data source: Danish National Insurance	No	Yes (Monetary)	HCA: Gross wages per hour for six different employment groups (including employment overheads and benefits)	CEA: Effect CBA: Total monetary benefit S.A: Wages reduced by

[2] Valued and treated as a cost in the economic evaluation

[3] Valued and treated as a monetary or non-monetary benefit in the economic evaluation

			Service			Data source: Danish Statistics Bank	25%
Driessen et al (2012) The Netherlands	CEA CBA	Societal Employer	Number of work hours on sick leave Data source: Single-item question about all cause sick leave in past 3 months	Yes	Yes (Monetary)	FCA: Average hourly wage rate based on age and gender Data source: Dutch Healthcare Insurance Board: Costing manual Friction period of 154 days Elasticity of 0.8	CEA: Not clear CBA: Total monetary benefits
Grahn et al (2004) Sweden	CEA	Societal	Number of sick days Data Source: Social Insurance Office of Sweden	Yes	No	HCA: Income including employers' costs per lost working day) Data Source: Not clear	CEA: Total Cost
Hagen et al (2003) Norway	CBA	Societal	Number of sick days Data source: National Insurance Registers	No	Yes (Monetary)	HCA: Gross wage payments including employment taxes Data Source: Not clear	CBA: Monetary benefit as the social value of production

Haldorson et al (2002) Norway	CBA	Societal	Number of sick months Data source: Social insurance register	No	Yes (Monetary)	HCA: Gross wage payments including employment taxes Data Source: Observed earnings (Not clear)	CBA: Monetary benefit as the social value of production
Hlobil et al (2007) The Netherlands	CBA	Employer	Number of calendar days workers completely or partially sick-listed (gross lost productive days) Partial lost productive days counted for its % work absence (net lost productive days) Data source: Electronic database of the	No	Yes (Monetary)	HCA: Mean daily wage inflated by 80% to account for secondary benefits Data Source: Not clear	CBA: Total monetary benefit

			occupational services				
Jensen et al (2013) Denmark	CEA CBA	Healthcare Taxpayer	Number of sick leave days Data source: Danish Register for Evaluation of Marginalisation register (DREAM) register of transfer	Yes	No	HCA: Compensation payment from society per week Data Source: DREAM register of transfer income	CEA: Not included CBA: Total cost
Lambeek et al (2010) The Netherlands	CEA CUA CBA	Societal	Number of days on sick leave Data source: Patient recorded absence data within 12 months. Pre-12 months, absence data collected using occupational services	Yes	Yes (Monetary)	HCA: Estimated price of production loss of a worker per day based on age and gender Data Source: Not clear	CEA: No CUA: Total costs CBA: Total monetary benefit
Leon et al (2009) Spain	CBA	Not Clear	Number of sick leave days Data source:	No	Yes (Monetary)	HCA: Average daily wage Data Source: Not clear	CBA: Total monetary benefit

			Recorded by rheumatologist				
Meijer et al (2006) The Netherlands	CEA	Individual Societal	Paid work: Number of sick leave hours Data source: Participants answered 4 productivity related questions Unpaid work: Extra hour needed for domestic activities Data source: Participants answered 4 productivity related questions	Yes	Yes (Monetary)	HCA: Gross hourly wage level Data Source: Dutch Healthcare Insurance Board: Costing manual Unpaid work: Net hourly wage Data Source: Dutch Healthcare Insurance Board: Costing manual	CEA: Effect and total cost
Niemisto et al (2005) Finland	CEA	Societal	Number of days absent from work (half and full days) Data source: Questionnaire to participants	Yes	No	Method not clear. Used average 2000-year wage level Used a 50% lower estimate of average wage level Data Source: Not clear	CEA: Total cost
Phillips et al (2012) UK	CUA	Healthcare	Days Sickness Absence	No	No	Multivariate regression analysis	CUA: No

			Data source: Not clear			Data Source: Not required	
Roelofs et al (2010) The Netherlands	CEA CUA	Societal	Paid: Number of days on sick leave Data source: Organisation register Unpaid: Number of hours absent Data source: Not clear	Yes	Yes (Non-monetary)	HCA: Corresponding salary scale in the collective labour agreement of the home care organisation Data Source: Company Records Unpaid: Shadow price €8.40 per hour Data Source: Dutch Healthcare Insurance Board: Costing manual	CEA: Effect and total cost CUA: Total cost
Schweikert et al (2006) Germany	CUA	Societal	Number of days off work Data source: Sickness insurance records	Yes	No	HCA: Age and gender adjusted labour costs including salaries and social insurance premiums paid by employer and employee Data Source: Not clear	CUA: Total cost
Spekle et al (2010) The Netherlands	CEA CBA	Societal Employer	Number of sick leave hours Data source: Company records	No	Yes (Monetary)	FCA: Mean income of the Dutch working population according to age and gender Data Source: Dutch Healthcare Insurance Board: Costing manual	CEA: No CBA: Total monetary benefit

						Friction period 154 days Elasticity of 0.8	
Squires et al (2011) UK	CEA CUA	Societal Healthcare Employer	Number of sick leave days Data source: Literature review	Yes	No	FCA: National average wage Friction period: 10 weeks based on national average replacement period Data Source: Department for Work and Pensions (Government)	CEA: Total cost CUA: Total cost CUA: No S.A: Used HCA
Steenstra et al (2006) The Netherlands	CEA CUA	Societal	Number of sick leave days Partial return to work: Worker performing at 50% per calendar day Data source: Occupational Health Service Unpaid: Number of hours Data Source: Not clear	Yes	No	FCA: Estimated price of production loss per worker per day based on age and gender Friction period: 122 days Data Source: Dutch Healthcare Insurance Board: Costing manual Unpaid work: Fee of €8.93 per hour Data Source: Dutch Healthcare Insurance Board: Costing manual	CEA: No CUA: Total cost S.A: Three variations 1) Used mean cost of production loss for entire population 2) Used number of calendar days off sick 3)Used HCA

Tompa, Dolinschi and Natale (2013) Canada	CBA	Employer	Number of days casual absenteeism Long-term sickness absence (number of absences and duration) Data source: Human Resource (company) records	No	Yes (Monetary)	Not Clear Data Source: Not Clear	CBA: Total monetary benefit
Torstensen et al (1998) Norway	CBA	Not Clear	Not stated	Yes	No	Not Clear Data Source: Social insurance offices	CBA: Total cost

4.4.3.2 Summary of methods used to measure absenteeism

Of the 20 studies that reported the inclusion of estimates of absenteeism from paid work, fourteen identified and measured absenteeism as the number of days away from work (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Hagen et al. 2003; Hlobil et al. 2007; Jensen et al. 2013; Lambeek et al. 2010; Leon et al. 2009; Niemistö et al. 2005; Phillips et al. 2012; Roelofs et al. 2010; Schweikert et al. 2006; Squires et al. 2012; Steenstra et al. 2006; Tompa, Dolinschi, and Natale 2013). Four studies identified and measured the number of hours on sick-leave (Bültmann et al. 2009; Driessen et al. 2012; Meijer et al. 2006; Speklé et al. 2010) and one study identified and measured absenteeism as the number of sick months (Haldorsen et al. 2002). One study did not explicitly report the units used to measure absenteeism (Torstensen et al. 1998).

The data sources used to identify and measure absenteeism included: social insurance records (n=8) (Bültmann et al. 2009; Grahn, Borgquist, and Ekdahl 2004; Haldorsen et al. 2002; Hagen, Grasdal, and Eriksen 2003; Jensen et al. 2013; Roelofs et al. 2010; Schweikert et al. 2006; Torstensen et al. 1998); company records (n=5) (Bernaards et al. 2011; Hlobil et al. 2007; Speklé et al. 2010; Steenstra et al. 2006; Tompa, Dolinschi, and Natale 2013); and data collected within the study (n=3) from participants (Driessen et al. 2012; Meijer et al. 2006; Niemistö et al. 2005) and professionals, in this case a rheumatologist (Leon et al. 2009). One study collected the number of sick days away from work using two data sources: 1) self-reported counts from participants within the study; and 2) sickness records from occupational services (Lambeek et al. 2010). One study did not report the data source used to measure absenteeism (Phillips et al. 2012).

Of the three studies that measured absenteeism from unpaid work: two identified and measured absenteeism as the number of hours due to sickness (Roelofs et al. 2010; Steenstra et al. 2006); and one study collected extra hours needed for domestic activities by asking the participants within the study (Meijer et al. 2006). Two studies by Roelofs et al. (2010) and Steenstra et al. (2006) did not clearly report how they collected absence data related to unpaid work.

4.4.3.3 Valuing absenteeism as a cost

Thirteen studies reported that they valued absenteeism as an input into the cost element of the economic evaluation (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Jensen et al. 2013; Lambeek et al. 2010; Hlobil et al. 2007; Leon et al. 2009; Meijer et al. 2006; Niemistö et al. 2005; Roelofs et al. 2010; Schweikert et al. 2006; Squires et al.

2012; Steenstra et al. 2006; Torstensen et al. 1998). Nine studies reported using the HCA (see chapter two, section 2.3.5) (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Hlobil et al. 2007; Jensen et al. 2013; Lambeek et al. 2010; Leon et al. 2009; Meijer et al. 2006; Roelofs et al. 2010; Schweikert et al. 2006). The nine studies that applied the HCA used similar time units: eight studies used number of days on sick leave (Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004; Hlobil et al. 2007; Jensen et al. 2013; Lambeek et al. 2010; Leon et al. 2009; Roelofs et al. 2010; Schweikert et al. 2006); and one used number of sickness hours (Meijer et al. 2006). One study did not report clearly report the methods used to value absenteeism as a cost (Torstensen et al. 1998).

The unit price (wage) used to value absenteeism as a cost differed across the ten studies that used the HCA. Three studies used daily wage rates (Leon et al. 2009; Hlobil et al. 2007; Jensen et al. 2009), and the remaining seven studies used: mean income by age and gender (Bernaards et al. 2011); employee income including the employers' cost per lost working day (Grahn, Borgquist, and Ekdahl 2004); compensation payments by society (Jensen et al. 2013); price per day based on age and gender (Lambeek et al. 2010); gross hourly wage rate (Meijer et al. 2006); salary scale (Roelofs et al. 2010); and labour costs adjusted for age and gender including salaries and social insurance premiums (Schweikert et al. 2006). Bernaards et al., (2011) also applied an elasticity rate of 0.8 which assumed full productivity was equal to 80% and not 100%, implying that production losses were 80% in employees with 100% sickness absence.

Two studies used the friction cost approach (FCA) to value absenteeism as a cost (Squires et al. 2012; Steenstra et al. 2006). Both studies used different "friction periods", including 122 days (Steenstra et al. 2006) and 10 weeks (Squires et al. 2012), and used different wage units to value the costs: one study used national average wages (Squires et al. 2012); and one study used the price of lost production based on gender and age (Steenstra et al. 2006).

Two studies did not clearly report the methods used to value absenteeism as a cost (Niemistö et al. 2005; Torstensen et al. 1998).

4.4.3.4 Valuing absenteeism as a monetary benefit

Ten studies valued absenteeism as a benefit that was measured in monetary terms (monetary benefit) (Bültmann et al. 2009; Driessen et al. 2012; Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Lambeek et al. 2010; Philips et al. 2004; Speklé et al. 2010; Tompa, Dolinschi, and Natale 2013; Leon et al. 2009; Hlobil et al. 2007). The most common method used to value the impact of absenteeism as a monetary benefit was the HCA (n=5) (Lambeek et al. 2010; Hagen, Grasdal, and Eriksen 2003; Leon et al. 2009; Hlobil et al. 2007; Haldorsen et al. 2002). Of the five studies, four used sick days off (Lambeek et al. 2010; Hagen, Grasdal, and Eriksen 2003; Leon et al. 2009; Hlobil et al. 2007) and one study used the number of calendar months returned to work (Haldorsen et al. 2002). All five studies that valued absenteeism in monetary terms using the HCA applied different price conversions: one study used wages based on age and gender (Lambeek et al. 2010); one study used gross wages including taxes (Hagen, Grasdal, and Eriksen 2003); one study used the average daily wage (Leon et al. 2009); one study used observed earnings two years after the employee returned to work (Haldorsen et al. 2002); and one study used mean daily wage inflated by 80% to account for secondary benefits (Hlobil et al. 2007). Speklé et al., (2010) also applied an elasticity rate of 0.8.

The FCA was used in two studies and both of which used a friction period of 154 days (Driessen et al. 2012; Speklé et al. 2010). The volume of lost productivity was valued using average wages adjusted for age and gender. The remaining study that valued absenteeism as a monetary benefit did so by applying regression techniques (Tompa, Dolinschi, and Natale 2013). Regression techniques were used to investigate the impact WPIs had on a number of outcome measures including number of days of casual absenteeism. The study was unclear in its reporting about how the number of days of casual absenteeism was converted into monetary-benefit units.

The data sources used to obtain unit costs (wage levels) to estimate the value of productivity in monetary terms from paid work were primarily obtained from national statistics (n=7) (Bernaards et al. 2011; Bültmann et al. 2009; Meijer et al. 2006; Roelofs et al. 2010; Steenstra et al. 2006; Speklé et al. 2010; Squires et al. 2012). Of the nine studies conducted in the Netherlands, six studies obtained wage estimates using the same data source: Oostenbrink, Koopmanschap, and Rutten (2002) (Bernaards et al. 2011; Bültmann et al. 2009; Driessen et al. 2012; Meijer et al. 2006; Speklé et al. 2010; Steenstra et al. 2006).Two studies used company records (Jensen et al. 2013; Roelofs et

al. 2010). Eleven studies were unclear about the data source used to obtain the wage rates used to estimate the value of productivity loss from paid work (Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Hlobil et al. 2007; Lambeek et al. 2010; Leon et al. 2009; Schweikert et al. 2006; Tompa, Dolinschi, and Natale 2013; De Bruijn et al. 2007; Grahn, Borgquist, and Ekdahl 2004; Torstensen et al. 1998; Niemistö et al. 2005). All four studies that valued productivity loss from unpaid work used the Dutch Costing Manual (Oostenbrink, Koopmanschap, and Rutten 2002) (Steenstra et al. 2006; Roelofs et al. 2010; Meijer et al. 2006; De Bruijn et al. 2007).

4.4.3.5 Valuing absenteeism as a non-monetary benefit

Two studies valued absenteeism as a non-monetary benefit (Roelofs et al. 2010; Phillips et al. 2012). Roelofs valued absenteeism as the number of days on sick leave during the 12 month intervention period. Phillips et al., (2012) sought to establish an association between health status and productivity. Health status was measured using the EQ5D (EuroQol Group 1990b), Short Form-12 (SF12) (Jenkinson and Layte 1997), and the General Health Questionnaire (GHQ) (Jackson 2007) and productivity, included as absenteeism, was measured using the number of days away from work. The associations between health status and absenteeism were examined using multivariate analysis. The results showed a statically significant relationship between the SF-12 and absenteeism. However, a non-statistically significant relationship was identified between the EQ5D and GHQ, respectively. The methods used to assess the relationship between health status and productivity was poorly reported and it was unclear how absenteeism was measured.

4.4.4 Presenteeism

Four studies have included presenteeism in their economic evaluations of WPI-MSDs (Driessen et al. 2012; Hlobil et al. 2007; Phillips et al. 2012; Tompa, Dolinschi, and Natale 2013). Table 4.6 summarises how each study identified, measured, and valued, and incorporated presenteeism into the analyses conducted.

4.4.4.1 Summary of methods used to measure presenteeism

Of the four studies that measured presenteeism (Hlobil et al. 2007; Philips et al. 2004; Tompa, Dolinschi, and Natale 2013; Driessen et al. 2012), three studies measured presenteeism in terms of the level of work performance (Driessen et al. 2012; Hlobil et al. 2007; Phillips et al. 2012) and one study measured the number of outputs produced correctly (Tompa, Dolinschi, and Natale 2013). Driessen et al. (2012) asked participants

to rate their work performance level using a scale from 0 (worst performance) to 10 (best performance) taken from the Health Productivity Questionnaire (HPQ) (Kessler et al. 2003). Hlobil et al., (2007) measured work performance as the percentage decrease (25 or 50%) in productivity. The percentage decrease in productivity reflected the possibility that workers, who have partially recovered or work in a different role as a result of ill-health, performed at a lower level than usual. Philips et al., (2012) measured work performance by asking the participants to think back over the past 6 months and to state whether they worked: 1) usual hours and duties; 2) usual hours but not usual duties; 3) usual duties but not usual hours; and 4) usual hours but help needed. The study by Tompa, Dolinschi, and Natale (2013) conducted a study in a clothes manufacturing company and used two objective measures of presenteeism: 1) percentage of shirts produced correctly, without mistakes or needing to re-do work; and 2) percentage efficiency. Percentage efficiency was measured as a the ratio of the actual number of shirts produced to the target number shirts that should be produced in one week (Tompa, Dolinschi, and Natale 2013). The authors of the study did not clearly report how the data used to calculate presenteeism were collected (Tompa, Dolinschi, and Natale 2013).

Table 4.5 Reported inclusion of presenteeism in economic evaluations of WPIs

Author (Year) Country	**Type of EE**	**Perspective**	**Identify and Measurement method**	**Cost**[1]	**Benefit**[2] (Monetary or Non-Monetary)	**Valuation Method**	**Incorporated in: baseline or Sensitivity analysis (S.A.)**
Driessen et al (2012) The Netherlands	CEA CBA	Societal and Employer	Work Performance: WHO-HPQ % performance (0% worst and 10% best) Data source: Questionnaire to participants	No	Yes Non-monetary	Not appropriate	CEA: Yes How: Unit of effect CBA: No
Hlobil et al (2007) The Netherlands	CBA	Employer	Work performance: 25-50% decrease in at work productivity Data source: Database of occupational services	No	Yes Monetary	Paid: HCA: Mean daily wage inflated by 80% to account for secondary benefits Data Source: Not clear	CBA: No S.A: Decrease in productivity (work performance) by 25 and 50% treated as monetary benefit
Phillips et al (2012) UK	CUA	Healthcare	Work performance: Unclear Data source: Not clear	No	Yes Non-monetary	Not required Data Source: Not Required	CUA: No
Tompa, Dolinschi and Natale 2013 Canada	CEA CBA	Employer	% products made first time (no mistakes) % efficiency Data source: Not clear	No	Yes Monetary	Paid: Regression methods Data Source: Not clear	CEA: Not clear CBA: Yes How: Total monetary benefit

4.4.4.2 Valuing presenteeism

All four studies valued presenteeism as an input into the consequences (benefits) element of the economic evaluation. Two studies valued presenteeism as a monetary benefit: one study used the HCA (Hlobil et al. 2007); and the other study did not report the methods clearly (Tompa, Dolinschi, and Natale 2013). The data sources used to value presenteeism as a monetary benefit were not reported clearly (Hlobil et al. 2007; Tompa, Dolinschi, and Natale 2013). Driessen et al., (2012) and Phillips et al., (2012) valued presenteeism as a non-monetary benefit. Driessen et al., (2012) used percentage levels of work performance as the unit of effect in their CEA and Philips et al., (2012) used multivariate regression methods to test the association between work performance and health, as measured by the EQ5D, SF-12, and GHQ. Philips et al., (2012) found that physical health status was an important factor in determining an individual's capacity to work. However, the study did not find a statistically significant association between presenteeism and health status (EQ5D) and the methods used to assess the association between the EQ5D and presenteeism were poorly reported.

4.4.4.3 Reporting Quality of Productivity in Economic Evaluations of WPI-MSD

Ten studies described productivity as a cost and included it as a cost in the final analysis (Bernaards et al. 2011; Hlobil et al. 2007; Niemistö et al. 2005; Roelofs et al. 2010; Schweikert et al. 2006; Squires et al. 2012; Steenstra et al. 2006; Meijer et al. 2006; Grahn, Borgquist, and Ekdahl 2004; Torstensen et al. 1998). Four studies described lost productivity as a (dis)benefit and treated it as a (dis)benefit throughout the analysis (Hagen, Grasdal, and Eriksen 2003; Haldorsen et al. 2002; Tompa, Dolinschi, and Natale 2013; Philips et al. 2004). Two studies that used more than one type of economic evaluation analysis explicitly and clearly explained how lost productivity was treated (as a cost and then monetary benefit) and incorporated in each analyses (Lambeek et al. 2010; Speklé et al. 2010). Lambeek et al., (2010) explained that they had included lost productivity as an effect (non-monetary benefit) in their CEA, as a cost in their CUA, and as a monetary benefit in their CBA. Speklé et al., (2010) reported lost productivity as the unit of effect in their CEA and as a monetary benefit in their CBA. Four studies were inconsistent and described lost productivity as cost (benefit) and treated it as a (dis)benefit (cost) in the final analysis (Bültmann et al. 2009; Driessen et al. 2012; Jensen et al. 2013; Leon et al. 2009).

4.4.4.4 Double-Counting

Two CEA studies may have potentially double-counted productivity in their economic evaluations (Driessen et al. 2012; Meijer et al. 2006). Driessen et al., (2012) reported work performance as one of the outcomes used in their CEA ICERs. The total costs used included the cost of health care, intervention, and lost productivity. Lost productivity costs were measured using the number of days on sick leave and valued using the FCA. Potentially, double-counting of productivity may have occurred in this part of the analysis where presenteeism (work performance) was included as an outcome in the numerator of the ICER and absenteeism was included as a cost in the denominator. However, the authors did not include lost productivity costs when estimating the CEA ICER for all cause sick-leave; only healthcare and intervention costs were included. Meijer et al., (2006) used an extra half-day return to work as the measure of effect but also seemed to have included the impact of lost productivity, calculated as the number of sick leave hours multiplied by the gross hourly wage, in their total cost.

Six studies excluded productivity costs from their CEA (Jensen et al. 2013; Lambeek et al. 2010; Roelofs et al. 2010; Speklé et al. 2010; Steenstra et al. 2006; Bültmann et al. 2009) and four studies explained the reason was to avoid double-counting (Jensen et al. 2013; Lambeek et al. 2010; Speklé et al. 2010; Steenstra et al. 2006). The remaining CEA studies did not need to consider the potential for double-counting because they used effects measures that did not involve productivity (Niemistö et al. 2005; Bernaards et al. 2011; Grahn, Borgquist, and Ekdahl 2004).

At present, there are no published reporting criteria to inform best practice of the rationale, estimation, incorporation, and reporting of absenteeism and presenteeism in economic evaluations of WPIs. Standardising the reporting of productivity in economic evaluations would allow problems and issues to be highlighted providing a clear indication of where further research is needed. Reporting criteria would also potentially help researchers to avoid common mistakes, such as double-counting. By improving the reporting standards of productivity in economic evaluations of WPIs decision-makers, in this case the employer would have a clearer understanding of the impact WPIs have on the costs or benefits of absenteeism and/or presenteeism. As a starting point some additional reporting criteria is listed in table 4.7. The suggested reporting criteria may be used alongside existing reporting standards, such as; CHEERS (Husereau et al.

2013). The criteria promotes consistency and transparency of the methods used to identify, measure, value, and incorporate productivity in economic evaluations. However, further empirical work, using accepted methods as proposed by Moher and colleagues, is needed to confirm that these suggested reporting criteria are sufficient (Moher et al. 2010).

Table 4.6 Potential reporting criteria for productivity in economic evaluations

Criteria	Description
Rationale for inclusion	State the reasons why productivity is being included in the economic evaluation
Type of productivity included and the definition used	What type of productivity is being included in the evaluation? For example, absenteeism, presenteeism or both. Clearly state the definition of absenteeism and presenteeism used in the study
Productivity included as a cost or benefit	How is productivity categorised? Cost, *or* monetary benefit, *or* non-monetary benefit (Do not include productivity as a cost *and* benefit in the same analysis; this will lead to double-counting)
Methods to identify and measure selected types of productivity and state the data collection methods and sources used	State the methods used to identify and measure absenteeism and/or presenteeism For example, absenteeism was identified and measured as the number of days away from work Presenteeism was identified and measured using a self-report survey (name the survey) Data collection methods: Describe how the data was obtained. For example, through national databases or through self-report surveys using a study sample
Methods to value selected types of productivity and state the data collection methods and sources used	State the methods used to value absenteeism or presenteeism as a cost, monetary benefit, or non-monetary benefit If valued as a cost or monetary benefit state the method used, for example the HCA or FCA. Also, state the price, for example wage levels, used to value productivity Report the data sources used to obtain the price If valued as a non-monetary benefit, state the unit of effect used, for example number of days on sick leave or work performance as a %. State the data sources or collection methods used to obtain the non-monetary benefit used
Justify the selection of methods used to identify, measure, and value productivity	Explain why each method, used to identify, measure, and value absenteeism or presenteeism, was selected
How is productivity is included in the analysis	State how productivity is included in the analysis. For example, as a cost in the total cost, a monetary (non-monetary) benefit in the total monetary benefit, or as a cost or benefit (monetary or non-monetary) in the sensitivity analysis

4.5 Discussion

This systematic review confirmed productivity, in the form of absenteeism, was considered to be a key component in published economic evaluations of WPI-MSDs. Absenteeism was included in all 20 identified economic evaluations of WPI-MSDs; presenteeism, however, was included in four studies only. Analysts consistently failed to explicitly describe the rationale for quantifying and including the impact of absenteeism in their analyses. On the four occasions where presenteeism was included, only one study provided a clear rationale to support the need to quantify the impact of presenteeism (Hlobil et al. 2007). It seems paradoxical that presenteeism was rarely considered in economic evaluations of WPI-MSDs given the consensus that the potential impact of presenteeism is greater than that caused by absenteeism (Hemp 2004).

There are at least two potential explanations for the limited number of economic evaluations of WPI-MSDs that included the impact of presenteeism. First, the concept of presenteeism, compared to absenteeism which has been widely understood and clearly defined for some time, is still fairly new with no consensus regarding its definition (Johns 2010). Doubts over the definition of presenteeism may have caused analysts to be wary of including it in their analyses. Second, this systematic review confirms that the methods used to quantify presenteeism varied to a greater extent than those used to quantify the impact of absenteeism; a finding that is in line with results in previously published systematic reviews (Kigozi et al. 2017; Mattke et al. 2007). Kigozi et al., (2017), in particular, suggested, and the authors of this systematic review agree, that there was a general lack of confidence and, potentially, confusion regarding the selection of a method to be used to quantify the impact of presenteeism suitable for their economic evaluation.

Guidelines developed to inform researchers of the approach and methods they should use when conducting and reporting economic evaluations of pharmaceutical based healthcare interventions differ across global healthcare systems (ISPOR 2017). All countries represented this systematic review, with the exception of England and Wales, have guidelines that allow for the inclusion of the impact of productivity in economic evaluations of pharmaceutical interventions (see ISPOR 2017). Guidelines developed by NICE in England and Wales explicitly reject the inclusion of productivity in the assessment of health care technologies (HTA) and diagnostics (NICE 2013). The

recommendation is justified on the premise that the National Health Service (NHS), in England and Wales, is funded from a fixed annual budget; therefore, economic evaluations of healthcare technologies are to be performed from the perspective of the healthcare sector meaning that only those costs and benefits that fall on the healthcare system should be included (Claxton et al. 2010). At present, relevant guidelines for the design and conduct of economic evaluations do not exist when alternative budgets to healthcare are being used to invest in health promoting intervention; for example employers funding WPIs. With that said, NICE public health guidelines acknowledge the limitations of adopting a healthcare perspective and suggest adopting a sufficiently wide perspective to allow for the inclusion of non-health outcomes (NICE 2012b).

Consistent with the NICE methods guideline for the appraisal of technology appraisals and diagnostics (NICE 2014), the two UK based studies adopted healthcare as the primary perspective of their in the base case analysis and did not include the impact of productivity (Phillips et al. 2012; Squires et al. 2012). Nevertheless, both studies conducted additional analyses that allowed them to assess the impact WPI-MSDs had on productivity. Squires et al., (2012) conducted two further analyses, one from the perspective of society and the other form the perspective of employer. The impact on productivity (absenteeism) was included as a cost in the total costs of both analyses. Phillips et al., (2012) incorporated the impact on productivity by exploring the relationship between productivity and health status as a separate investigation alongside the main cost-effectiveness analysis (Phillips et al. 2012). However, given that the primary aim of the study was to assess the cost-effectiveness of an occupational health service, Phillips et al., (2012) reported limited information about the methods used to analyse the relationship between health status and productivity. Two further studies have explored the association between health status and productivity (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014) (see chapter six, section 6.1) the results indicate that there is a quantifiable relationship but more research is needed to confirm the extent to which there is a relationship. Further investigation of the potential relationship between health status and productivity (absenteeism and presenteeism) would be useful for the development of methods that can be used to quantify the impact of ill-health, in this case MSDs, on productivity (see chapter six). Presenting separate analyses is a useful pragmatic approach to take when evaluating the economic impact of WPIs as it seems illogical not to investigate the impact WPIs have on productivity (ability to work). However, there is a lack of consistency of the techniques used to incorporate and

quantify the impact of productivity in economic evaluations of WPI-MSDs contributing further to the problem of incomparability of results across studies.

4.5.1 Strengths

This systematic review collated all evidence for the inclusion of absenteeism and presenteeism in economic evaluations of WPI-MSDs. The study has contributed to the literature in five novel ways. This is the first systematic review that has critically appraised the methods used to identify, measure, and value absenteeism and presenteeism in economic evaluations of WPI-MSDs conducted from all potential perspectives (societal, healthcare, employer, patient). This review adds to the literature by comparing the difference in methods used to quantify absenteeism and presenteeism in current practice and the extent to which both concepts were included in economic evaluations of WPIs. The dates of this systematic review overlap with Kigozi et al., (2017) review, but this study has identified four additional studies that incorporate presenteeism as a benefit providing evidence that the literature considers presenteeism as benefit as well as a cost. This review has clearly and explicitly investigated the extent to which absenteeism and presenteeism are included as a cost, monetary benefit, and non-monetary benefit in economic evaluations of WPIs. The analysis conducted goes beyond labelling productivity as a cost or saving by reporting how productivity is incorporated (or not) into the final total cost as a cost, or total benefit as a monetary benefit, or non-monetary benefit. A critical appraisal of the reporting quality of productivity in economic evaluations of WPIs identified the need for a published set of standardised reporting guidelines, especially concerning the categorisation and inclusion of absenteeism and presenteeism as a cost, monetary benefit, or non-monetary benefit. This review has identified double-counting of productivity as a potential issue in economic evaluations of WPIs. It was identified that the potential for double-counting productivity was considered in some economic evaluations, but not all which may have led some to have counted the impact of productivity as a cost and benefit. This review has provided a set of potential reporting criteria that may be useful to aid the standardisation of estimating, incorporating and reporting of productivity in economic evaluations.

4.5.2 Limitations

There are two limitations in this systematic review. This systematic review was designed to identify economic evaluations of WPIs for individuals with MSDs. The definition of WPI used was very specific, potentially meaning that some relevant studies

that used broader definitions may have been missed. However, due to the nature of the aim of this systematic review it was appropriate to adopt the specific definition of WPI. The review by Hillage et al (2008) sought to identify all economic evaluations of workplace interventions used HEED and NHS EED databases. Due to the closure of these databases, it was only feasible to update the searches in Medline and Embase and there may a small possibility that some studies were not identified. To ensure no studies were missed, references of all identified studies were screened for potential relevant studies.

4.6 Conclusion

This systematic review identified lost productivity as a key factor included in published economic evaluations of WPI-MSDs. Absenteeism was included in all identified studies; however, presenteeism was included in only four. Variation in the methods used to identify, measure, and value the impact of absenteeism and presenteeism was identified as a potential challenge for the interpretation of the results of economic evaluations of WPI-MSDs. A simple first step would be to improve the standardisation of reporting of why, and how, the impact on productivity (absenteeism and/or presenteeism) is included in economic evaluations. Further development of standardised reporting criteria for the identification, measurement, and valuation of absenteeism and presenteeism in economic evaluations may help reduce variation in the methods and avoid mistakes such as double-counting. There is also a need to further develop methods that identify, measure, and value presenteeism. A potential starting point would be to further explore the associated relationship between health and productivity (absenteeism and presenteeism) which may inform the design of alternative methods to quantify the impact MSDs have on the ability to work.

Chapter 5 Exploring the impact of health status and capability of employees with inflammatory arthritis on presenteeism in the workplace: A Qualitative Study

5.0 Introduction

This chapter describes a study designed to explore the extent to which selected existing measures of health status and capability (collectively termed measures of health-related quality of life (HRQoL) are potentially able to capture the impact of rheumatoid arthritis (RA), ankylosing spondylitis (AS) and psoriatic arthritis (PsA) (see chapter one, section 1.6) have on the ability to work (presenteeism). This study used a qualitative approach to complement the quantitative study presented in Chapter Six. Qualitative research methods were used to understand how employees with RA, AS or PsA are affected in their ability to work effectively and how this aligns with the face validity of available selected measures of HRQoL, appropriate for use in economic evaluations, to capture the impact of presenteeism.

Section 5.1 explains the rationale for this study and why it is important to test selected measures of HRQoL for face validity. Section 5.2 presents the aims and objectives of the study. Section 5.3 describes the study sample, the interview schedule, and the qualitative methods used to collect and analyse the data. Section 5.4 describes the results of the qualitative analysis and sections 5.5 and 5.6 present a discussion and conclusion, respectively.

5.1 Background

Chapter 4 identified a study (Philips et al., 2012) which sought to quantify the relationship between measures of HRQoL and productivity (absenteeism and presenteeism). Philips et al. assessed the association between productivity and various health measures including the EQ5D, the general health questionnaire (GHQ), and the Short-Form 12 (SF-12) which prompted the rationale for this chapter; can alternative measures of HRQoL be used to capture an individual's ability to work given the presence of a health condition?

Evidence, although limited, suggests there is a relationship between HRQoL and presenteeism (Krol, Stolk, and Brouwer 2014; Lamers et al. 2005) (see chapter six, section 6.1); however, the literature has focused its attention solely on the association

between productivity and the EQ5D. Also, no empirical research has been conducted to explore the face validity of HRQoL measures, including the EQ5D, and their capacity to capture those aspects of a health condition that limit an individual's ability to work.

Psychometric properties are used to assess the extent to which instruments designed to measure HRQoL are reliable and valid (Ginty 2013). Chapter three defined and described the different types of psychometric properties that can be tested (see chapter three, section 3.4.6). This chapter focuses on the face validity, also called content validity, of selected measures of HRQoL. Face validity is fundamentally important because it establishes the extent to which an instrument measures the construct of interest.

Therefore, as a useful first step, qualitative methods were used to begin to explore the potential content validity of three selected measures of HRQoL: the EuroQol Five Dimensions Five Level (EQ5D-5L) (Herdman et al. 2011); the Short-Form Six Dimensions (SF6D) (Brazier, Roberts, and Deverill 2002); and the ICEpop CAPability measure for Adults (ICECAP-A) (Al-Janabi et al. 2013), and their ability to capture the impact on presenteeism caused by RA, AS, and PsA (see chapter one, section 1.4).

5.2 Aims and Objectives

The aim of this study was to explore the extent to which selected measures of health status and capability captured aspects of RA, AS or PsA that impact on levels of presenteeism.

To meet this aim, three objectives were addressed:

1. To explore how employees with RA, AS or PsA may be affected in their ability to work;
2. To compare the qualitative responses from employed individuals with RA, AS or PsA to the domains included in the three selected measures (EQ5D; SF6D; and ICECAP-A)
3. To provide evidence of the face validity of the EQ5D-5L, SF6D, and ICECAP-A and their ability to capture presenteeism

5.3 Methods

This study used qualitative methods to explore whether selected measures of HRQoL were able to capture those aspects of RA, AS, and PsA that have an impact on levels of

presenteeism. Ethical approval was granted by The University of Manchester [ref: 16144].

5.3.1 Selection of measures of HRQoL

The EQ5D-5L, SF6D and ICECAP-A were selected as examples of measures of HRQoL because of their potential use as measures of outcome suitable for economic evaluations of WPIs. The measures of HRQoL are made up of a set of dimensions used to describe an individual's health status or capability. All three measures are generic and have published tariffs that allow them to be used in economic evaluations. This section describes each HRQoL measure and why they were selected for use in Chapter 5 and 6.

5.3.1.1 EQ5D-5L

The EQ5D was chosen because it is the measure of health status that is recommended by the National Institute for Health and Care Excellence (NICE) for use in economic evaluations as part of their appraisal and guideline programmes (NICE, 2013). In August 2017, NICE released a position statement about the EQ5D-5L and recommended that, at the moment, it should not replace the EQ5D-3L currently being used in economic evaluations of healthcare interventions (NICE 2017b). However, there is emerging interest in the use of the 5 level-version.

The EQ5D is a generic measure of health status used to value the impact of changes in health status in economic evaluations. The EQ5D is made up of five dimensions (see Table 5.1). Each dimension is broadly defined allowing the EQ5D to be a suitable measure for patients with different conditions; for example, the definition of 'anxiety/depression' is not restricted to those with a diagnosed condition, it is also relevant for people who feel low or depressed and have not received a diagnosis (Gudex 2005).

Table 5.1 EQ5D dimensions

Dimensions	Definition
Mobility	How well can the individual walk around
Self-care	Can the individual wash or dress themselves
Usual Activities	How well can the individual work, study, do housework, spend time with family or do leisure activities
Pain/Discomfort	How much pain or discomfort is a person experiencing
Anxiety/Depression	How anxious or depressed is the individual feeling

Source: Gudex (2005)

The newer EQ5D-5L version was not immediately released with its own set of preference weights. In order to satisfy immediate research needs in an interim set of weights was estimated using mapping techniques (van Hout et al. 2012). In January 2016, the Office of Health Economics (OHE) released a research paper detailing the methods they had used to generate weights specifically for the EQ5D-5L (OHE 2016). Two methods were used to elicit preferences weights: 1) the TTO method and 2) discrete choice experiments (DCEs). The TTO method elicited values for each state with 1 (full health) and 0 (dead) defined as anchor points. DCEs generated binary data allowing for the derivation of a scale of non-anchored relative values. Interviews were designed to ask respondents to self-report their health using the EQ5D-5L, including the VAS, to answer questions about their background, and to complete ten TTO tasks or seven DCE tasks. Respondents who completed the TTO task were asked to imagine living for ten years in a given EQ5D-5L health state followed by death. The value for each health state was derived by identifying the number of years in full health (zero to ten years) they considered to be equivalent to each health state they were presented with. The idea was that more severe health states would require more years in full health to be sacrificed. For very poor health states all years in full health may be traded off indicating a value state that is less than or equal to zero (worse than dead).

The DCE tasks presented respondents with two health states with no reference to the duration of each state. Respondents were asked to choose the ‘better’ health state; an indifference option was not included. The study used a sample of 912 individuals from the UK general population. The results of the TTO and DCE study suggested that pain/discomfort and anxiety/depression should receive the greatest weight. The EQ5D-5L is also anchored on a scale of one (perfect health) and zero (death), however, like the EQ5D-3L, the EQ5D-5L also accounts for health states worse than death. The minimum value the EQ5D-5L can score is -0.281 compared with -0.594 by the EQ5D-3L.

5.3.1.2 SF6D

The SF6D is an alternative measure of health status that can be used to value health in economic evaluations of healthcare interventions. The SF6D evolved from the SF-36 a generic survey designed to measure an individual’s overall health status (Ware and Sherbourne 1992). The SF-36 is a generic health survey designed to assess patient health and was developed to aid the collection of data relating to a patient’s experience of a disease or treatment programme. The original SF-36 health survey was not

designed with its own set of preference weights which meant that it could not be used to estimate quality adjusted life years (QALYs) needed for economic evaluations of health-related interventions (Brazier, Roberts, and Deverill 2002).

In 2002, Brazier, Roberts and Deverill modified the SF-36 and estimated preference weights for each of the six dimensions included in the SF6D (see table 5.2). Each of the six dimensions of health (see chapter five, section 5.4.1.2) were assigned two to six levels allowing for a total of 18,000 health states to be defined. The valuation of a sample of 249 health states was completed using a sample of 836 members of the general public. All respondents were asked to value six of the 249 states using the standard gamble (SG) technique. The SG technique asks respondents to make a choice between two or more alternatives based on risk; for example, risk of death or risk of developing a chronic condition (Drummond et al. 2015). The SF6D is also scored between zero (worst health state) and one (full health state).

Table 5.2 SF6D dimensions

Dimensions	**Definition**
Physical Functioning	Does the individual's health limit their ability to do vigorous activities, moderate activities, or bathing or dressing? *Vigorous activities is defined as running, lifting heavy objects, participating in strenuous sports* *Moderate activities is defined as moving a table, pushing the vacuum cleaner, bowling or playing golf*
Role Limitation	Does the individual have problems with work or carrying out other daily activities because of physical or emotional health problems?
Social Functioning	Does the individual's health problem limit their social activities? *Social activities include visiting friends and family*
Pain	How much pain interferes with the individual's ability to do both work inside and outside of the home?
Mental Health	How often does the individual feel tense or downhearted?
Vitality	How often does the individual feel they have lots of energy?

Source: Brazier, Roberts, and Deverill (2002)

The SF6D was selected for use in this study because it is a frequently used measure of HRQoL in populations with inflammatory arthritis (Tugwell, Izenda and Wells 2008; Fransen and van Riel 2009).

5.4.1.3 ICECAP-A

The ICECAP-A is part of a suite of measures of capability developed by researchers based at the University of Birmingham (ICECAP-A 2017). The ICECAP-A was selected because of its ability to capture HRQoL beyond health status, such as the ability to lead an independent life (Coast 2004). There is growing concern that quality-adjusted life years (QALYs) (see chapter two, section 2.1.4) estimated using the EQ5D or SF6D are not sensitive enough to capture changes in health and well-being in areas including mental health, social care, end-of-life, and public health (Coast, Smith, and Lorgelly 2008b). Sen (1993) advocated an alternative approach where the evaluative space consists of functioning and capabilities to function. The branch of extra-welfarism that is used by NICE restricts the focus to health (see chapter two, section 2.2.5) and functioning with no consideration for choice and the ability to improve health. Coast, Smith, and Lorgelly (2008b) argued that the capability approach offers a broader evaluative space that allows other dimensions to be valued. Sen did not state possible dimensions of capability; however, others have suggested capabilities including bodily health, integrity, imagination and thought, emotions, and the ability for an individual to control their environment (Nussbaum 2003).

A study by Grewal et al., (2006) conducted in-depth interviews with older people and found their lives were limited through their reduced *ability* to obtain various aspects that make up quality of life. The findings were linked to the capability approach and subsequently the ICEpop CAPability-O was developed for older people (ICECAP-O) (Coast et al. 2008a). Following the development of the ICECAP-O, in-depth interviews were held with adults to identify capabilities that determine their HRQoL and the evidence led to the development of the ICEpop CAPability measure for Adults (ICECAP-A: see Table 3) (Al-Janabi, Flynn, and Coast 2012). The ICECAP-A has been recommended by NICE public health guidelines as a potential measure to capture non-health related benefits in the form of capability (NICE 2012b). More information about the development of the ICECAP-A can be found in chapter six, section 6.3.3.

Table 5.3 ICECAP-A capabilities

Dimensions	Description
Feeling settled and secure (Stability)	The absence of dramatic change and stress in life. A person who feels settled is likely to be employed, in a long-term relationship, have a good income and live in a good area
Love, friendship and support (Attachment)	An individual's ability to interact and have good quality relationships with other people
Being Independent (Autonomy)	An individual's ability to look after themselves and make their own decisions
Achievement and Progress (Achievement)	An individual's ability to move forward and achieve their goals
Enjoyment and Pleasure (Enjoyment)	An individual's ability to enjoy life and feel content.

Source: Al-Janabi et al. (2013)

Preference weights for the ICECAP-A were then estimated using a type of DCE named 'best-worst scaling' (BWS) (Flynn et al. 2015). A sample of 413 individuals from the UK general population was asked to choose, from a set of hypothetical scenarios, the best and worst attributes within a profile. For example (taken directly from Flynn et al., 2015), a profile may include the highest levels of autonomy and achievement, middle levels of stability and lower levels of attachment and enjoyment. Individuals were asked which attribute they felt was the best and the worst. The resulting capability values were estimated as a function of how frequently each attribute was selected as the best or worst. Results from the BWS experiment were adjusted for heterogeneity and a population tariff for ICECAP-A capabilities was estimated and reported by Flynn et al., (2015).

The use of measures of capability has not yet been specified in a resource allocation decision-making context. The goal of maximising health benefits translates to a decision-rule where the intervention that produces the greatest amount of QALYs per cost should be funded. Applying the health maximisation principle to the capability approach may not make sense (Coast, Smith, and Lorgelly 2008b) and instead it has been proposed that rather than aiming to provide maximum benefits to society the aim should be to provide a decent minimum level of capability (Coast, Smith, and Lorgelly 2008b). At present, a major limitation of applying the capability approach to economic evaluations is that it is not clear how capabilities should be compared (Cookson 2005).

As a consequence, no agreed decision rule for capability currently exists (Lorgelly 2015).

5.3.2 Rationale for using Qualitative Methods

This study applied qualitative methods to assess the extent to which three selected measures of HRQoL could be used to capture the impact of WPIs on changes in presenteeism. Qualitative methods are used to understand the 'what', 'why', and 'how' of a research question rather than the 'how many' or 'how much' questions addressed using quantitative studies (Greene and Thorogood, 2009). For example, a patient-satisfaction survey generates quantitative data that can be used to analyse trends; however, understanding why the patient has rated their satisfaction level as good or bad requires further explanation from the respondent. Qualitative methods also allow researchers to explore issues from alternative perspectives highlighting how and why individuals may interpret situations in differently.

5.3.2.1 Qualitative Data Collection Methods

There are a number of different types of qualitative data collection methods including: focus groups, ethnography, and one-to-one interviews (Greene and Thorogood 2009). Focus groups are used to bring together relevant individuals to discuss a specific issue. The aim of a focus group is to collect opinions regarding a specific issue; for example, to evaluate the delivery of a certain type of healthcare service (Greene and Thorogood 2009). The disadvantage of using focus groups is that they are unlikely to generate in-depth accounts of the specific issue being discussed (Berg 2001).

Ethnography requires the researcher to observe and record the behaviour of individuals in their particular environment. Typically, the researcher spends time with groups of interest and records behaviour patterns taking into account the social, cultural and environmental context. The advantage of ethnography is that it allows the researcher to gain a deeper understanding how contextual factors influence the area of research being investigated (Greene and Thorogood 2009). However, to successfully conduct an ethnographic study a substantial amount of time is needed to set-up and conduct the study, collect, and analyse the data (Greene and Thorogood 2009).

The purpose of one-to-one interviews is to gain insight and opinions about a specific research topic (Greene and Thorogood, 2009). There are broadly three types of interview methods that can be utilised when carrying out qualitative research: structured

interviews; unstructured (or depth) interviews; and semi-structured interviews. Structured interviews follow a strict set of questions and are most useful for researchers who already have a clear understanding of the research area they are investigating. The questions designed for use in structured interviews are based on two assumptions: 1) all questions included cover all relevant information that needs to be extracted from the participant; and 2) all questions asked are interpreted and understood in the same way by all participants involved in the study (Berg 2001).

Unstructured (or depth) interviews are opposite to structured interviews; no questions are set prior to the interview. Researchers tend to use this type of interview technique when they are unfamiliar with the research area under investigation and are, therefore, unable to fully anticipate all relevant questions that need to be asked in advance of the study taking place (Berg 2001). Unstructured interviews are commonly used during the initial stages of research to begin exploring the research area of interest (Berg 2001).

The third type of interview method is the semi-structured interview. Semi-structured interviews are a useful way to explore and gain a deeper understanding of a specific research area (Greene and Thorogood 2009). Semi- structured interviews are, by design, half-way between a structured and unstructured interview. Semi-structured interviews, like structured interviews, use a set of pre-determined questions; however, and like unstructured interviews, the researcher is free to pursue areas of interest and ask questions (probing questions) that were previously not defined in order to find out more (Berg 2001).

The nature of the principle research question for this study required a specific set of questions to be answered; however, because of the exploratory nature of this study, a degree of flexibility was needed to allow respondents to elaborate and discuss potential issues that were previously unforeseen. Therefore, given the requirements of the study, semi-structured interviews were selected as the most appropriate data collection method.

5.4.1 Relevant Study Population: Employees with Inflammatory Arthritis

The relevant study population were employed adults currently working with RA, AS or PsA. It was important to recruit individuals who were employed at the time of the interview in order to capture relevant and accurate experiences and opinions of individuals who were currently living in that situation. It was necessary to exclude

individuals who do not work or who no longer work in paid jobs because of the potential to capture distorted accounts. As time passes, individuals who no longer work in paid employment may look back and view their situation differently compared to when they were actually living through the experience.

5.4.1.1 Selected study sample

To represent the views of the relevant study population, a purposive sample of working individuals, aged 18 years and above, and diagnosed with RA, AS, or PsA was recruited. For the purposes of this study, a working individual was defined as a person who is employed or self-employed working in full-time or part-time for pay. Individuals who volunteer were excluded from the study because pressures to continue working at a high level was potentially less pronounced compared to those working for pay.

Participants with RA, AS or PsA were recruited via advertisement of the study on support group websites including The National Rheumatoid Arthritis Society (NRAS), The National Ankylosing Spondylitis Society (NASS) and The Psoriasis Association (PsAZZ). The advertisement provided brief information about the topic of the study and an email address to contact if they were interested in taking part in the study. Further information about the study was sent by email to those individuals who expressed an interest to participate. Once the potential participant agreed to take part a date and time was arranged to conduct the interview (Appendix 5.1). Participants who did not respond within a week were sent a second email asking the participant to state a suitable time and to conduct the interview. If no response was received no further contact was made. The recruitment strategy aimed to recruit individuals based across the United Kingdom, therefore, it was necessary to conduct the semi-structured interviews over the telephone, rather than face-to-face.

5.5.2 Data Collection

Informed consent was taken over the telephone prior to the interview taking place. All interviews were audio-recorded and transcribed verbatim (by one researcher, CJ).

5.5.2.1 The interview schedule

The interview schedule (Appendix 5.2) was developed through discussions with two expert researchers in the areas of presenteeism and inflammatory arthritis. A protocol for this study was written up and sent to researchers that specialise in presenteeism, qualitative methods and inflammatory arthritis based within the Centre for

Musculoskeletal Health and Work (CMHW). Their comments and feedback were discussed and where necessary incorporated into the overall design of the study. Patient and public involvement (PPI) describes research that is being carried out *with* patients or members of the public rather than to, for or about them (INVOLVE 2018). PPI for this study would have involved employees working with RA, AS or PsA.

The interview schedule consisted of three sections. The first section included questions that asked the respondent about their rheumatic condition. Information about their rheumatic condition included: 1) when they received a diagnosis; 2) what medication they take and whether they thought it was effective; 3) how severely they believe their condition affects them on a day-to-day basis; and 4) whether they experience fluctuations and, if so, how frequently.

The second section asked questions about the respondents' jobs in terms of: 1) the tasks they complete; 2) whether they work independently or as part of a team; and 3) whether colleagues were able complete/help them with their work.

The third section asked respondents to discuss whether, and how, their rheumatic condition affects their ability to work. The first question, 'bearing in mind your condition, do you feel you are able to achieve what you want to achieve at your job?' was designed to have a broad interpretation allowing the respondent to describe whether they felt able to achieve everything they would like to at their job. The term 'achieve' was not specifically defined permitting respondents to define 'achievements' in their own way. The reason 'achievement' was not explicitly defined was to avoid directing answers that could have been associated with the selected HRQoL measures thereby reducing the potential for researcher bias (where researchers influence the results to identify a certain outcome).

The second question of the third section asked respondents to think about how fluctuations in their disease activity affect their ability to work. Respondents were asked to comment on whether their condition affects the quantity and/or quality of work they were able to produce. If the quantity of work was reported to be reduced, respondents were then asked to explain what happened to the work that was left uncompleted; were colleagues able to pick-up extra work or did they work overtime (compensation mechanisms, see chapter two, section 2.3.4)

Respondents were also asked to describe instances where they changed or altered their job because of their condition. This question was designed to gain insight into the historical development of their condition and their working experience, how they reacted to the changes made, and how employers managed their employees' needs to help them to continue to work. The penultimate question asked the respondent to describe any further changes they would like to help them to work more effectively. The final question asked the respondent to state whether they disclosed their condition to their employer and why.

5.5.5 Qualitative Data Analysis Methods

Broadly, there are three methods that can be used to analyse data collected from semi-structured interviews. The choice of analysis method depends upon the aim of the study. If the aim was to identify common concepts related to a specific subject thematic content analysis methods are recommended (Greene and Thorogood, 2009). Thematic content analysis compares participants' accounts within and between each other to identify common concepts and ideas. The concepts identified are then categorised into a set of themes that describe the key issues raised.

In contrast, if the aim of the study was to derive a new theory then grounded theory methods (GTM) would be recommended (Urquhart 2013). GTMs involve systematically analysing participants' accounts, word by word, to identify key concepts. Data collection and analysis are performed simultaneously. The identification of new concepts that emerge from the analysis of the data informs the researcher to continue sampling and collecting data. The process of sampling and analysis is repeated until 'saturation' (where no further concepts are introduced) is reached at which point data collection ceases.

A third qualitative analysis method is the Narrative Analysis Method (NAM). NAMs aim to understand the experiences and points of view through participants' stories. The method takes account of the relationships between the individual experience and cultural context (Clandinin 2006).

Selection of appropriate data analysis method

The aim of this qualitative study was to identify aspects of RA, AS and PsA that make work challenging and to assess whether those aspects are captured by measures of HRQoL. Thematic content analysis, specifically the framework analysis method

(FAM), was selected to analyse the data because it offered a way of identifying and categorising aspects of RA, AS, and PsA that affect ability to work. The FAM is used to identify common themes within the data that describe the key issues raised (Greene and Thorogood 2009). The method is most frequently used to analyse data collected from semi-structured interview transcripts (Gale et al. 2013). An advantage of FAM is that it allows the data to be analysed systematically though the production of a matrix made up of cases (participants) and codes (themes).

Themes (codes) may be developed using one, or a combination of, two methods: 1) deductive analysis methods; and/or 2) inductive analysis methods. The deductive approach relies on a pre-existing theory or literature from which the themes are pre-selected and the inductive approach involves analysing the data and identifying new themes that emerge. In situations where the deductive approach is used qualitative researchers also recommend performing an inductive analysis to ensure no further important codes (themes) are missed (Gale et al. 2013). For the purposes of this study, both deductive and inductive methods were used to identify key themes. The domains included in the EQ5D-5L, SF6D and ICECAP-A were used to develop themes needed for the deductive analysis (see section 5.3.1).

The degree to which the three selected outcome measures conceptually overlapped was potentially high. The EQ5D-5L and SF6D are two measures of health status that include similar dimensions the describe health; for example, the SF6D includes physical functioning and the EQ5D-5L includes mobility. Health status may be viewed as a subset of capability where an individual needs a certain level of health to attain a high level of quality of life (Lorgelly 2015). Certain capabilities included in the ICECAP-A, for example ‘Enjoyment and Pleasure’, conceptually overlap with ‘mental health’ (SF6D) and ‘anxiety and depression’ (EQ5D). Therefore, to prepare for the deductive analysis, each measure was analysed to identify the extent to which the dimensions and capabilities conceptually overlapped. The concepts that overlapped across two or three of the selected measures were collected to generate a group of independent themes.

Data were analysed on an iterative basis and data collection continued until saturation was reached after which data collection ceased. All qualitative data was uploaded into the computer software package NVivo (version 11.0) ready for analysis.

5.4 Results

Twenty-two employed individuals with RA, AS, or PsA were interviewed: 10 with RA; 9 with AS; and 3 with PsA. Four participants were males and 18 were female. Due to the small number of males whom expressed interested to participate in the study, a second advertisement of the study modified to specifically recruit males was released. The advertisement was unsuccessful and no further males with RA, AS or PsA were recruited. Five participants worked in a manual job, or had a significant manual component, and the remaining 17 participants worked in non-manual roles. Table 5.4 summarises the characteristics of each participant.

Table 5.4 Sample Characteristics

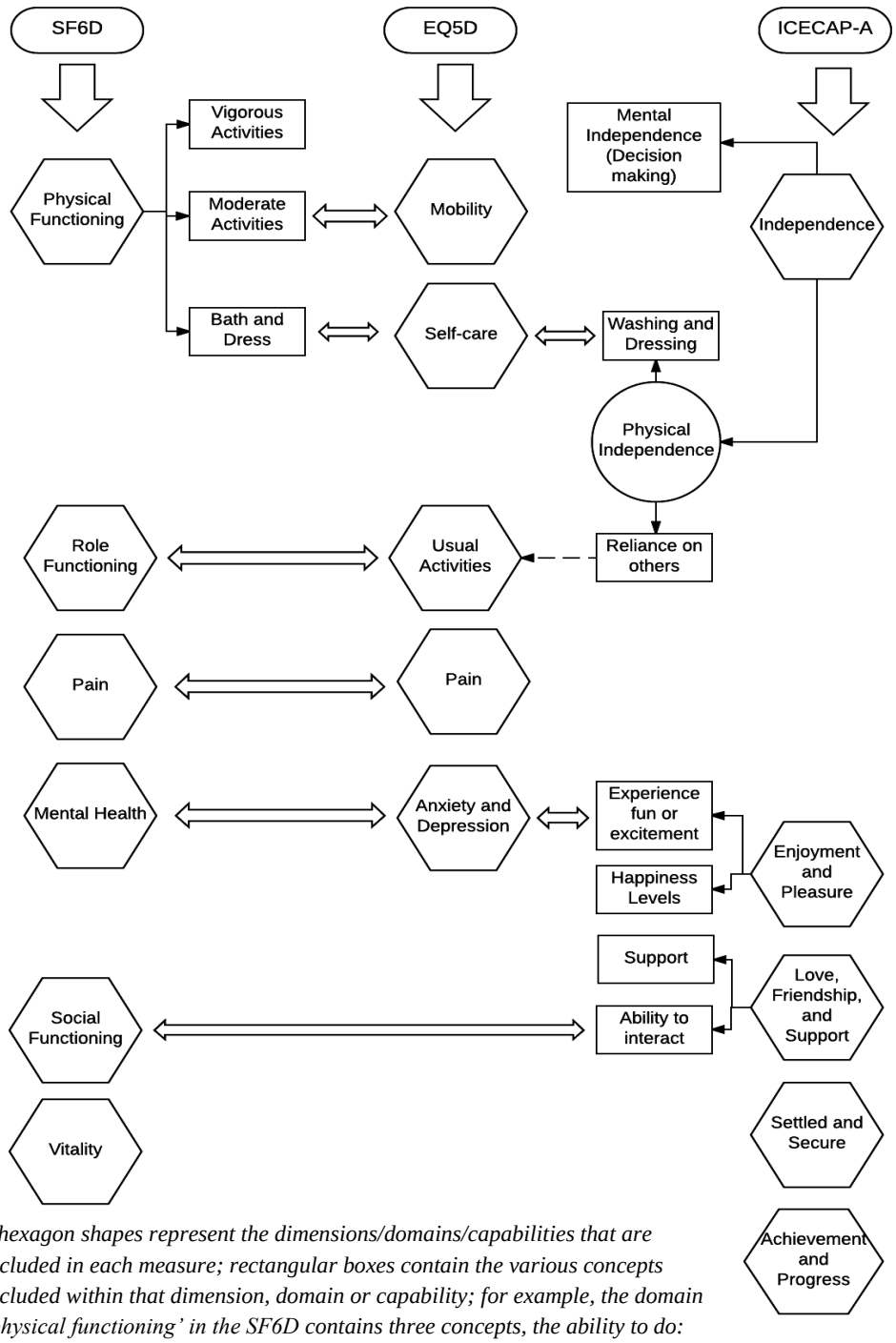
Participant Number	Condition	Gender Female (F) or Male (M)	Job	Industry	Manual or Non-Manual or Both
P101	AS	F	Police Officer	Police	Manual
P102	PsA	F	Business Support Officer	Local Government	Non-manual
P103	RA	F	Research Project Officer	Research	Non-manual
P104	RA	F	Relationship Support Manager	Banking	Non-manual
P105	AS	F	Specialist Behaviour Practitioner	Care	Both
P106	AS	M	Store Manager	Retail (Electrical)	Manual
P107	AS	M	Joiner	Construction	Manual
P108	AS	F	Lecturer	University Education	Non-manual
P109	RA	F	Teacher	Education	Non-manual
P110	PsA	F	Solicitor	Engineering	Non-manual
P111	AS	F	NHS Co-ordinator (Admin)	Healthcare	Non-Manual
P112	RA	F	PA at Stately Home		Non-manual
P113	AS	M	Cardiac Pacemaker	Healthcare	Both
P114	RA	F	Health Inspector	Healthcare	Non-manual
P115	RA	F	Pharmacy	Healthcare	Non-manual
P116	AS	M	Training Manager	Professional Services	Non-manual
P117	RA	F	Biomedical Scientist	Healthcare	Non-manual
P118	RA	F	Document Controller (Admin)	Engineering	Non-manual
P119	RA	F	Account Clarke	Engineering	Non-manual
P120	RA	F	PA in Finance	Banking	Non-manual
P121	RA	F	Product Merchandiser	Retail	Non-manual
P122	PsA	F	Locum GP	Healthcare	Non-manual

RA: Rheumatoid Arthritis; AS: Ankylosing Spondylitis; PsA: Psoriatic Arthritis: GP; General Practitioner

5.4.1 Themes identified for subsequent deductive analysis

Figure 5.1 illustrates the concepts included in each measure and where they overlap.

Figure 5.1 Conceptual overlap between EQ5D, SF6D, and ICECAP-A [a]



[a] *hexagon shapes represent the dimensions/domains/capabilities that are included in each measure; rectangular boxes contain the various concepts included within that dimension, domain or capability; for example, the domain 'physical functioning' in the SF6D contains three concepts, the ability to do: 'vigorous activities', 'moderate activities', and to 'bath and dress'; circles indicate that there is a link between two concepts: however, the concept in the circle covers a broader definition which is then split into its separate conceptual parts. The dashed arrow connecting 'Usual Activities' to 'Reliance on others' represents a partial overlap of concepts.*

‘Physical functioning’, included in the SF6D, comprise three concepts including the ability to do ‘vigorous activities’, ‘moderate activities’, and to ‘bath and dress’. The EQ5D-5L included two domains that conceptually overlapped with ‘physical functioning’, ‘mobility’ and ‘self-care’. The ICECAP-A also conceptually overlapped with ‘self-care’ through the ‘independence’ domain where part of the definition was the ability to look after one’s-self (self-care). Three separate themes emerged: 1) vigorous activities; 2) moderate activities and mobility; and 3) self-care (bathing and dressing one-self). Independence was split down into two further themes: 1) mental independence, defined as the ability to make decisions; and 2) physical independence, defined as the ability to look after themselves without having to rely on others. Physical independence was also split into two separate themes: 1) the ability to wash and dress; and 2) reliance on others to carry out tasks. The latter of the two physical independence themes partially overlapped with ‘Usual activities’ and ‘Role Functioning’ (represented by the dashed arrow); however, reliance on others also depicts a loss of independence and, hence, was kept as an independent theme.

The ICECAP-A included the concept of being able to have ‘enjoyment and pleasure’ in life and was divided into two sub-sets: 1) the ability to experience life as fun or exciting; and 2) happiness levels. ‘Mental health’ (SF6D), ‘anxiety and depression’ (EQ5D) and happiness levels (ICECAP-A) were combined into one theme labelled ‘mental health’. A separate theme ‘Enjoyment’ defined as the ability to enjoy activities in life was also identified. ‘Role functioning’ (SF6D) and ‘usual activities’ (EQ5D) were concepts that were very similar and therefore were combined to create one theme. ‘Pain’, from the EQ5D-5L and SF6D, are conceptually the same and were represented by one theme.

‘Love, friendship and support’, included in the ICECAP-A, was divided into two concepts: 1) the ability to interact with people; and 2) to feel supported. ‘Social functioning’ (SF6D) was described as the ability to take part in social activities, for example visiting friends and family, which overlapped with the ICECAP-A’s ability to interact with people. Two separate themes emerged: 1) the ability to interact (social functioning); and 2) the ability to gain support. There were three domains that did not overlap: ‘vitality’ (SF6D), ‘feeling settled and secure’ (ICECAP-A), and the ability to ‘achieve and progress’ (ICECAP-A). Table 5.5 lists the 14 independent deductive themes identified from the HRQoL measures. Each theme was categorised as a symptom, activity or capability.

Table 5.5 Deductive Themes

Concept/theme	Definition	Symptom, Activity or Capability
Pain	How much pain is the individual experiencing and does it interfere with their work	Symptom
Mental Health	How anxious, depressed or tense the individual feels and how often do they feel this way	Symptom
Vitality	How often does the individual feel they have lots of energy	Symptom
Moderate activities and mobility	The individual is able to walk, use a vacuum cleaner, play light sports such as golf, lift objects such as a table	Activity
Self-care (washing and dressing)	The individual is able to was and dress themselves without help from others	Activity
Usual activities and role functioning	The individual can perform their usual activities including housework and work	Activity
Vigorous activities	The individual can run, lift heavy objects and participate in strenuous sports	Activity
Social Interaction	The individual's health limits their ability both physically and mentally to engage in social situations	Activity
Achievement and progress	Can the individual attain goals in their lives that are important to them?	Capability
Enjoy Activities	Ability to enjoy life and experience exciting things	Capability
Mental Independence (decision making)	The individual is able to make decisions for themselves without having others make decisions for them	Capability
Physical Independence (Reliance on others)	The individual relies on others to complete tasks such as house work and tasks at their job	Capability
Settled and secure	The individual is able to obtain a stable life where dramatic changes and stress is not present	Capability
Support	Does the individual feel supported by family, friends, and people they work with	Capability

5.4.2 Framework Analysis

In total, 14 independent themes were identified from the health status and capability measures. The deductive analysis involved analysing the 22 qualitative accounts to identify evidence that supported the 14 themes of the measures of HRQoL. The evidence was then used to assess the extent to which the three selected measures were able to capture the impact RA, AS or PsA had on presenteeism. The inductive analysis involved identifying evidence of themes that were missed by the EQ5D-5L, SF6D, and ICECAP-A. One extra theme was identified to affect levels of presenteeism: 'mental clarity'. Two further inductive themes were identified to have an impact on absenteeism and not presenteeism; these are reported in Appendix 5.3.

The next section presents evidence of the themes that were captured by the health status and capability measures. The themes are presented in three subgroups: symptoms; activities; and capabilities. The analysis method used to identify each theme (deductive or inductive) is stated in brackets.

5.4.3 Themes related to symptoms of RA, AS, or PSA

This section reports the qualitative evidence identified in support of each symptom theme. There were three themes capturing symptoms including: pain; vitality; and mental health. The symptom 'stiffness' was also identified but was reported within the 'activities' section because the effects of 'stiffness' were appropriately and adequately captured by domains such as mobility and usual activities.

Theme 1: Pain (Deductive)

Pain was explicitly captured by the EQ5D and the SF6D. Many participants stated that they constantly feel some level of pain which may become worse when experiencing a flare. Pain was reported as a reason for being unable to perform work as usual. One interviewee, a general practitioner (GP), stated that the pain she experienced had prevented her from being able to perform invasive procedures that were part of her job:

"*I stopped putting in coils and stopped doing any invasive procedures apart from taking blood, because I have decided that if I was my patient I wouldn't want someone in pain and therefore distracted from doing this"* (P122, PsA, female)

Another interviewee, a front line police officer, said that the pain level of pain she sometimes feels makes her job unsafe and difficult to do:

"it's dangerous isn't it… because I am potentially running after people, fighting with people, and I know deep down, if I'm limping then I'm struggling with walking, but I know it's unsafe for me if I'm fighting…." (P101, AS, female)

Interviewees also explained that that whilst pain is a difficult symptom to cope with, they are, however, able to adapt and can ignore pain to some extent:

"I mean, I've got used to the pain now…that sounds a bit bizarre, but when you have pain every day, you do, kind of, get a limit that you can handle and ignore" (P111, AS, female)

However, some explained that whilst pain is controllable they rely heavily on analgesics to control the pain:

"my pain is pretty well controlled, although I think I would have a major meltdown if I thought I didn't have any co-codamol nearby" (P115, RA, female)

Theme 2: Mental Health (Deductive)

Mental health was captured by all three measures. Mental health was not identified to directly affect ability to work; however, many interviewees stated that they felt anxious, guilty and worried about having to take time off work because of their condition:

"I switch my work phone off all other days of the week and none of my other colleagues do….so they answer calls outside of their on-call day, they end up fielding my calls outside of my on-call day urm and….and I actually feel really guilty, I feel like I'm not doing my job properly but actually I'm, sort of, protecting myself" (P122, PsA, female)

Another interviewee explained that, because of her condition, she works many hours to help her avoid feeling anxious about not being able to do her job well:

'yeah…I know I work too hard, I work too many hours, but I just…I suppose I'm scared that I won't be able to do my job as well as everybody else in the team, I don't want to be one who everyone makes allowances for urm…I want to be up there, hitting the targets and doing what I'm supposed to do, you know, the job that I do, I feel it's really important' (P114, RA, female)

On the other hand, one interviewee explained that working actually made her feel better, because it allowed her to get her mind away from her condition:

'I'm trying to hold down the job…trying to…working has a lot of other things… it's not just about money… it's not about money for me… for me it's about making a difference, I enjoy the work, I enjoy the people, it takes your mind off things and you feel better for working than you do actually not working' (P102, AS, female).

Theme 3: Vitality (Deductive)

Fatigue is the symptom caused by RA, AS, and PsA and is captured in the SF6D through the vitality domain. Interviewees described fatigue as the symptom that is most difficult to work with, adapt to and avoid:

"*it does vary, urm the hardest thing I find to live with is the fatigue urm…I can cope with pain and there's ways of getting round doing things urm and at work, people will help me if I can't do something, but the fatigue nobody can really help with that"* (P117, RA, female)

Interviewees felt that they need more energy to work fast and effectively:

"*to be fully effective in my job I would need to be able to write a lot more than I can and I would need to have more energy to do extra things"* (P109, RA, female)

The interviewees also explained that fatigue limits their ability to work full-time hours:

"The fatigue I suffer from….it really is that because those actual… that's why I don't work full-time. I couldn't imagine now, having to work full-time hours, I can only do these couple of days… it feels like I drag myself through those couple of days, if that makes sense…" (P102, PsA, female)

Fatigue also reduces their ability to work extra hours needed for their job, for example some people need to attended professional development courses during the evenings:

"*we have to stay up to date, continuing my professional development […]the local health authority will have an up-date on their latest policy on antibiotics, or something like that, and I might say 'I'm sorry, I can't go to that, I just feel too ill", or "I'm still too tired"* (P115, RA, female)

Some interviewees stated that the fatigue they experienced had forced them to change job:

"*I had to give up being a solicitor in a private practice urm…because I was too*

knackered, I was just exhausted all the time and having issues[….], and I just had to step away from that entirely" (P108, AS, female)

Interviewees also explained that fatigue makes getting up early for work particularly challenging. Some interviewees had the advantage of working flexible hours and most choose to arrive at work later in the morning.

Participants also said that they used a lot, if not all, of their energy at work which left them with very little to complete tasks at home impacting on the amount of work that could be done at home:

"*I think, what happens is that just because you get so tired urm…that once you've done a day's work, when you get home you are not fit for anything else so because you have to do the work it's things at home that urm… don't get done to the standard that you'd want them to do because, you know, you have to pace yourself and stuff, so if you were less time at work you could get more stuff done at home"* (P120, RA, female)

5.4.4 Themes capturing the impact on activities and daily tasks

Activities are represented by those domains that describe a person being able to complete a task such as walking, daily activities, and self-care. Five activity themes were identified.

Theme 4: Moderate Activities and Mobility (Deductive)

Mobility and moderate activities are captured by relevant domains in the EQ5D and SF6D. Mobility problems can have a direct impact on an individual's ability to do their job. One participant explained that, as part of her role she was expected to visit the farms they work with; however, her condition has reduced her capacity to walk:

"No, it would be physically going round their premises…there are one or two that I go to where we have a meeting in the house and discuss financial stuff ur…which is fine, that's manageable, but obviously, I can't go traipsing round fields and things like that" (P102, RA, female)

Another participant explained that having a flare and not being able to walk at a key event would stop her from being able to do her job:

"if we've got an event, urm…if I was suffering a flare at the time of the event, then that would be really bad urm…thing to have to deal with because it's only me who could do

it and I, you know, if it was too bad I wouldn't be able to be there but luckily that's never actually happened, I haven't had a flare coincide with a time when I've had to be up on my feet and walking about" (P120, RA, female)

Many participants also struggle to climb up and down stairs which causes problems at work if there are no lifts:

"I couldn't do stairs and my office is upstairs so… I couldn't, I couldn't get up and down the stairs, well I can in the house because I crawl up and come down on my bum, which is fine, but you can't be coming down into a banking hall, in a bank, on your bum [laughs]…" (P104, RA, female)

Interviewees largely agree that being able to walk frequently throughout the working day helps to manage levels of stiffness. Interviewees frequently stated that long meetings are difficult because they cannot sit for that length of time without becoming very uncomfortable.

About half of the interviewees stated that their limited levels of mobility create problems with commuting. Interviewees agreed that stiffness, caused by the conditions, is worse during the morning than at any other point during the day. Two interviewees explained that they have to get up very early in order to have time to stretch out and get moving ready for work:

"I'll set my alarm, kinda, an hour before I need to get up…urm see how I'm feeling and I might nod-off again if I'm feeling alright, if not I might need to get up and do some yoga or some stretches" (P101, AS, female)

Another stated that they start work later because of their decreased mobility:

"*I am altering the actual hours I work a little, in self-returns, so I'm starting a little bit later in the mornings because… the arthritis makes you very stiff in the morning… from… whereas the team starts work at 9am, I'll start at 9:30"* (P102, PsA, female)

Theme 5: Self-care (Deductive)

Self-care relates to an individual's ability to wash and dress and is captured by relevant domains in the EQ5D and SF6D. Some participants find it difficult to get washed and dressed for work in the morning. One participant explained that it is easier for her to work from home because she does not have to deal with getting dressed and washed

quickly in the morning and being ready for work:

"you know fact is you've got to get yourself ready to go to work... that would be, you know, that would be awful. I mean, as it is, like I said, I can just haul myself out of bed and sit down here in my pyjamas and carry on working and answer my phone and no-one would know any different" (P104, RA, female).

Theme 6: Usual Activities / Role Limitations (Deductive)

Usual activities or role limitations are captured by relevant domains in the EQ5D and SF6D. Interviewees who work in non-manual jobs stated that they struggle to write and type which negatively impacts on the amount and quality of work they can produce:

'yeah, absolutely, urm if it's very bad and everything is very stiff… I can have trouble using my hands and urm… and you know, a lot of the work involves typing and stuff and I do have a dictation thing, but it's not great, and I have work arounds for it' (P108, AS, female)

Interviewees in manual jobs also explained that they have difficulty performing physical tasks required for their jobs. One interviewee, a joiner, said that he struggles with his usual tasks; such has having to stretch to reach things and bend down to pick things off the floor:

'it's really stiff urm…my...my lower back, it well, its fused my lower back now, so I do struggle to pick up objects off the floor, I…r…rather than bend from the hip now […]...I fit windows and doors, skirting boards, things, for kitchens, urm..ur…flooring, I do, which [laughs] as you can imagine, it is a struggle' (P107, AS, male)

Many interviewees also stated that their condition makes driving difficult, which can make commuting to work a challenge. One interviewee, who used to take the train to work, can no longer do so after she experienced a major flare and now she has to work from home.

Numerous interviewees commented on how their condition affects their ability to perform everyday tasks at home, such as cooking and cleaning the house. The condition also had an impact on non-work roles. One participant explained that she feels guilty because she struggles to fulfil her role as a mother and wife:

"It affects me less now because my children are both at University, but when they were younger and at home I felt dreadful because they had after school activities and they, you know, I cook when I get home and sometimes I, they would, my husband would have to cook and so they would get something with chips [laughs] or they would get fish and chip or McDonalds so…healthy eating went out of the window" (P114, RA, female).

Theme 7: Vigorous Activities (Deductive)

The ability to perform vigorous activities is captured by a relevant domain in the SF6D. The inability to do vigorous activities affects working life and leisure. First, participants in manual jobs explained that they struggle to lift heavy objects and fulfil highly physical tasks. One participant, who used to work with the army, explained that his condition stopped him from being able to take part in certain aspects of the training:

"P116: *So going to Afghanistan, for example, you'll do about 7 months' worth of preparation training which will involve firing riffles and also doing role play in a scene, or a mock-up village urm…so, you know, you kind of patrol through the village and have a mocked attack that you would have to react to, so some of that I couldn't get involved in and the bits I could get involved in, I suddenly found that, actually, I can't do a lot of this, which again was really really frustrating for example, a casualty…imagine we've got somebody who's pretending to be a casualty and that person has to be carried by four people, to safety, that I found really, really difficult and, at times, I just couldn't do it"*(P116, AS, male)

Theme 8: Social Interaction (Deductive)

Social interaction is captured by relevant domains in the SF6D and ICECAP-A. Many participants felt that their condition stops them from being able to interact positively with work-colleagues and, as a consequence, they are more likely to isolate themselves:

"R: I'm probably not as urm…I wouldn't say I'm not cheerful, but I probably isolate myself more…I shut my office door and that's a clear sign to leave me alone" (P103, RA, female)

Many participants explain that pain and fatigue affects their mood causing them to come across as angry or grumpy to co-workers, potentially affecting working relationships:

"And because people don't actively see it, because you don't see the disease, they think you're being a bit mardy and that kind of thing….so, you know, it can make… then you

end up feeling a bit precious…so… yeah it probably affects relationships a little bit and general how much you can achieve in one day." (P102, PsA, female)

One participant, who works from home because of her condition, finds that she misses the social interaction working in an office brings:

"Sometimes you just miss a little bit of banter across the desk and urm.." (P104, RA, female)

Interviewees also explained that working extra hours during evenings and weekends, or simply feeling exhausted after a working day, can have a negative impact on relationships within families:

" yeah so I'm going to get this work done, it'll be fine, urm…because tomorrow then, you know, everything will be good because tomorrow I'll have less work to do, I'd have caught up, but the other side of it is that I don't get to spend time doing weekend things urm…at the moment my daughter is studying as well, so it's alright, so she's in her room and I'm on the computer but yeah, it's a bit of a thing" (P114, RA, female)

One interviewee described how her condition has had a negative impact, not just on work, but also in family life:

"I think the really big thing is home life urm…it's suddenly becomes really important, you're in the middle with your illness and you've got to make some sort of decision on whether you prioritise work or home because neither of them are going to go particularly well and it almost feels like 'Oh my god, which one am I prepared to lose, my work or relationships with my family, friends and my husband?' and I would say…and actually that my husband, his aunt had rheumatoid arthritis from the age of 20, so he is quite aware of what it is, but we still got to the point where, just before Christmas, we were talking to solicitors about divorce so…you know, we are back together again but that's purely because I have been seeing a psychologist and we're working through it" (P122, PsA, female)

Theme 9: Mental Clarity (Inductive)

Mental clarity is defined as the ability to concentrate, remember tasks, and think quickly and clearly. The definition of mental clarity describes an individual's ability to do something, therefore, mental clarity was categorised as an activity, rather than a symptom. Mental clarity was not explicitly captured as an activity in the EQ5D or

SF6D; however, the effects of poor mental clarity may be captured by the ICECAP-A. Interviewees with reduced mental clarity explained that it can affect concentration and the ability to remember:

"*yeah there is the concentration part, I keep mentioning this brain fog thing, it's quite heavily linked with the condition where you do, you're concentrating but it's difficult to think of too many things at once, so I think that sometimes affects me because I have to write everything down and make sure I don't forget something […], like today, I've, sort of, forgotten a few things and it's been pulled up*" (P113, AS, male)

Interviewees also explained how the condition limits their ability to think quickly:

"my thought processes are affected, I have, what I term 'brain fog', considering I used to be the multi-task queen, I'm still coming to terms with the fact that getting one or two things done in the day is a good thing" (P122, PsA, female)

5.5.5 Capabilities

Capabilities describe a person's ability to obtain certain aspects that make up quality of life, for example being able to have good relationships and achieve their goals. The following themes were captured by the ICECAP-A only.

Theme 10: Achievement and Progress (Deductive)

The domain 'achievement and progress' can be spilt into two separate parts: 1) achievement and 2) progress. Achievement is a broad theme that has the ability to capture the impact on ability to work caused by the following themes; pain, fatigue, usual activities and role limitations, mobility, moderate and vigorous activities, and mental clarity. The majority of interviewees working in non-manual roles stated that they were able to achieve what they would like to achieve in their current jobs because their jobs are not physically demanding.

Researcher: 'Bearing mind your condition, do you feel you are able to achieve everything you want to achieve at your job?'
Interviewee: 'Ur…yes I do actually, purely because there's not a lot, I don't have to do physical labour, so yeah…it's ok' (P116, AS, male)

In contrast, three interviewees who have worked or were currently working in manual jobs stated that they struggled with the physical aspects of their jobs, such as heavy

lifting, bending and moving quickly, limiting what they were able to achieve during a working day.

'it can be really hard work and especially if some of the work is a bit more active, you have to really [something] today, that's the problem with it, you think 'right, I gotta sort this, this and this' and actually there is no way I can physically do that, so… I'd have to re-arrange it' (P106, AS, male)

The three interviewees stated that they have had to change their jobs for something less physical because they could not cope with the physical demands.

'I discovered that I couldn't carry heavy shopping which was the whole prompt of going to the GP and finding out what was wrong with me and of course, finding out then that put 'kai-bousch' on going back to France to physical labour again' (P112, RA, female)

All interviewees explained how their condition has made doing their job more difficult, with the exception of one interviewee who explained that developing RA, in some ways, made her a better pharmacist because she is more able to advise her patients:

"in some respects having RA has made me better at my job because I have spent a lot of time counselling patients who are newly diagnosed or are frightened about going on biologics and I will take them into the consultation room and I will talk it through with them and I'll say 'I can't promise it'll work for you, but it's given me my life back' and urm…if they're struggling with something that…I can say 'well try this, it worked for me' or 'I tried that and that didn't work, but a friend said such and such' through the national rheumatoid arthritis society, people I know through that urm…I've been able to help patients with the same condition" (P115, RA, female)

Progress captures a person's ability to move forward with their lives, in this case to progress in terms of their careers. Interviewees also explained that their condition prevents them from progressing further in their careers. One interviewee described her current situation where she feels she cannot apply for a promotion because her condition limits her ability to do tasks needed for the job:

"I mean my next move [job wise] is a research business manager, but I don't want to do it because I'll be sat all day, in meetings, I'll earn a lot more money, but if I'll be sat all day in meetings, that's not good for me, I need a job where I can get up and move

about, you know and that kind of thing, so I do look at jobs as to how it will make me feel physically, and therefore, mentally" (P103, RA, female)

Some interviewees explained that their condition has forced them to take a step-down in their career because they were unable to fulfil the duties required of the job:

"*I used to work at a higher level, but I've stopped that because of the health condition… it was too much, I had to travel a lot more, when I was working at a higher level and work more hours… so it's stopped me from doing that"* (P102, PsA, female)

One interviewee, a front line police officer, explained that she may be prevented from progressing towards certain roles, such as joining the riot team, because her condition stops her from being able to complete the required fitness tests.

Theme 11: Enjoyment (Deductive)

Overall, interviewees agreed that they enjoy working and want to continue working, despite their condition:

"I mean I absolutely adore the job I do, in terms of getting out and meeting people and doing that thing of just trying to make a difference and trying to make peoples life more meaningful and fulfilling, urm…so that aspect I really enjoy" (P105, AS, female)

However, other participants said that the condition can make them feel dis-interested about doing anything and working with the condition can take away the time they have to enjoy themselves:

"I wake up in the morning [and] *feel as though I haven't slept, body feels heavy and I really can't and I, you know, I still try and work those days, but really if it's a weekend and I feel like that, like last weekend is an example, I spent the whole day between the bed and my armchair, and even my armchair was uncomfortable and just urgh! I spent most of the day in bed and that's a rubbish day really, it's not good, especially on a Saturday, it's not good"* (P108, AS, female,).

Some interviewees who are approaching retirement age were concerned about their future ability to enjoy their retirement once they finished working:

"I said I want to continue working, I need to do something to keep my brain active, but he [husband] *said just think about it, if you keep on working until you can't work no longer, he said, you won't be able to enjoy your retirement either because you'll be on*

burn out. So, he is right, in that respect, so, I suppose I have to pick a happy medium where I finish work on a high, sorta thing, but I've still got enough left in me to do things to enjoy myself so… I've got to, sort of, strike a balance really" (P104, RA, female)

Theme 12: Mental Independence (decision-making)

The condition itself was not reported to hinder interviewees' abilities to make their own decision regarding how they live their lives. However, one participant reported the condition takes away her right to make some decisions regarding her career. She explained that her condition prevents her from doing specific fitness tests needed to qualify to do certain tasks. However, despite her willingness and, in her opinion, her ability to complete the tests the occupational health team have said that they are not willing to risk putting her through the tests in case she causes injury. The participant is frustrated by this because she feels that, ultimately, it is her decision to do the fitness tests. The same participant also said that she does not disclose exactly how much her condition affects her at work because she does not want to be put on lighter duties (less physical tasks):

"*I don't tell my boss* [when her condition is bad] *that because they'll put me on light duties and put me in an office, no-one wants that…"* (P101, AS, female)

Theme 13: Independence (not washing or dressing) (Deductive)

Independence describes a person's ability to carry out tasks, excluding washing and dressing and decision-making, without the need to rely on help from others. In the workplace, participants in manual and non-manual jobs frequently stated that they rely on staff members to help them carry out particular tasks including physical tasks, such as heavy lifting, and non-physical tasks, such as operating a stapler.

A couple of interviewees explained that they need to continue working in order to remain financially independent:

'Interviewee: So I was down to my one day a week, it's all gone horribly wrong, at the point at which they diagnosed me, because I was being told, you know, I may need to give up work entirely, urm..and all the other stuff that comes with that
Researcher: and how did that make you feel?
Interviewee: Absolutely dreadful
R: working for you is a big part of your life?

Interviewee: Yeah, it's not so much the working; I think it's the financial independence and it's the feeling that I'm contributing something to urm..you know, the household, living costs, urm… and more than that, it's the structure, I'm not by nature a very structured person, so actually having a job in an office works very well for me' (P108, AS, female)

Interviewees also frequently commented that they struggle to drive themselves to work and many rely on their partners to take and fetch them from work:

"if I have a very bad day then urm…my hus….I don't feel safe to drive, it's too painful for me to change gear and my hands, my left foot is very painful on the clutch and I feel I can't turn my neck very well, turn my head very well because my neck is stiff, and my husband, who has retired, he's older than me, uh he will drive me to work and he will pick me up and bring me home again" (P115, female, RA)

Another interviewee explained her condition has forced her to rely on her husband to help her carry out simple tasks at home:

"[I]can't stand long enough to boil the kettle..urm... so I have a bar stool that I sit on , peeling vegetables it very difficult at times, and if I have to, sort of, lift boiling pots or anything big, my husband has to do it for me, I can't, I can't do that" (P104, RA, female)

Theme 14: Settled and Secure (Deductive)

Interviewees worry about their job security because of the conditions they have. Many live in fear that they may be told that they cannot do the job anymore:

"I've got no aspirations to move anywhere else, urm…I kinda of enjoy my job that much, if my condition does advance in any way I think I would have to start having conversations around working hours, remote working and things like that, possibly, but certainly it's not a job that I want to stop doing or kinda move from…urm… I had to undertake extra training and do all kinds of exams and studying dissertations to get the extra qualification to do what I do…so … I think I would be absolutely devastated and heart-broken if at any point I was told I couldn't do it anymore..urm… but it would be something I would have to consider if working got to a point where I was having more bad days than good days…" (P105, AS, female)

One interviewee stated that she feels she has to work no matter how bad she is feeling because she does not want to lose her job:

"I've tended to work even however ill I am, I've tended to work up until then, because I don't want to jeopardise my job" (P109, RA, female)

Many interviewees have either changed their jobs or have started to retrain for another career that would suit them and their condition better. One interviewee was a qualified and practicing nurse, but her condition forced her to be redeployed to her current position in administration. One interviewee, who worked in human resources for the army, believes that his condition prevented him from continuing to work with the army, and, as a result, has had to start working in a less physical job. Other participants have altered part of their job to make it more suitable for their condition; for example, a pharmacist became a stop-smoking counsellor allowing her to sit for part of the day rather than standing up for the entire day.

There was also the belief that developing the condition when working in a job was better than if you develop the condition when looking for a job. One interviewee explained:

"I developed it once I was working and I think that's the ideal scenario, somebody with RA, or another time I was put at-risk, I did secure another role in the bank, but there was a period that I was looking outside and I went for two jobs, one of them I didn't have to mention it on any of the applications, and I was offered that role, and the other one I was asked on the application form to go into, quite detail, about what it was and how it affected me and I decided to disclose it and I really think that that would've just put them off totally, I didn't even get an interview for that one" (P120, female, RA)

Theme 15: Support (Deductive)

Many interviewees felt that they receive support from their employers with some giving them time off to attend hospital appointments and other guaranteeing flexible working hours. Some employers were reported to go to great lengths help their employees to continue working, especially if they were known to be good workers. For example, an employer of one of the interviewees gave their employee a budget to hire an extra person to help him with the physical tasks of his job. Another participant described the support he received from the army to help him deal with his AS:

"it's a rehab centre for all our, sort of, trauma injuries, so anybody who may have lost a limb or something will go in there to do rehabilitation, they also do a lot of other stuff and that involved a doctor who looks after all the AS patients within the armed forces and so once, every 18 months, I would go back and have a two week rehab programme, which was absolutely fantastic, and so we would have, maybe, 12 or 15 guys on a course, everybody had been diagnosed with ankylosing spondylitis, and that was my first, sort of, 'goodness, lots of people have got this, what is going on?', and we'd spend two weeks doing mobility exercises, hydro-pools, CV, weights, stretching, stretching pool, and then back out on to the streets again, so it was very closely monitored, urm…pretty much the same medication as well, so it was really useful" (P116, AS, male).

One interviewee explained that her condition caused her to have her eye removed. Her employer supported her throughout this time by providing her with 5 months full paid leave to recover. As a result, she says that there is no way she could leave the company for another job because of their kindness. Also, one interviewee explained that the support she gets from her current employer is fantastic, but is fearful of leaving:

"Well, a bit frightened to leave… I'm just not convinced anywhere else would be as accommodating" (P114, RA, female).

One participant described how a change of employer can make a big difference to an individual with RA, AS or PsA working life. The interviewee explained that she applied for a job as a part-time teacher working with up to six children at any one time which suited her condition. Her initial employer was also very supportive and did all she could to accommodate her and her condition. The first employer left and was replaced with someone who was not as accommodating. As a result, the participant's job now involves teaching a full-size class and she feels she is now struggling to complete the work.

Many participants also felt supported by their colleagues through their offers to help him/her complete tasks at work. One interviewee said that he feels particularly supported by his team because they are willing to cover the early hours of his working shift when he is feeling ill.

In terms of other types of external support, one interviewee explained that there was very little support to help working people come to terms with contracting a chronic

condition and how to handle it when wanting to continue to work:

"I searched and searched and searched for support and guidance urm…and…all the doctors charities and all the doctors support are for people with addiction problems, or alcohol problems, or mental health problems…urm…I could only, I had occupational health and to be honest occupational health knew less about my condition than I did urm…and that was it really, my named GP role, they gave me 6 sessions of counselling through occupational health, to get my head round it all but no….nothing….it's like you hit a brick wall, our trade union they offer counselling but it's just very generic, there's no support, the only job stuff and work stuff I was offered was why don't you re-train as.. do law stuff so you can work for XXX and XXX" (P122, PsA, female).

Also, the conditions may increase the probability of employees making mistakes in their work. It is important and helpful to have an employer who understands why mistakes sometimes happen. Employers who do not understand how RA, AS or PsA affects an individual's ability to work may offer them job opportunities that are unrealistic for the employee:

"I did have one manager, one regional manager, going back a while ago…urm… and a… I said… I got offered an opportunity and whatever else and I said 'nah, I can't really do that' and ur…I think that's, physically, going to be a bit too much for me, urm… and then I got an email back and I thought you haven't really read what I have said and…you know… just glossing over it and not really understanding how it impacts on me" (P106, AS, male)

5.4.6 Summary of themes captured by the outcome measures

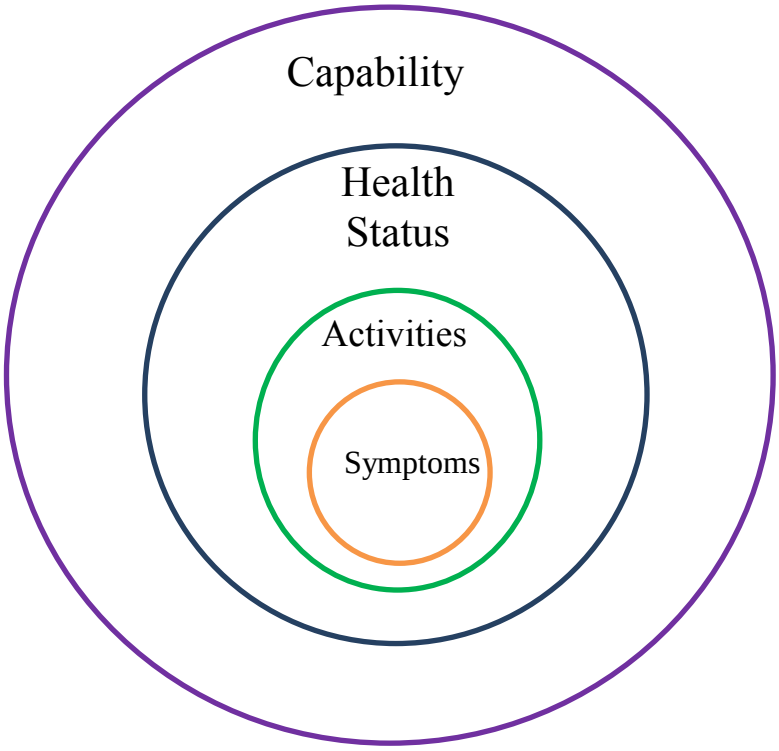
The SF6D captured two aspects that the EQ5D does not that affect levels of presenteeism, the two aspects were: 1) vitality; and 2) social functioning. The ICECAP-A, because of its broad design, was found to capture all aspects of the conditions that affected employees ability to work. Table 5.6 lists those factors beyond health status that were captured by the ICECAP-A that were reported to impact an individual's ability to work. For example, 'Achievement and Progress' was able to capture all symptoms and activities, including mental clarity, that were reported to have an impact on presenteeism. The same capability was also able to capture the impact conditions have on the ability to progress and obtain higher level jobs.

Table 5.6 Factors beyond health captured by ICECAP-A which impact levels of presenteeism

ICECAP-A Capabilities	Factors beyond health status captured by ICECAP-A
Achievement and Progress	Ability to progress in career
Enjoyment and Pleasure	Ability to enjoy working
Independence	Ability to complete work tasks independently
Love, Friendship, and Support	Impact on relationships with colleagues
	Ability to gain support from employer
Settled and Secure	Ability to continue working and avoid job loss

Figure 5.2 illustrates a potentially useful way of interpreting how symptoms, activities, health status, and capability are connected. Health-related symptoms of a condition affect an individual's ability to carry out certain activities. Symptoms and activities, together, make up health status, as defined by the EQ5D and SF6D health status measures. Finally, capability includes health status as one dimension that influences an individual's ability to obtain a high quality of life (Lorgelly 2015).

Figure 5.2 The relationship between symptoms, activities, health status and capability



5.5 Discussion

This study provides qualitative evidence of the face validity of the ability of three selected measures of HRQoL (EQ5D, SF6D, and ICECAP-A) to capture the impact RA, AS and PsA has on levels of presenteeism. In general, the interviewees explained that symptoms caused by the conditions affect their ability to complete activities which ultimately affects their career in broader terms, for example the ability to progress.

The EQ5D and SF6D were able to capture symptoms caused by RA, AS, and PsA and the impact they have on the ability to perform activities. The EQ5D and SF6D are sensitive measures that capture specific factors (symptoms and activities) that cause presenteeism. The SF6D, compared to the EQ5D, is arguably a more suitable measure because it explicitly takes into account fatigue. Fatigue was reported as the symptom that was the most difficult to work with and adapt to indicating that it is an important factor that may substantially influence levels of presenteeism. Nevertheless, the EQ5D and SF6D were both unable to capture the impact 'mental clarity' had on ability to work. The interviewees explained that poor mental clarity prevents them from working quickly and increases the amount of mistakes they make, ultimately increasing levels of presenteeism. Evidence, from another qualitative study that explored the impact of RA on work performance also concluded impaired mental clarity (termed mental cognition

in the study) and was a key factor that negatively impacts work performance in terms of number of mistakes made, the ability to concentrate and think quickly (Connolly et al. 2015). It may, therefore, be argued that 'mental clarity' is an important factor in determining levels of presenteeism and consequently the EQ5D and SF6D may be hindered in their ability to accurately capture presenteeism caused by RA, AS and PsA; although, this hypothesis would need to be tested empirically.

In contrast to the EQ5D and SF6D, the ICECAP-A was identified as a broad measure that was able to capture symptoms, activities, and factors beyond health that impact presenteeism as a result of working with one of the three conditions. The ICECAP-A captures broader aspects associated with an individual's ability to work, for example, the ability to progress or maintain a stable and secure job. The results of this study suggested that the ICECAP-A was able to capture many factors that determine levels of presenteeism; however it was not able to capture the extent to which symptoms of RA, AS, and PsA, and their influence on the ability to perform activities, has on levels of presenteeism. The measures of health status and capability are shown to capture different factors that determine that levels of presenteeism. The evidence here suggests that, assuming all three measures were able to adequately capture the impact on presenteeism, it would be beneficial to use one measure of health status and a measure of capability to capture the full impact of RA, AS and PsA has on presenteeism. One measure of health status and one measure of capability may be viewed as complements rather than tools that are mutually exclusive.

The results of this study were in-line with those reported in the wider literature. A systematic review collated qualitative evidence of the factors that challenge employees with IA's ability to remain employed or return to work (Hoving et al. 2013). The review identified ten qualitative studies, eight of which focussed on RA, one on AS, and one on IA. The review identified seven key challenges (themes) employees with inflammatory arthritis experience when working:

1. Symptoms and disease effects on work;
2. Managing arthritis and consequences
3. Social/economic factors (e.g. job insecurity, support, part-time work opportunities);
4. Working conditions and adaptations (e.g. support from employer);
5. Emotional challenges (e.g. anxiety, guilt, self-confidence);

6. Interpersonal issues and choices affecting work and family life (e.g. relationships with colleagues, family and work life balance);
7. Meaning of work (e.g. work important to well-being).

The experiences shared by the participants working with RA, AS, and PsA, in this study, provides more evidence of the importance of taking into account not just symptoms and how these affect an employee's ability to do certain activities, but to incorporate the impact wider factors have on ability to work.

Double-Counting of Productivity in Incremental Cost Effectiveness Ratios

The original Washington Panel stated that there was a possibility that productivity was captured in preference-based measures, such as the EQ5D, because respondents consider the effect productivity has on their income when valuing their health states (see chapter two, section 2.2.7). An update of the Panel's reference case now states that, given the evidence, income effects are unlikely to be taken into account when patients value their health-states (Sanders et al. 2016); this may be particularly relevant in the context of presenteeism where the vast majority of employees' are still paid meaning they are unlikely to experience a decrease in their level of income because of presenteeism. It may, however, be argued that individuals take into account their ability to work when valuing health states. For example, someone who is required to walk a great deal for their job and is experiencing problems may value their mobility level as more severe compared to someone with the same mobility problems who does not have to walk for their job. It was clear from the findings in this study that an individual's job is an important aspect of peoples' lives, it may, therefore, be justified to assume that individuals take into account their ability to work when filling out their health states. If this assumption is correct then there is potential for productivity to be double counted through self-report presenteeism instruments and preference-based measures as a cost in the numerator of the ICER, using self-reported measures of presenteeism, and also as a consequence in the denominator, using a self-reported measure of outcome, potentially leading to double-counting. Further research is needed to explore this assumption in greater detail. One way could be to conduct studies involving think-aloud methods. Think-aloud methods are used to explore what individuals think or understand about a certain subject (Dillman 2006). In this case, think aloud methods may be used to find out what individuals consider when they rate their health status or capability. The

Washington Panel also advocates further research on the extent to which productivity effects health state evaluations and therefore QALYs (Sanders et al. 2016).

5.5.1 Strengths

This study contributes to the literature in four important ways. This study is the first to investigate whether available preference based measures have sufficient face validity to capture presenteeism. The results from this study indicate preference based measures of health and capability are able to capture the impact on presenteeism caused by RA, AS, or PsA. However, it is important to note that measures of health status and capability capture different factors that determine levels of presenteeism.

This is the first study to have analysed and identified the concepts that overlap across three measures of HRQoL. The ICECAP-A was found to have the ability to incorporate all concepts included in the EQ5D and SF6D. The result confirms the ICECAP-A as a broader measure of quality of life with health forming one essential part needed to attain a good standard of quality of life.

The evidence provided by this study may contribute towards the debate concerning the potential for double-counting of productivity in the ICER; not as income, as was described by The Washington Panel, but as the ability to work and how that potentially affects the way in which individual's rate their health status.

5.5.2 Limitations

The evidence generated from this study cannot be used to inform normative arguments (see chapter two, section 2.1) regarding what measure ought to be used to capture the impact of RA, AS, and PsA on productivity. The results address the positive stance suggesting that the SF6D is, potentially, more likely to be a suitable measure to capture presenteeism compared with the EQ5D, and that the ICECAP-A is able to capture presenteeism beyond the effects of symptoms and limited ability to perform activities.

The interviews had to be conducted over the telephone rather than face-to-face for logistical reasons. There is potential that further information given by a participant's body language may have been missed which may have added more depth and understanding to the data collected but such data are unlikely to have made a difference to the main findings and conclusions of this study.

A potential opportunity was missed to ask each participant to answer the EQ5D-5L, SF6D, and ICECAP-A. A comparison of an individual's qualitative responses and their valued health states may have provided some insight about how individuals interpret their situation and rate their current health status.

The participants interviewed for this study worked in a wide and varied selection of jobs; however, only five participants worked in manual jobs which may limit generalisability of the results to non-manual based work only.

The study was only able to recruit and interview a small number of males, which may limit generalisability of the impact RA, AS or PsA has on men and their ability to work. However, the responses from males compared with females within this study were similar; therefore, there is little reason to believe that the results from this study would have been substantially different had more males been included.

5.6 Conclusions

This study has explored how RA, AS and PsA limits an employee's ability to work in terms of completing tasks and progressing in their career. All three measures had face validity in terms of being able to capture the impact on ability to work caused RA, AS or PsA. The EQ5D and SF6D were identified as sensitive measures able to capture the impact of symptoms and activities had on levels of presenteeism. The ICECAP-A, on the other hand, was identified as having the ability to capture the impacts RA, AS and PsA may have on an individual's ability to work and career. Regression based research is needed to explore the extent to which each measure can capture presenteeism. The next chapter investigates the relationship between health status (EQ5D; SF6D) and capability (ICECAP-A) with presenteeism using econometric methods.

Chapter 6 To develop a prediction model for presenteeism using existing measures of health status and capability

6.0 Introduction

One objective of this PhD was to develop a prediction model for presenteeism using measures of health status and capability (see chapter one, section 1.1.5). A study designed to address this objective is now presented in chapter six. The first section of this chapter (section 6.1) provides a summary of the existing evidence that has explored the relationship between health status, or capability, and presenteeism. Section 6.2 presents the aims and objectives of the study. Section 6.3 describes the methods, sections four and five presents the results and discussion, respectively. The chapter ends with a conclusion (section 6.6)

6.1 Background

Workplace interventions (WPIs) are a set of programmes, activities and equipment designed to help individuals manage their condition in terms of improving symptoms and functional ability (Public Health England 2014b). Employers, frequent funders of WPIs, have limited access to information of the costs and productive benefits WPIs may produce. Economic evaluations (see chapter two, section 2.1), used to compare the costs and consequences of healthcare interventions (Drummond et al. 2015), may also be used to produce evidence of the costs and consequences of WPIs to inform funding decisions made by employers.

Existing methods of presenteeism have predominately focussed on valuing it as a cost by multiplying the volume of presenteeism by the wage (human capital approach (HCA) (see chapter two, section 2.3.6) (Kigozi et al. 2017b). In a systematic review of methods used to measure and value presenteeism, Mattke and colleagues (2007) reported that there is no established and validated method that can be used to derive the monetary value of the cost of lost productivity. Since publication of the Mattke et al's systematic review, little progress has been made to develop and validate methods that can be used to monetise the value of presenteeism (see chapter 4, Kigozi et al. 2017). The HCA is most often used to value presenteeism as a cost, a method which understood to overestimate the value of presenteeism (Koopmanschap et al. 1995; Zhang et al. 2012; Pauly et al. 2008). Pauly et al. (2008) state that current estimates of

the value of presenteeism have failed to be accepted by employers and they remain sceptical of the magnitude of the cost of lost productivity.

Also, from the perspective of conducting an economic evaluation including productivity as a cost or monetary benefit, valued using wages, is problematic because it may influence cost-effectiveness decisions towards those who are paid high salaries, causing potential equity issues that employers would wish to avoid (Olsen and Richardson 1999). Finally, employers have strong incentives to reduce lost productivity of their workforce (Hemp 2004) and therefore, from an extra-welfarist perspective, it may be argued that productivity be chosen as the maximand, in the same way health is set as the maximand for the healthcare sector (see chapter 2 section 2.2.2). There is a need to develop a method that avoids valuing presenteeism as a cost using wages.

The results from chapters three and four of this thesis summarised the existing methods used to quantify the impact WPIs have on productivity and concluded they are inadequate to provide robust evidence for employers who fund WPIs. Therefore, this study focused on developing a method, consistent with extra-welfarist views, that values presenteeism as a consequence using non-monetary measures suitable for use in economic evaluations of WPIs from the perspective of the employer. Measures of health status, such as the EuroQol Five Dimensions (EQ5D) (EuroQol Group 1990a), and capability, such as ICECAP-A, are used in economic evaluations to value the consequences of healthcare interventions. A potential way to measure presenteeism would be to develop a model that predicts levels of presenteeism given an individual's health related qualify of life. A prediction model for presenteeism using measures of HRQoL may be useful for two reasons. First, WPIs are designed to improve an individual's ability to manage their health condition and subsequently improve their ability to work (see chapter 1, section 1.3). A prediction model for presenteeism using HRQoL data would allow the employer to monitor their employees' health and presenteeism levels. Second, a point which is directly related to the first and more practical from the employers' perspective, it would be very difficult to ask employees to fill out a self-reported measure of presenteeism honestly for fear of being judged and potentially losing their job. Therefore, employers may be more inclined to ask their workforce about their health and indirectly link this to levels of presenteeism. In practice, the implementation of such a prediction model would require some thought and possible legislation regarding how information should and should not be used

offering some protection for employees; this issue is beyond the scope of this PhD, however it is important and should be stated.

Existing Research

Limited research has been carried out to formally estimate the associated relationship between HRQoL and presenteeism. Brouwer et al. (2002) initially explored the extent to which a relationship between health status and productivity existed using a sample of 51 Dutch employees working in a trade firm. The employees were asked to complete a survey which recorded their self-reported levels of presenteeism and health status. Presenteeism was measured using the quantity and quality (QQ) method (see chapter three, section 3.4.7.1); a self-reported questionnaire that asks the respondent to, separately, score the quantity and quality of work they have produced using a zero to ten Likert scale. Health status was measured using the EQ5D visual analogue scale (VAS), a measure of general health where the respondent is required to state their level of health using a scale from 0 to 100 where zero presents worst health and 100 represents best health. Forty-one (80%) respondents with an average stated EQ5D-VAS score of 81 stated that they did not experience presenteeism in terms of the quantity and quality of work they produced. Six (12%) respondents scored an average of 69 and reported increased levels of presenteeism upon return to work. Four respondents with an average EQ5D-VAS score of 80 reported decreased levels of presenteeism when they had first returned to work after a period of absence but after some time described productivity levels as having increased to levels before they were absent from work. The results indicate a potential relationship between health and productivity; however, the authors stated that further research using larger sample sizes was needed to confirm the relationship.

Two studies have since empirically estimated the associated relationship between health status, measured using preference weighted surveys (EQ5D-3L; three level) and presenteeism (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014). In 2005, Lamers et al., collected health status and presenteeism data as part of a cluster randomised control trial (RCT) designed to evaluate physiotherapy guidelines for low back pain in a sample of 483 Dutch patients. Health status was measured using the EQ5D-3L (applying UK preference weights), absenteeism using the Health and Labour Questionnaire (HLQ) (L. van Roijen et al. 1996) (see chapter three, section 3.4.7.1); and presenteeism using the quality and quantity (QQ) method (Brouwer, Koopmanschap, and Rutten 1999) (see

chapter three, section 3.4.7.1). The questionnaires were completed at baseline, 6 weeks and 12 weeks follow-up. Presenteeism results were reported as efficiency losses (see chapter three, section 3.4.7.1, table 3.7).

Lamers et al., (2005) identified individuals who had been absent from work over the past two weeks reported mean EQ5D-3L scores of 0.48 whereas those who were not absent from work reported mean EQ5D-3L scores of 0.71, a significant difference of one standard deviation. Individuals without absence from work reported a mean efficiency loss of 0.2 (standard deviation 0.25). Individuals who reported their last day at work after a period of absence had a mean EQ5D-3L score of 0.62 (standard deviation 0.24) with a mean efficiency loss of 0.38 (standard deviation 0.3.) Lamers et al., (2005) also formally tested the correlation between the EQ5D-3L and absence from paid and unpaid work. The correlation coefficients, assessed using Spearman's rank correlations, were low (0.25) to moderate (0.40) (see section 6.3.4). Lamers et al., (2005) concluded that the correlation coefficients for the EQ5D-3L and absenteeism were too low to recommend the development of a prediction model for productivity costs associated with low back pain. Analysis of the correlation between the EQ5D-3L and presenteeism was not examined.

A more recent study, conducted in 2014, aimed to develop prediction models for absenteeism and presenteeism using the EQ5D-3L as the measure of health status (Krol, Stolk, and Brouwer 2014). The aim was to develop a model to enable the prediction of productivity costs associated with a given health state. Using an online questionnaire, 1,013 employed members of the Dutch population were asked to state their *expected* level of productivity based on different health states selected from the EQ5D-3L (specific population preference weights used was not disclosed in the study). Using block design, six subsets of 16 different EQ5D-3L health states were defined and each respondent was assigned to one of the six subsets. A prediction model for presenteeism was estimated using generalised estimating equations (GEEs). Using the Quasi Likelihood under Independence Model Criterion (QIC) and the Corrected Quasi Likelihood under Independence Model Criterion (QICC), a negative binomial distribution with a log link was demonstrated to be the best fitting model. The dependent variable of the prediction model was presenteeism measured as the expected level of functioning and was measured using the quantity question of the QQ method (see chapter three, section 3.4.7.1, Table 3.7). Covariates included in the prediction model were age and gender.

Krol et al's., prediction model was validated using an external dataset collected by Lamers et al., (2005). The average level of functioning when attending work was 0.75 and the predicted was 0.66 (zero represents unable to complete work and 10 represents normal functioning). The prediction model was found to overestimate levels of presenteeism which may have been caused by the hypothetical nature of the study. Krol et al., asked respondents to imagine being in 16 different health states and then provide information on their expected levels of presenteeism. Asking individuals to imagine being in 16 different health states is cognitively challenging and the respondents may have found it difficult to accurately predict what their own levels of presenteeism might be. For example, one factor which may lead to under or overestimation of presenteeism is the respondents' potential ability to comprehend how they might, or might not, adapt their working style to suit their condition.

The results from Krol et al., (2014) study also suggested that females were expected to be less productive compared to their male counterparts. The result is questionable due to systematic differences in reporting made by males compared to females. It is widely understood that males are more likely overestimate their abilities and women are more likely to underestimate their abilities (Visser, Ashton, and Vernon 2008; Lenney 1977) which may help to explain why, in hypothetical circumstances, women appear to be potentially less productive compared to men.

The two studies that have examined the relationship between HRQoL and presenteeism provide a positive indication that there may be a quantifiable relationship between HRQoL and presenteeism (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014). The most substantive limitation of the literature at present is that it has only examined the relationship and predictive ability of one measure of health status (EQ5D-3L) and productivity (the QQ method). Other HRQoL measures, such as the Short-Form Six Dimensions (SF6D) (Brazier, Roberts, and Deverill 2002); or the ICEpop CAPability measure for Adults (ICECAP-A) (Al-Janabi et al. 2012), and alternative measures of presenteeism may exhibit a stronger relationship allowing for more accurate predictions of presenteeism. Existing evidence has focussed on quantifying and developing prediction models for presenteeism using Dutch populations with IA. It is, therefore, necessary to investigate whether the relationship between HRQoL and presenteeism exists within a UK population and whether other HRQoL measures can be used to predict presenteeism. Importantly, the prediction model for presenteeism developed by Krol, Stolk, and Brouwer (2014) is based on hypothetical expectations of presenteeism

given different health states rather than data collected from individuals living and working with specific health conditions. The current evidence is also not generalisable to rheumatoid arthritis (RA) or ankylosing spondylitis (AS) populations.

6.2 Aims and Objectives

The aim of this study was to develop a prediction model for presenteeism using measures of HRQoL, using three standardised measures (EQ5D-5L; SF6D and ICECAP-A).

To meet this aim, four objectives were specified:

1. To describe the relationship between HRQoL and presenteeism;
2. To develop prediction models for presenteeism and health status, or capability, measured using standardised measures;
3. To assess the performance of each prediction model;
4. To assess whether average predictions of presenteeism can be applied across specific subgroups of the population.

6.3 Methods

This study used regression-based methods to analyse data collected from a cross sectional survey of working individuals with RA or AS. The objective of the methods used in this study was to predict levels of presenteeism using measures of HRQoL. It was not feasible, within the time frame of this thesis, to collect a longitudinal dataset to address questions about the causal relationship between HRQoL and presenteeism.

6.3.1 Measures of presenteeism, health status and capability

This section reports the methods used to quantify the relationship between presenteeism and health status or capability. A description of the standardised measures selected for this study is explained in chapter 5, section 5.3.1.

6.3.1.1 The selection of measure of presenteeism

Chapter three of this thesis identified 24 self-report presenteeism instruments suitable for use in RA and AS populations. The Short-Form Health and Labour Questionnaire (SF-HLQ) (L. van Roijen 2010) (not included in chapter three because it was identified as a modified survey from the original Health and Labour Questionnaire) was also considered because it is shorter than the original HLQ. The selection of the most appropriate self-reported presenteeism measure for this study involved judging all 25 presenteeism measures against pre-specified inclusion/exclusion criteria listed in Table

6.1. The inclusion criteria were informed by the pragmatic requirements of the design of this study.

Table 6.1 Inclusion criteria to inform the selection of the most appropriate presenteeism instruments

Inclusion Criteria	Description
Numeric Score	Measured using quantitative descriptions
Non-time units	Measures presenteeism using a scale or percentage
Length of Survey (number of questions asked)	Survey that asks more than 10 questions and cannot be used without excluding irrelevant questions

Given one objective of this study was to predict levels of presenteeism it was necessary to use a measures that produced a single numerical score. Some presenteeism instruments use qualitative statements such as ‘agree’ or ‘disagree’, or ‘a little’ or ‘somewhat,’ and were subsequently excluded from the study (see Appendix 6.1). It was important to select a measure of presenteeism that did not use time related units. Measuring presenteeism as a unit of time may have potentially introduced double-counting of the patient or employees time if the method were to be used alongside a cost-utility analysis. HRQoL is measured using QALYs where an individual’s remaining life time is measured in years (see chapter two, section 2.1.4); therefore, incorporating a patient’s working time as a result of the impact of a WPI would potentially lead to the double-counting of time through: 1) the number of years of life (time) estimated in the QALY; and 2) the amount of time spent working. To prevent the overall survey from becoming too long and burdensome for the respondent the presenteeism measure selected had to be short using ten questions or fewer. Table 6.2 presents the inclusion/exclusion results of each self-report presenteeism instrument considered for use in this study.

Table 6.2 Presenteeism instruments appropriate for this study

Presenteeism Instrument	**Numeric score Yes (Y) / No (N)**	**Non-Time Units used Yes (Y) / No (N)**	**Length of survey (≤ 10 questions) Yes (Y) / No (N)**	**In/Out**
The Functional Status Questionnaire/ Work Performance Scale (WPS) Jette et al (1986)	N	N	N	Out
Osterhaus Technique (OT) Osterhaus et al (1992)	Y	Y	N	Out
Work Productivity and Activity Impairment Instrument (WPAI) Reilly et al (1993)	Y	Y	Y	In
Health and Labour Questionnaire (HLQ) van Roijen et al (1996)	Y	N	N	Out
Endicott Work Productivity Scale (EWPS) Endicott et al (1997)	N	N	N	Out
The Quality and Quantity Method (QQ) Brouwer, Koopmanschap and Rutten (1999)	Y	Y	Y	In
Work Productivity Index (WPI) Burton et al (1999)	Y	N	Y	Out
Work Role Functioning Questionnaire (WRFQ) Amick et al (2000)	N	N	Y	Out
The Work Limitations Questionnaire (WLQ) Lerner et al (2001)	N	Y	Y	Out
Stanford/American Health Association Presenteeism Scale (SAHAPS) Pelletier and Koopman 2001	Y	Y	N	Out

Stanford Presenteeism Scale – 6 (SPS-6) Koopman et al (2002)	N	N	Y	Out
Work Instability Scale (WIS) Gilworth et al (2003)	N	N	N	Out
Work Productivity Short Inventory (WPSI) Goetzel et al (2003)	Y	N	Y	Out
The World Health Organisation Work Performance Questionnaire (HPQ) Kessler et al (2003)	Y	Y	N	Out
Health-related Productivity Questionnaire – Diary (HRPQ-D) Kumar et al (2003)	Y	Y	Y	In
Health and Work Questionnaire (HWQ) Shikar et al (2004)	Y	Y	N	Out
*Work Health Interview (WHI) Stewart et al (2004)	-	-	-	-
Stanford Presenteeism Scale – 13 (SPS-13) Turpin et al (2004)	Y	Y	N	Out
Work Ability Index (WAI) Ilmarinen et al (2007)	Y	Y	N	Out
Work Productivity Survey for Rheumatoid Arthritis (WPS-RA/AS) Osterhaus, Purcaru and Richard (2009)	Y	Y	Y	In

Short Form Health and Labour Questionnaire (SF-HLQ) van Roijen et al (2010)	Y	Y	N	Out
* Well-being Assessment for Productivity (WBA-P) Prochaska et al (2011)	-	-	-	-
The Valuation of Lost Productivity (VOLP) Zhang et al (2012)	Y	Y	N	Out
Composite Work Functioning Approach Boezeman et al (2015)	N	Y	Y	Out
iMTA Productivity Cost Questionnaire (iPCQ) Bouwmans et al (2015)	Y	Y	N	Out

*The WHI and WBA-P were unavailable

After applying the inclusion criteria, four presenteeism instruments were identified as potentially suitable measures for use in this study. The four measures included the: WPAI; QQ method; HRPQ-D; and WPS-RA. All four measures were explored in more detail and narrowed down to two. The HRPQ-D was excluded based on requirements for completion. The HRPQ-D requires the respondent to complete the survey every day for one working week. The design of the HRPQ-D was incompatible with the design of this cross-sectional study and was therefore excluded. The WPAI and the WPS-RA (or AS) have been recommended by OMERACT for measuring lost productivity at work caused by rheumatic conditions (Beaton et al. 2016). The WPS-RA was excluded because it adopts a clinician perspective, a perspective that was not consistent with people adopted by the survey which used a patient perspective. For this reason, the WPAI was selected over the WPS-RA (or AS). The WPAI was selected for this study because it met the inclusion criteria, used a patient perspective, and was recommended as a measure of presenteeism for individuals with rheumatic conditions. The QQ method, whilst not recommended by OMERACT, was included because it had been used to quantify the link between presenteeism and HRQoL in previous studies (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014).

6.3.1.2 Presenteeism Questions

Figure 6.1 presents the questions and scales used to measure presenteeism in the WPAI and QQ method surveys. The WPAI records levels of presenteeism using a zero to ten Likert scale where zero indicates 'Rheumatoid Arthritis (or Ankylosing Spondylitis) had no effect on my work' and ten indicates 'Rheumatoid Arthritis (or Ankylosing Spondylitis) completely prevented me from working'. The QQ method also uses zero to ten Likert scales to measure productivity except zero indicates completely unable to work and ten indicates able to work as normal. The WPAI and the QQ method were registered for use in this study by contacting the originators of each measure.

Figure 6.1 Presenteeism questions included in the two selected presenteeism measures

WPAI

Consider only how much Rheumatoid Arthritis (or Ankylosing Spondylitis) affected productivity while you were working.

Rheumatoid Arthritis (or Ankylosing Spondylitis) had no effect on my work	0	1	2	3	4	5	6	7	8	9	10	Rheumatoid Arthritis (or Ankylosing Spondylitis) completely prevented me from working

QQ Method

1. Could you indicate how much work you actually performed today during regular hours as compared to normal on the scale below?

Practically nothing	0	1	2	3	4	5	6	7	8	9	10	Normal quantity

2. Could you indicate the quality of the work you performed today as compared to normal on the scale below?

Very poor quality	0	1	2	3	4	5	6	7	8	9	10	Normal quality

3. Please circle on the scale below the degree of efficiency you consider yourself to have worked with on the days you did go to work during the past week while suffering from rheumatoid arthritis (or ankylosing spondylitis).

Very inefficiently	0	1	2	3	4	5	6	7	8	9	10	As efficient as normal

6.3.2 Data collection

A cross-section study design was used and data were collected accordingly using a bespoke survey. This section describes the relevant study population, sampling frames and the data collection methods used.

6.3.2.1 Study Population

The relevant study population were employed individuals with RA or AS. Individuals eligible for inclusion into this study had to meet the following criteria:

1. Participants must have been diagnosed with RA or AS by a rheumatologist;
2. Participants must be employed (including self-employed) in paid work on either a full-time or part-time basis;
3. All participants must be at least 18 years old.

6.3.2.1 Sampling frames

Two strategies were used to recruit samples of employed individuals with RA and AS (i) an internet panel provider and (ii) a patient support organisation. The internet panel provider, ResearchNow, was contracted to deliver completed surveys from individuals who matched the inclusion criteria set for this study. ResearchNow (ResearchNow 2017) specialise in data collection by approaching registered individuals and inviting them to complete surveys in return for a small fee. The advantage of using an internet panel provider is their ability to quickly identify a large number of individuals who are able to complete surveys allowing time spent collecting data to be kept to a minimum (within two weeks). This recruitment strategy has been successfully used in previous studies (Vass, Rigby, and Payne 2017). Appendix 6.2 provides further information regarding the recruitment process used by ResearchNow.

A sample of 550 working individuals with RA was obtained via ResearchNow. ResearchNow charge the client a fee per survey successfully answered. A maximum sample of 550 individuals working with RA was recruited. A sample size of 550 was considered acceptable based on previous studies that have attempted to quantify the impact of presenteeism and other mapping studies. Lamers et al. (2005) used a sample of size 483. Previously published mapping studies, explored using the HERC database of mapping studies (Dakin 2013), that have mapped health surveys to preference weight health measures, such as the EQ5D, in patient populations with RA used a sample size as small as 206 (Carreño et al. 2011).

Individuals with AS were recruited through a support group called The National Society for Ankylosing Spondylitis (NASS). The study was advertised on the NASS website with the aim of recruiting a sample size of 100 working individuals with AS. Individuals who responded to the AS survey were not remunerated for their time.

A different recruitment method for individuals working with AS was adopted for two reasons: 1) the response from AS working individuals via NASS for the qualitative study was very high indicating a strong level of interest from employed people with AS to participate in research regarding AS and work; and 2) it was not possible for financial reasons to recruit more working individuals through ResearchNow.

6.3.3 Data Collection

An online bespoke survey was designed to ask working individuals about their HRQoL and levels of presenteeism. Two surveys were designed for: 1) people with RA; and 2) people with AS. The only difference between the two surveys was the disease being referenced where the RA survey asked questions specifically referencing RA and the AS survey asked questions specifically about AS (see Appendix 6.3).

6.3.3.1 Survey design

The survey was designed using information from chapters three and five of this thesis and published reviews of the determinants of presenteeism caused by RA and AS. Chapter three was used to identify relevant presenteeism instruments that could be used to quantify the impact of presenteeism caused by RA or AS. Chapter five highlighted the importance of asking questions about the individual's job type and whether or not they had disclosed their condition to their employer. Other published reviews regarding the determinants of presenteeism caused by RA or AS informed questions regarding their: 1) characteristics including their gender, age and level of education; 2) contextual factors regarding the type of job they have; and 3) medication they currently receive (Verstappen et al. 2004; de Croon et al. 2004; van Vilsteren et al. 2015).

The online survey comprised five sections. Section one explained the purpose of the study, provided contact details for respondents who may have had further questions, and confirmed the respondents consent to take part in the study. Informed consent was taken before the respondent completed the survey. Respondents were informed that they could

leave the survey at any point if they decided they no longer wished to participate; however, it was also explained that once they clicked 'submit' at the end of the survey it would not be possible to withdraw their answers from the study.

The survey began with two screening questions used to confirm the respondents' suitability for inclusion into the study: 1) have they been diagnosed with RA or AS by a rheumatologist (doctor); and 2) they were currently working in paid employment on a full-time or part-time basis. It was confirmed through expert advice (personal communication: Professor Anne Barton, Consultant rheumatologist, The University of Manchester, 2017) that individuals cannot have both RA and AS. Only those participants who answered 'yes' to both questions were permitted to complete the full online survey. Those who did not match the criteria were thanked for their interest in taking part and told that because they did not meet the requirements of the study were unable to continue.

The main part of the survey was made up of four sections requiring information about the respondents: individual characteristics; job characteristics; severity of RA or AS; measures of HRQoL; and levels of presenteeism (Appendix 6.3).

Section two: Individual Characteristics

This section asked individuals' to provide information about their demographics including age, gender, level of education and ethnicity.

Section three: Job Characteristics

This section asked respondents to provide details of their current occupation in terms of their job title, sector, employment status, average number of days worked per week, average number of hours worked per week, type of job tasks they complete, whether they work indoors or outdoors, the type of profession they work in, size of the company they currently work for, whether their colleagues are able to help them complete work, and level of income.

A list of job sectors was put together using two sources: 1) the PriceWaterHouseCooper list of industries (PWC 2017); and 2) the UK Standard Industrial Classification of Economic Activities (ONS 2007). Two sources of lists of UK sectors were used to ensure that both private and public sector industries were taken into account. Both lists were cross-referenced to avoid overlapping sectors. Employment status was presented using four options: 1) full-time employee; 2) part-time employee; 3) self-employed; 4)

other, please specify. The average number of days and hours individuals worked per week was included to more fully understand an individual's work situation. Job tasks were broken down into four categories using the definition for sedentary to heavy manual work provided by the Social Security Administration based in the United States of America (SSA) (SSA 2017). The four categories for non-manual and manual work were defined as follows: 1) sedentary, defined as lifting 10 pounds or less (4.5 kg), and walking or standing for less than two hours per eight hour work day; 2) light work, defined as often lifting 10 pounds (4.5kg) and occasionally lifting 20 pounds (9kg) and standing or walking for six hours of an eight hour work day; 3) medium work, defined as constant standing or walking and kneeling, squatting, climbing and often lifting 25 pounds (11kg) and sometimes lifting 50 pounds (23kg); and 4) heavy work, defined as constant standing or walking and kneeling, squatting, climbing and lifting up to or more than 100 pounds (45 kg). A question about whether the individual worked indoors or outdoors was included because it was expected that those worked predominantly outdoors were affected differently compared with those who worked indoors. Three options were given: 1) outdoors; 2) indoors; and 3) 50% indoors and 50% outdoors.

A list of professions was presented based on the Standard Occupational Classification 2010 (ONS 2010). The size of the company the respondent worked for was asked to understand whether the size of the company had an indirect effect on their ability to work. Large companies are more likely to employ individuals who do the same job; therefore colleagues may be able to help complete work (see chapter 2, section 2.3.4). Five different company size definitions were presented: 1) micro business, fewer than 10 members of staff; 2) small, more than 10 and fewer than 50; medium, more than 50 and fewer than 250 member of staff; and 4) large, over 250 members of staff (European Commission 2017). The next question then asked respondents to state whether colleagues were able to help them complete their work if necessary, the responses included: 1) yes, other colleagues can help me with my work load; and 2) no, I am the only person who can complete the work I do. Finally, respondents were asked to indicate of their level of income. Income was categorised into nine bands increasing in increments of £10,000 up to £80,000 and more.

Section four: Health Conditions

This section asked questions about the participant's health condition; RA or AS. Disease severity was measured using the Routine Assessment of Patient Index Data three (RAPID3) questionnaire (Andersen et al. 2012).The RAPID3 questionnaire was

chosen because it has demonstrated its validity for use in both RA (Hendrikx et al. 2016) and AS populations (Danve et al. 2015). The RAPID3 was also recommended for use in this survey by an expert in the field (personal communication Dr. Mark Harrison, The University of British Columbia, 2017). The RAPID3 contained 15 questions. The first thirteen questions asked the individual to state their ability to complete different types of activities from a set of responses: 1) without any difficulty; 2) with some difficulty; 3) with much difficulty; and 4) unable to do. The fourteenth and fifteenth questions ask the respondent to state how much pain they have experienced and how much the condition has affected them over the past week; both questions are answered using a Likert scale from 0 to 10, increasing by increments of 0.5 where zero indicates 'no pain' and they are coping 'very well' and 10 indicate pain is as 'bad as it could be' and they are coping 'very poorly'. The final score of the RAPID3 was calculated by totalling the scores from the first ten questions, converting the score onto a zero to ten scale using pre-defined weights, and then adding the scores from the pain and overall affect scales. Each RAPID3 score was categorised into one of four levels of severity: 1) remission; a RAPID3 score of less than 3.1; 2) low severity, a score between 3.1 and 6.0; 3) medium severity, a score between 6.1 and less than 12; and 4) high severity, a score of 12 and above (Pincus, Yazici, and Bergman 2009).

Section five: measures of HRQoL

This section asked respondents to record their overall health status, using the EQ5D-5L and the SF6D, and capability, using the ICECAP-A.

Section six: Productivity

This section asked respondents to complete questions about how and to what extent RA and AS affects their ability to work. The primary aim of this section was to measure presenteeism using self-reported instruments suitable for use in RA and AS. The questions used to measure presenteeism are discussed in detail in section 6.3.1.3.
The overall survey also included some additional questions related to productivity. Respondents were asked to state whether they were currently on sick leave, and if yes on what date did their current sick leave period begin and for how many days it lasted. This information was collected to confirm individuals were reporting levels of presenteeism within the recall periods stated.

The WPAI also included a question related to absenteeism: 1) 'During the past 7 days, how many hours did you miss from work because of RA (or AS)?'; 2) 'During the past

7 days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?'; and 3) 'During the past 7 days, how many hours did you actually work?'. The WPAI also included a presenteeism question related to unpaid work: 'during the past 7 days, how much did you rheumatoid arthritis (or Ankylosing Spondylitis) affects your ability to do your regular daily activities, other than work at a job?' The unpaid presenteeism question was measured using the same Likert scale presented in Figure 6.1.

Two final questions were included in the survey. The first asked whether the respondent had disclosed their condition to their employer, the responses included: 1) yes, and he/she is very supportive; 2) yes, and he/she is not supportive; 3) no; 4) I do not have an employer; and 5) prefer not to answer. Disclosure and the degree of support provided by the employer may affect how well an employee with RA or AS is able to work. The second question asked respondents to state whether they considered themselves to be work disabled. The survey included the definition of disability as stated by the Equality Act 2010 (Equality Act 2010). The responses to the question about disability included: 1) yes, because of RA or AS; 2) yes, because of another condition they might have which they were also asked to specify; 3) yes, because of RA or AS and another condition they may have, also asked to specify; or 4) no, they do not consider themselves to be work disabled. The question was included because it may have implications for the behaviour of the employer where they are bound, by law (Equality Act 2010), to provide support for their employee which may impact on levels of presenteeism. The final section of the survey asked respondents to provide their feedback about the survey.

6.3.3.2 Pilot testing

A two-phase pilot study was used to assess the phrasing of questions, logical flow and technical ability to accurately record data using the on-line survey. The first phase involved asking two working individuals with RA, identified from the Rheumatoid User Group (RUG) based at The University of Manchester (RUG 2017) to read through a paper copy of the questionnaire and provide feedback on the following;

1. Do the questions make sense to you (the respondent)? If not, please explain why.
2. Do the questions flow in a logical manner?
3. Any further comments you may have about the questionnaire.

The two individuals reported problems with the phrasing and possible ways of responding to standardised and validated questionnaires. For example, one individual suggested people were likely to differ in their assessment of the questions included in the SF6D. The second individual wanted to see Likert scales from 0 to 100 and not 0 to 10. These comments were discarded because the questions and Likert scales they referred to are standardised questionnaires that have been validated and cannot be changed. Both individuals thought the question regarding income was too intrusive. In economics, inclusion of income in a dataset is standard practice; however, in order to prevent participants from becoming upset with this question the option 'I prefer not to say' was added. The two individuals added that they felt positive about the questionnaire and its potential use in presenteeism related to RA.

The second phase involved uploading and running the on-line survey as intended in the main study, using Sawtooth software to produce the on-line version (Sawtooth Software 2017), to check for potential technical faults that may occur. A sample size of 30 working individuals with RA was recruited via ResearchNow to complete the pilot questionnaire. The online survey software's ability to skip questions where appropriate and its ability to record and export the data to an excel file was also assessed. Finally, the amount of time it took for respondents to complete the survey was also assessed; the survey aimed to take a maximum of 15 minutes to complete. The results of the pilot study are reported in Appendix 6.4.

6.3.4 Analysis: Mapping Methods

The purpose of this study was to develop a prediction model between measures of HRQoL and presenteeism. Krol et al. (2014) developed a prediction model for presenteeism using a negative binomial distribution with a log link. This model was reported to be the best fit model based on Quasi Likelihood under Independence Model Criterion and Corrected Quasi Likelihood under Independence Model Criterion. Given information reported in the study, it was difficult to understand whether a negative binomial model with log link was the best model to use. The model selected potentially over-fitted the data collected by Krol et al. causing problems for prediction when used out of sample. Also, because Krol et al. used hypothetical data from which to select their model, a different approach to model selection was thought necessary. Limited research has been done on the development of prediction models for presenteeism using self-reported measures of HRQoL. This study was also explorative in nature and

therefore adopts a different approach to the development of a prediction model for presenteeism; namely mapping methods.

In health economics, mapping methods have been frequently used to develop prediction algorithms that map responses from disease-specific measures of health to generic measures of health, such as the EQ5D (Brazier et al. 2009). Currently, no existing research has been conducted where mapping methods are applied to predict a concept beyond health, for example presenteeism. To understand whether mapping methods could be used to predict a concept beyond health a conversation with a specialist in the application of mapping methods was asked to state whether it was possible and it was confirmed that it would be worth conducting the study to explore this further (Personal communication with Dr. Louise Longworth, health economist at EuroQol).

This study used mapping methods to develop a prediction model for presenteeism. The Decision Support Unit (DSU) published a document that details of the technical stages needed to develop a prediction model (Longworth and Rowen 2013). Stage one involved the assessment of the degree of conceptual overlap between the dependent and independent variable. The second stage involved developing multiple prediction models specifying: 1) the covariates included in the model; and 2) appropriate regression models based on the type of dependent variable being predicted (discrete, continuous, and ordinal). The third stage involved assessing all of the developed potential prediction models performance and selecting the 'best' performing model by measuring the magnitude of errors between the observed value and predicted value of presenteeism. The model with the smallest errors was selected as the best performing prediction model.

6.3.4.1 Analysis: Conceptual Overlap

Conceptual overlap describes the degree to which two measures (or descriptive systems) measure the same concept (Brazier et al. 2009). If measures exhibit a weak correlation it provides an indication that the prediction model will not be able to adequately capture a relationship between the two measures (Longworth and Rowen 2013) and further analysis to quantify the association between the measures would be inappropriate (Brazier et al. 2009). The degree of conceptual overlap was used to inform the design of the prediction model for presenteeism. The survey designed for this study included two measures of presenteeism: the WPAI and the QQ method. It was decided, *a priori*, that the presenteeism instrument, the WPAI or the QQ method, that exhibited the strongest

correlated relationships with HRQoL measures would be the one selected and used to develop the prediction model summarising the quantified relationship.

To assess the degree of conceptual overlap between HRQoL, as measured by the EQ5D, SF6D and ICECAP-A, and presenteeism, as measured by the WPAI and QQ method, descriptive analysis methods were applied. A two stage process was used. Stage one generated visual scatterplots of the measures of HRQoL and presenteeism to explore the strength of the relationship between the two concepts.

Stage two involved formally estimating Spearman's rank correlation coefficients (Zar 2014) to investigate the degree to which measures of HRQoL overlapped with the measures of presenteeism. Spearman's rank correlation is a non-parametric technique; it was selected to assess the correlation between HRQoL and presenteeism because measures of health status and presenteeism typically do not follow a normal distribution (Leggett et al. 2017). The strength of the correlation was described by categories defined *a priori*: very weak (0 to 0.19); weak (0.2 to 0.39); moderate (0.40 to 0.59); strong (0.6 to 0.79); and very strong (0.8 to 1) (Swinscow 1997).

6.3.5 Analysis: Developing the prediction model

A total of 18 potential model specifications were defined *a priori*. The selection of regression models was informed by: 1) the type of dependent variable the model was used to predict (ordinal, discrete, continuous); 2) the distribution of the dependent variable; and 3) models used in previous literature. The dependent variable for this analysis was defined as the level of presenteeism and the independent variables included a measure of HRQoL (EQ5D-5L, SF6D or ICECAP-A). To promote the model's applicability for use across a wide range of datasets, which is important when applying the prediction model to potentially already existing datasets that have been collected by employers, only two variables, age and gender, alongside one measure of HRQoL, were included in each model (Longworth and Rowen 2013). Age and gender were added as co-variates to explore whether the combination of two variables increase the model's ability to predict presenteeism. Table 6.3 defines the specifications used for each model with presenteeism as the target outcome.

Table 6.3 Possible specifications for prediction model of presenteeism

Model	Health Status and Capability Measures	Covariates
1	EQ5D Index	-
2	EQ5DIndex	Age, Gender
3	EQ5D Index	Age*Gender
4	EQ5D Domains	-
5	EQ5D Domains	Age, Gender
6	EQ5D Domains	Age*Gender
7	SF6DIndex	-
8	SF6DIndex	Age, Gender
9	SF6DIndex	Age*Gender
10	SF6D Domains	-
11	SF6D Domains	Age, Gender
12	SF6D Domains	Age*Gender
13	ICECAP-A Index	-
14	ICECAP-A Index	Age, Gender
15	ICECAP-A Index	Age*Gender
16	ICECAP-A Domains	-
17	ICECAP-A Domain	Age, Gender
18	ICECAP-A Domains	Age*Gender

EQ5D: EuroQol Five Dimensions; SF6D: Short-Form Six Dimensions: ICECAP-A; ICEpop Capability measure for Adults

6.3.6 Description of the specific Regression Models

The following three sections describe the selected regression models that were tested in the analysis to identify the degree of association between measures of HRQoL and presenteeism.

6.3.6.1 Ordinary Least Squares

The simplest econometric model is the ordinary least squares (OLS) (Wooldridge 2016). It has been advocated as good practice to model a basic linear relationship between two variables first and then consider other options if deviations away from linearity are found (Pasta 2009). The dependent variable (measuring presenteeism) in this study is ordinal, ranging from zero to ten and increasing by increments of one. OLS is designed to treat the distance between the scores as equivalent making it a particularly good model to predict presenteeism where is it understood that the gaps between zero and one are the same as nine and ten; however, where the dependent variable is ordinal issue of bias may occur. Mapping studies, within health economics, are most commonly used to predict utility based HRQoL, for example the EQ5D-3L, using non-preference based measures, such as the Bath Ankylosing Spondylitis Disease Activity Index, to allow for the calculation of quality adjusted life-years (QALYs) (Braizer et al. 2009).

Equation one presents a simple linear regression relating *y* to *x*, where *y* is the dependent variable, *x* is the independent variable, β_0 is the constant, β_1 is the effect *x* has on *y*, *i* is the individual observation and *u* is the error term. OLS predicts or estimates the value of the dependent variable y_i using independent variables x_i and estimated coefficients β_0 and β_1.

$$y_i = \beta_0 + \beta_1 x_i + u_i \qquad \text{(Equation 1)}$$

The OLS model finds $\hat{\beta}_0$ and $\hat{\beta}_1$ and minimises the residual, $\hat{u}_I$, where $\hat{y}_i$ is the predicted observation, see equation two;

$$\hat{u}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \qquad \text{(Equation 2)}$$

Taking the sum of squared residuals minimises the residuals, see equation three;

$$\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_1)^2 \quad \text{(Equation 3)}$$

OLS is an unbounded regression model meaning it can produce estimates that are below zero. Therefore, in studies where the dependent variable is bounded between two values OLS may produce implausible estimates (Longworth and Rowen 2013). The β parameter estimated by an OLS where the dependent variable is censored (left or right) may produce inconsistent estimators (Cameron and Trivedi 2010). To overcome the theoretical limitations of OLS, the Tobit model was used because of its ability to handle censored data.

6.3.6.2 Tobit

Tobit models are used to describe the relationship between a dependent variable (y_i) that is ≥ 0 and an independent variable (x_i) (Cameron and Trivedi 2010), see equation four. The model assumes there is a latent variable, y^*_i, that linearly depends on the independent variable, x_i. Beta (β) determines the linear relationship between the independent variable and the latent dependent variable. The error, u_i, captures random unobservables that affect the relationship between y^*_i and x_i.

$$y_i^* = \beta_0 + x_i \beta_i + u_i \qquad \text{(Equation 4)}$$

Tobit models are useful when there is left or right censoring of the dependent variable equal to and greater than or equal to and less than some threshold, for example $0 \leq y \geq$

10. (Leggett et al. 2017) Tobit models are highly sensitive to heteroscedasticity which can lead to inconsistent estimates and affecting the standard errors (Long and Freese 2006). Therefore, a CLAD model was used because it is less sensitive to skewed data and is robust to heteroscedasticity (Maddala 1999).

6.3.6.3 Censored Least Absolute Deviation (CLAD)

Least absolute deviation (LAD) is a linear model that estimates the effects of the independent variables on the conditional median and not the mean (Powell 1984). LAD uses the conditional median to reduce the impact extreme observations have on the final estimate and is therefore less sensitive to outliers in the data. The LAD estimator $\breve{\beta}$, is defined by equation five:

$$\breve{\beta} = \min_{\beta} \sum_{i=1}^{n} |y_i - x'_i \tilde{\beta}| \qquad \text{(Equation 5)}$$

$\tilde{\beta}$ denotes any estimate that is an arbitrarily selected vector of the same dimensionality as β, $x_i'\hat{\beta}$ is the matrix transpose so that $x_i'\hat{\beta}$ is the scalar resulting from the product of vectors x_i and $\tilde{\beta}$. The CLAD estimator is a generalisation of the LAD estimator (Powell 1984). Like the Tobit model, it also allows for censoring at the lower value of 0, defined by equations six and seven where y_i^* is the observed level of presenteeism and y_i is the bounded measure of the presenteeism score, where $x_i'\breve{\beta}$ is the transpose matrix of the LAD estimator;

$$y_i^* = x_i'\breve{\beta} + \varepsilon_i \qquad \text{(Equation 6)}$$

$$y_i = f(x) = \begin{cases} y_i^*, if\ y_i^* \geq 0 \\ 0, if\ y_i^* \leq 0 \end{cases} \qquad \text{(Equation 7)}$$

Because the CLAD models median values it therefore minimises the sum of absolute deviations rather than the sum of squares (Pullenayegum et al. 2010).

6.3.7 Analysis: Testing Model Assumptions

Accuracy of predictions made by a multivariable regression model may be subject to bias if there are too few events per variable (EPV) (number of observations per variable) (Concato et al. 1995). Peduzzi et al. (1996) advocated a sample size per observation of ten, sample sizes smaller than ten were found to produce bias estimates in both positive and negative directions (Peduzzi et al. 1996). In a more recent study, Austin and

Steyerberg (2014) (Austin and Steyerberg 2014) conducted a study that investigated the number of EPV for estimating out-of-sample validity of logistic regression models and the minimum EPV was reported to be 20. Therefore, to minimise bias in the prediction model, the EPV for each of the levels of the domains included in the HRQoL measures will be checked. If levels of the domains of HRQoL measures are less than 20, then models that use domain level variables will be discarded.

Health status and capability measures are made up of different domains, with differing levels of severity (section 6.3.1). The domains and levels of severity have specific preference weights attached to describe the degree to which one health state is better or worse relative to another. The preference weights can be added across a health state to provide a total index score that describes the overall health status or capability of an individual. Prediction models developed using health status or capability data may use the total index score or individual domains to predict the outcome of interest.

Heteroscedasticity describes the presence of unequal variance of the residuals which causes regression models, such as the OLS, to produce biased standard errors (Wooldridge 2016). Heteroscedasticity may cause systematic bias of the predictions, and whilst this is not a catastrophe when developing prediction models, it is necessary to consider how biased predictors may affect results when included in a cost-effectiveness analysis (Longworth and Rowen 2013). Therefore, it was important to test for and understand whether heteroscedastic errors were present in the data.

Heteroscedasticity can be test in two ways: 1) using graphical plots of the residuals; and 2) conducting a formal test such as Breusch-Pagan test (Cameron and Trivedi 2010). The presence of heteroscedasticity, using a graphical plot, is identified when the residuals are not randomly distributed. Any type of pattern in the residuals potentially indicates the presence of heteroscedastic errors. The Breusch-Pagan test tests whether the variance of the errors from a regression are dependents on the values of the independent variable (Cameron and Trivedi 2010). The Breusch-Pagan test uses a chi-square test statistic. If the chi-square test statistic has a significance value (p-value) below 5% it can be concluded that heteroscedasticity is present. Heteroscedasticity of the residuals was tested using both methods.

6.3.8 Assessing Model Performance

The performance and selection of the 'best' model was examined using three goodness of fit statistics: 1) the absolute error (AE); 2) the mean absolute error (MAE); and 3) the root mean square error (RMSE). Formal guidelines that recommend using AS, MAE or RSME exist; however, the MAE and RMSE are commonly used statistics to judge the performance of one prediction model against another (Longworth and Rowen 2013). Measures of explanatory power are not useful for assessing the accuracy of prediction models because they focus on how well the model explains the dataset (Brazier et al., 2009). The purpose of developing an algorithm is to predict the dependent variable, in this case presenteeism, out of sample.

The AE measures the absolute difference between the mean observed and mean predicted values and is defined by the following (equation nine);

$$AE = \left|\left(\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i\right) - \left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)\right| \qquad (9)$$

Where $\hat{y}_i$ is the predicted level of presenteeism, y_i is the observed level of presenteeism, and n is the sample size.

The MAE measures the average error between the predicted and observed values but unlike the AE the MAE takes into account dispersion of the data. The MAE is given by equation ten;

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \qquad (10)$$

The RMSE, similar to the MAE, also takes into account dispersion of the data. However, because the RMSE squares the errors and attaches greater weights to larger errors, effectively adding a penalty to larger errors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \qquad (11)$$

Ideally the AE, MAE and RMSE would be calculated by developing the model using one dataset and testing its ability to predict using a separate external dataset. However, another dataset containing measures of presenteeism, health status and capability was not known to exist. To overcome this limitation, methods were used to split the collected dataset into separate estimation and validation datasets. There are three

common methods that can be used to split dataset: the hold-out method; the K-Fold method; and bootstrapping (Lachenbruch and Mickey 1968).

The hold-out method has been used to validate regression models used in the context of mapping studies (Longworth and Rowen, 2011). The hold-out method involves randomly splitting the dataset into two samples; an estimation sample and a validation sample. Typically, the data is split into two-thirds for the estimation sample and one-third for the validation sample (Grootendorst et al. 2007; K. A. Khan et al. 2014). The estimation sample is used to generate the algorithm and the validation sample is used to test the performance of the algorithm. A disadvantage of using the hold-out method is that it may estimate misleading errors because only one estimation and test experiment can be conducted (Longworth and Rowen, 2011).

The K-Fold method splits the dataset into K sections. There is no 'gold standard' method for selecting the most appropriate number of folds, however ten is common practice (Anguita et al. 2012). The K-fold method has advantages over the hold-out method because it is less likely to produce misleading error rates through the allowance of more than one estimation and test experiment (Refaeilzadeh, Tang, and Liu 2009).

Bootstrapping is a statistical method that involves sampling from the dataset and forming a new sample (a bootstrap sample). The bootstrap sample is made up of observations using sampling with replacement. One observation is randomly taken from the dataset and the predicted dependent variable score and prediction error are recorded. The observation is then replaced back into the dataset. This process can be repeated numerous times (500 times in Grootendorst et al. 2007 and Xie et al. 2010). Bootstrapping methods require more computational effort taking more time for estimation and validation models to run with potentially little gain in terms of assessing the performance of each model. Based on the advantages and limitations of the three methods used to generate separate validation and estimation datasets the K-fold method was used selected for use in this study.

6.3.9 Selecting the best performing prediction model

For all three goodness of fit statistics lower numbers indicate better performing models. Therefore, the model specification with the lowest AE, MAE, and RMSE would be considered to be the best model. However, in the event that no single model derives the lowest AE, MAE and RMSE each goodness of fit statistic would be ranked where the

lowest errors score low and large errors score high; for example, the lowest AE would be given a score of one; the second lowest AE would score two, and so on. The ranked scores for each AE, MAE and RMSE would then be totalled to provide an overall ranking score. The model with the lowest ranking score would be considered to be the best performing model. If more than one model received the same lowest ranking score the model with the lowest RMSE would be selected as the best performing model because of the statistic's ability to penalise large errors. There exists no established method to select the best performing prediction model; however, the methods employed in this study are in line with a recently published mapping study by Davison et al (2017). It is recommended and considered good practice to report errors across a range of values of the dependent variable to show how well the model performs when predicting low, middle and high values (Longworth and Rowen 2013). The AE, MAE and RMSE of the best performing model overall were reported across quartiles of the presenteeism outcome variable. Reporting errors as a percentage over the range of the dependent variable (presenteeism) was recommended by Brazier et al., (2009).

6.3.10 Subgroup Analysis

A subgroup analysis was conducted to explore the extent to which average predictions made by the model using the available study sample could be applied to specific subgroups within alternative samples of people with RA or AS; for example, is it valid to apply average predictions produced by the model to individuals with RA or AS that are all on biologics. The analysis was performed using individuals with RA or AS except when RA or AS were separately specified as the subgroup of interest. The dataset collected contained many variables that could potentially have been selected to perform the subgroup analyses. The choice of subgroups was informed *a priori* by literature (de Croon et al. 2004; Verstappen et al. 2004; van Vilsteren et al. 2015) and common sense. Table 6.4 lists the variables highlighted by the literature as key characteristics that determine work disability in populations with RA.

Table 6.4 Subgroups for analysis

Variable	Defined in dataset	Reference
Level of Education	Degree or equivalent, A levels or equivalent, GCSEs or equivalent, level 1 and below	(de Croon et al. 2004; S. M. M. Verstappen et al. 2004))
Job type in terms of physical demands	Non-manual or manual	(de Croon et al. 2004; S. M. M. Verstappen et al. 2004; van Vilsteren et al. 2015))
Disease Severity	RAPID score: high severity, medium severity, low severity and remission	(Verstappen et al. 2004; van Vilsteren et al. 2015))
csDMARDs used	Currently taking csDMARDS	(Verstappen et al. 2004; van Vilsteren et al. 2015)
Biologics used	Currently taking Biologics	(Verstappen et al. 2004; van Vilsteren et al. 2015)
Health Status	EQ5D VAS	(Verstappen et al. 2004; van Vilsteren et al. 2015)

RAPID3: Routine Assessment of Patient Index Data Three; csDMARDS: Disease Modifying Anti-Rheumatic Drugs

Other variables included in the dataset that were potentially useful to explore included: individuals with RA or AS; number of hours actually worked over the past 7 days; whether individuals reported their medication to be effective (yes, fair, and no); and whether individuals, given they had disclosed their condition to their employer, had a supportive employer; see Table 6.5.

Table 6.5 Additional covariates available in dataset

Other Variables	Defined in dataset
Rheumatoid Arthritis (RA) or Ankylosing Spondylitis (AS)	RA = 1, AS = 0
Medication Effective	Yes, fair, no
Number of hours actually worked during the past 7 days	Number of hours
Supportive employer given disclosure of condition	Yes, no

6.4 Results

The final sample was consisted of 542 observations: 472 observations with RA and 70 with AS. A total of 600 individuals initially completed the online survey about working with RA (n=514) or AS (n=86). Figure 6.2 presents a flow diagram providing details of

the number of observations that were included and those dropped from the study and the reasons why. Respondents were excluded for four reasons:

(i) Individuals who answered "yes" to the following screening question "Do you work in a job that pays you?" (Question QB in questionnaire) and "no" to "are you currently working for pay? (Question WPAI1) were deleted from the sample because they provided inconsistent information about their current job status.

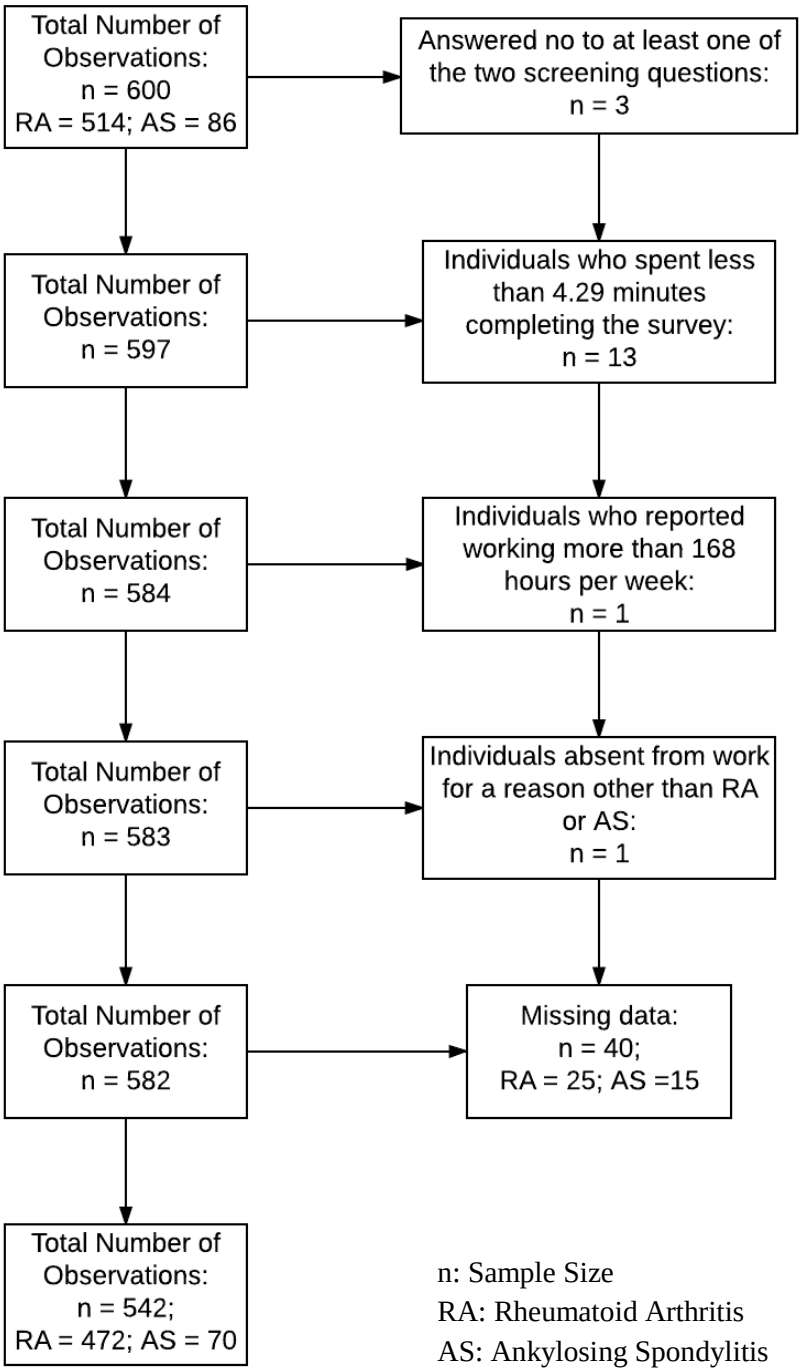
(ii) ResearchNow reject completed surveys if the individual takes less than 33% of the average time taken for the sample to complete the survey. The same rule was re-applied in the sample collected for this study in order to avoid collecting answers that were not adequately thought through. The survey in this study took an average of 13 minutes to complete; therefore, individuals who took less than 33% of the total time (4.29 minutes) to complete the survey were dropped from the final sample.

(iii) The number of hours in a seven day week is 168; therefore, individuals who report they worked over this number were excluded.

(iv) Individuals who reported being absent from work for any other reason unrelated to RA or AS were also excluded (in this case, it was one women on maternity leave).

All variables representing questions in the survey were manually checked for completeness. Only some responses for the WPAI presenteeism scale for paid work had missing data; this was because of a technical fault with the software used to record the data that was not highlighted as a problem in the pilot study. Twenty-five people with RA and 15 people with AS did not complete the WPAI and all 40 of these observations were dropped from the sample.

Figure 6.2 Observations dropped from sample



n: Sample Size
RA: Rheumatoid Arthritis
AS: Ankylosing Spondylitis

6.5.1 Descriptive Statistics

Table 6.6 describes the sociodemographic characteristics of the dataset (n=542 total; n=472 people with RA and n=70 people with AS). Females with RA made up 55% (n=297) of the RA sample which is in line with expectations from the population where women are more likely to develop RA at a ratio of 3:1 (female: male) (D. Symmons et al. 2002). AS affects more men and women at a ratio 2.5:1 to 5:1 ratio (Inman 2012), therefore the expectation would be see more males in the sample. However, males with AS were under-represented consisting of approximately 41% (n=29) of the AS sample (41%=29/70). Ages of the total sample were as expected with slightly more individuals reporting they were above the age of 45 years (Humphreys et al. 2013). White individuals were heavily over-represented in the total, RA and AS samples. Overall, the number of individuals with a University level qualification in the total and RA samples was approximately in line with the UK average at 38% (ONS 2013). The number of individuals with AS and educated to degree level were slightly over-represented (n=38) compared with the average population in the UK.

Table 6.6 Characteristics of survey respondents

Characteristics	**Total Sample (n=542)**		**Rheumatoid Arthritis (n=472)**		**Ankylosing Spondylitis (n=70)**	
	N	**%**	**N**	**%**	**N**	**%**
Total sample	542	100%	472	87.1%	70	12.9%
Gender, female	338	62.4%	297	54.8%	41	7.6%
Age bands, years						
18 – 34	64	11.8%	47	8.7%	17	3.1%
35 – 39	64	11.8%	50	9.2%	14	2.6%
40 – 44	60	11.1%	47	8.7%	13	2.4%
45 – 49	79	14.6%	68	12.5%	11	2.0%
50 – 54	96	17.8%	89	16.4%	7	1.3%
55 - 59	87	16.1%	81	14.9%	9	1.7%
60+	92	17.0%	90	16.6%	2	0.4%
Ethnicity						
White	485	89.5%	421	77.7%	64	11.8%
Other	57	10.5%	51	9.4%	6	1.1%
White Other	17	3.1%	14	2.6%	3	0.6%
Mixed: White/Black/Asian	10	1.9%	10	1.8%	-	-
Indian/Pakistani/Bangladeshi	17	3.1%	16	3.0%	1	0.2%
Chinese	2	0.4%	2	0.4%	-	-
Asian: Other	3	0.6%	3	0.6%	-	-
Black: African / Caribbean	5	0.9%	4	0.7%	1	0.2%
Any other ethnic group	2	0.4%	1	0.2%	1	0.2%
Prefer not to say	1	0.2%	1	0.2%	-	-
Education						
University Education and or Professional/Vocational equivalents	215	39.7	177	21.6%	38	7.0%
A levels/ Vocational equivalents	139	25.7	124	22.9%	15	2.8%
GCSEs/ Vocational equivalents	148	21.3	136	25.1%	12	2.2%
Level 1 and below and no qualifications	38	7.0	34	6.3%	4	0.7%
Prefer not to say	2	0.4	1	0.2%	1	0.2%

Table 6.7 summarises the stated job characteristics of the sample. Over 69% (n=372) of individuals reported that they were employed on a full-time basis working on average 38 hours per week. Full-time employees who worked in non-manual jobs made up 54% (n=290) of the total sample. Part-time employees made up 27% (n=146) of the sample, working on average 21 hours per week. Part-time workers in non-manual jobs made up 20% (n=109) of the total sample. Self-employed individuals made up 4.4% (n=24) of the sample working an average of 33 hours per week. Income levels for AS individuals were reportedly higher than those for RA.

Table 6.7 Job characteristics of survey respondents

Job Characteristics	Total Sample (n=542)		Rheumatoid Arthritis (n=497)		Ankylosing Spondylitis (n=70)	
	N	%	N	%	N	%
Full-time employee	372	68.6%	325	60.0%	47	8.7%
Hours per week	38 hours	-	38 hours	-	39 hours	-
Non-manual	290	53.5%	255	47.0%	35	6.5%
Manual	82	15.1%	70	12.9%	12	2.2%
Part-time employee	146	26.9%	132	24.4%	14	2.6%
Hours per week	21 hours	-	20 hours	-	30 hours	-
Non-Manual	109	20.1%	97	17.9%	12	2.2%
Manual	37	6.8%	35	6.5%	2	0.2%
Self-employed	24	4.4%	15	27.7%	9	1.7%
Hours per week	33 hours	-	31 hours	-	36 hours	-
Non-manual	20	3.7%	14	25.8	6	1.1%
Manual	4	0.7%	1	0.2%	3	0.6%
Income						
Less than £10,000	79	14.6%	72	13.3%	7	1.3%
£10,000 - £19,999	112	20.7%	95	17.5%	17	3.1%
£20,000 - £29,999	119	22.0%	105	19.4%	14	2.6%
£30,000 - £49,000	127	23.4%	100	18.5%	27	5.0%
£50,000 - £69,999	45	8.3%	40	7.4%	5	9.2%
£70,000 to £79,999	41	7.6%	41	7.6%	-	-
Prefer not to say	19	3.5%	19	3.5%	-	-

Table 6.8 describes the sample in terms of their self-reported severity of their primary condition (RA or AS). Almost half (n=269; 50%) of the sample was classified as having high disease severity. All individuals in the sample were asked to confirm whether they had received a diagnosis for RA or AS from a rheumatologist (doctor), and given that NICE guidelines recommend early aggressive treatment, the number of RA patients on biologics, csDMARDs, and combination therapy (biologics and csDMARDs) was far lower than would be expected (n =183; 39%). Sixty-four percent (n=45) of the AS population reported taking biologics, csDMARDs, or a combination of the two; this is lower than would have been expected.

Table 6.8 Self-reported disease severity and use of medicines

Health	**Total Sample (n=542)**		**Rheumatoid Arthritis (n=497)**		**Ankylosing Spondylitis (n=70)**	
	N	**%**	**N**	**%**	**N**	**%**
Disease Severity (RAPID)						
High	269	49.6%	236	43.5%	33	6.1%
Medium	169	31.2%	146	26.9%	23	4.2%
Low/Remission	104	19.2%	90	16.6%	14	2.6%
Years with condition						
> 5	213	39.3%	177	32.7%	36	6.6%
5 – 9	103	19.0%	93	17.2%	10	1.8%
10 – 14	70	12.9%	62	11.4%	8	1.5%
15 – 19	69	12.7%	65	12.0%	4	0.7%
20+	87	16.1%	75	13.8%	12	2.2%
Medication						
Biologics	111	20.5%	69	12.7%	42	7.7%
Biologics only	66	12.2%	32	5.9%	34	6.3%
csDMARDs only	117	21.6%	114	21.0%	3	5.5%
Biologics and csDMARDS	45	8.3%	37	6.8%	8	0.7%
csDMARDs	162	29.9%	151	27.9%	11	2.0%
Oral Corticosteroids	58	10.7%	57	10.5%	1	0.2%
NSAIDs	239	44.1%	207	38.2%	32	5.9%
Other Analgesics	106	19.6%	92	17.0%	14	2.6%
None	84	15.5%	77	14.2%	7	1.3%
Do not know	7	1.3%	6	1.1%	1	0.2%
No. of Medications						
0	84	15.5%	77	14.2%	7	1.3%
1	287	53.0%	250	46.1%	37	6.8%
2	118	21.8%	103	19.0%	15	2.8%
3	33	6.1%	25	4.6%	8	1.5%
4	11	2.0%	10	1.8%	1	0.2
5	2	0.4%	1	0.2%	1	0.2
Don’t know	7	1.3%	6	0.1%	1	0.2

RAPID3: Routine Assessment of Patient Index Data Three; csDMARDS: Disease Modifying Anti-Rheumatic Drugs

Health Related Quality of Life

Table 6.9 lists the scores from both measures of health status and capability generated by the total sample, and separately, for individuals with RA or AS. Figure 6.3, 6.4 and 6.5 present the distributions of HRQoL using the three measures (EQ5D, SF6D, ICECAP-A, respectively) for the total sample, and for the RA sample and AS sample, respectively. The distribution for the EQ5D for the whole sample and two respective sub-samples is highly skewed to the right. The distribution of the SF6D, again for all three samples, shows a slightly skewed distribution to the right but is close to a normal

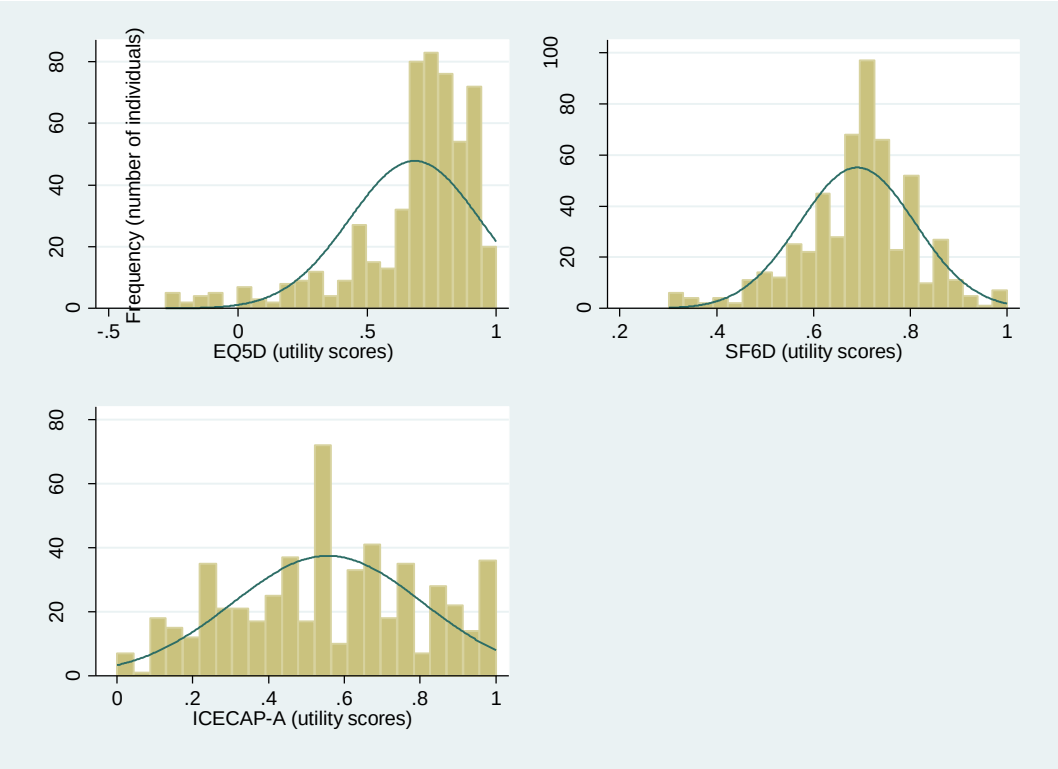
distribution. The ICECAP-A shows, for all three samples, a normal distribution however the responses are far more spread out creating a shallower 'bell' shaped curve.

Table 6.9 Health Status and capability scores of survey respondents

Health Status	Mean	Minimum	Maximum
EQ5D-5L			
Total Sample	0.6842	-0.2810	1
Rheumatoid Arthritis	0.6829	-0.2898	1
Ankylosing Spondylitis	0.6936	-0.1068	1
SF6D			
Total Sample	0.6905	0.301	1
Rheumatoid Arthritis	0.6935	0.301	1
Ankylosing Spondylitis	0.6705	0.319	1
ICECAP-A			
Total Sample	0.5569	0	1
Rheumatoid Arthritis	0.5544	0	1
Ankylosing Spondylitis	0.5735	0	1

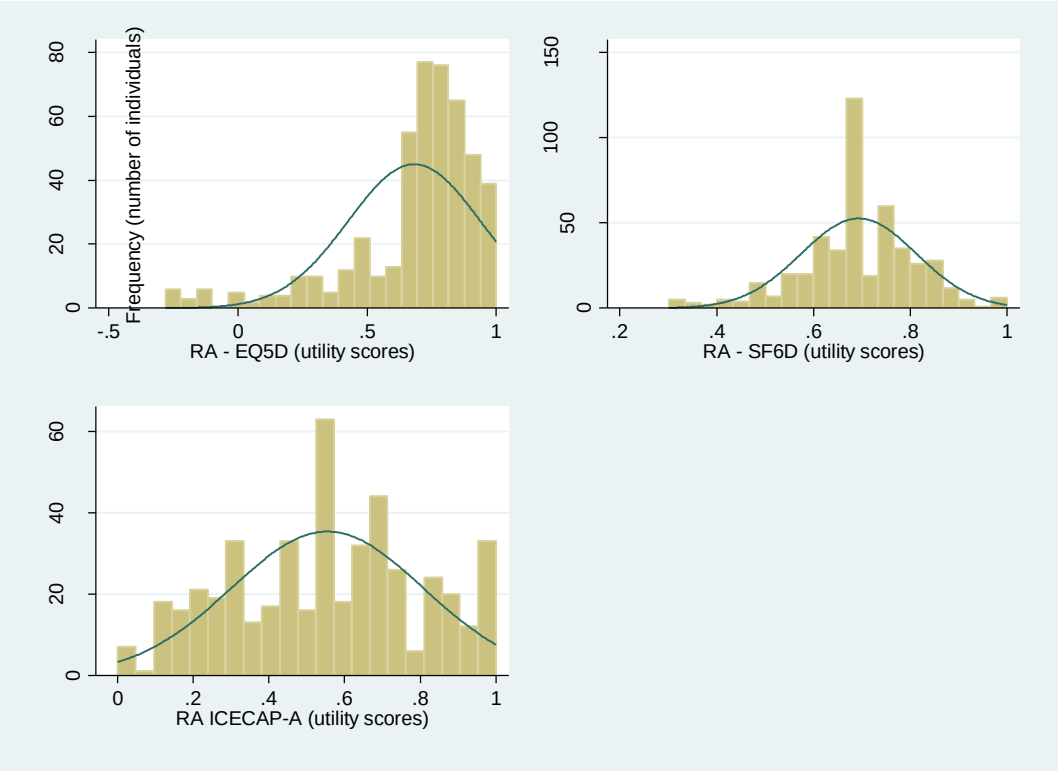
EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults

Figure 6.3 Distribution of HRQoL scores for the total sample



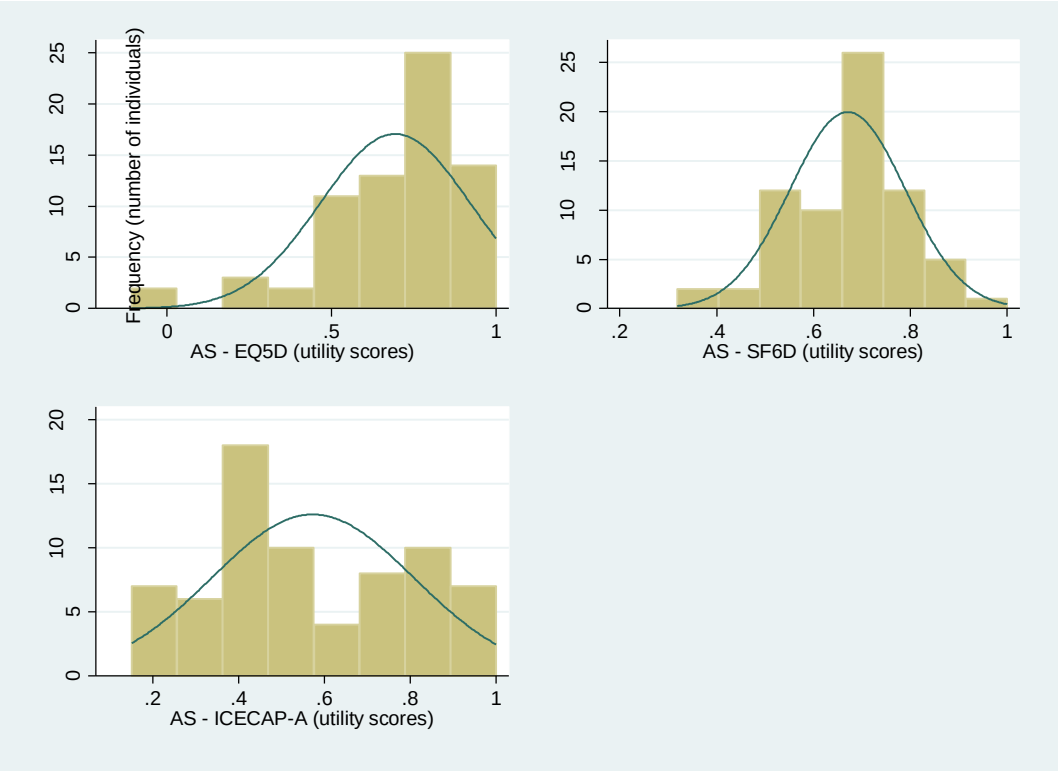
EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults

Figure 6.4 Distribution of HRQoL scores for individuals with rheumatoid arthritis



EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults; RA: Rheumatoid Arthritis

Figure 6.5 Distribution of HRQoL scores for individuals with ankylosing spondylitis



EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults; RA: Rheumatoid Arthritis; AS: Ankylosing Spondylitis

Presenteeism

For statistical purposes, it was important to ensure adequate sample sizes for each level of presenteeism were recorded. Table 6.10 presents the number of respondents per category of presenteeism within each measure (WPAI or QQ) from zero to ten. High levels of presenteeism for the WPAI (9 – 10) and QQ method (0 – 2) did not have sample sizes that were large enough to enable the development of an accurate prediction at these levels. To remedy the problem, presenteeism levels for the WPAI, values nine and ten, were recoded as eight, and the QQ method, values zero, one, and two, were re-coded as three.

Table 6.10 Response sample sizes for presenteeism categories

Level of presenteeism	WPAI Sample (n=542)	WPAI after re-coding (n=542)	QQ Method (total efficiency)[4] (n=542)	QQ Method (total efficiency) after re-coding[5] (n=542)
0	83	83	10	-
1	65	65	1	-
2	68	68	4	-
3	67	67	17	32
4	50	50	26	26
5	93	93	71	71
6	61	61	59	59
7	28	28	82	82
8	19	29	116	116
9	6	-	84	84
10	2	-	113	113

WPAI: Work Productivity Activity Impairment questionnaire; QQ Method: Quantity and Quality Method

Table 6.11 reports presenteeism scores from the WPAI and QQ method. Values from the WPAI and QQ method have opposite interpretations; zero from the WPAI means complete ability to work (100% productivity) and zero from the QQ method means completely unable to work (0% productivity), both scales report mean presenteeism levels which have a similar interpretation where 3.4 and 7.4 are approximately equivalent.

[4] The QQ method includes three questions regarding presenteeism: 1) quantity of work completed; 2) quality of work completed; and 3) the degree of efficiency the individuals worked at. Total efficiency refers to the third question asked by the QQ method.

[5] Re-coding refers to the responses collapsed to create sample sizes large enough for analysis

Table 6.11 reported levels of presenteeism

Presenteeism	Mean	Minimum	Maximum
WPAI			
Total Sample	3.404	0	8
Rheumatoid Arthritis	3.328	0	8
Ankylosing Spondylitis	3.914	0	8
QQ Method			
Degree of efficiency	7.448	3	10
Total Sample	7.538	3	10
Rheumatoid Arthritis	6.843	3	10
Ankylosing Spondylitis			
Quantity of Work			
Total Sample	7.646	3	10
Rheumatoid Arthritis	7.650	3	10
Ankylosing Spondylitis	7.614	3	10
Quality of Work			
Total Sample	7.886	3	10
Rheumatoid Arthritis	7.860	3	10
Ankylosing Spondylitis	8.057	3	10

WPAI: Work Productivity Activity Impairment questionnaire[6]; QQ Method: Quantity and Quality method[6]

Figure 6.6 shows the distribution of the levels of presenteeism recorded by the WPAI. The distribution appears to be negative with excessive numbers of respondents recording the value of five. The high number respondents reporting a value of five may be a result of the way in which individuals complete Likert scales. Dillman (2006) explains how external factors, such as culture, can influence individual's responses to scales. Scales with categories that exhibit a mid-point are potentially susceptible to increased response rates at the mid-point (Dillman 2006). In this case, the WPAI asked the individual to state their working performance using a scale between zero and 10; five is half way point.

[6] Scores range from 0 to 10

Figure 6.6 Distribution of the Work Productivity Activity Impairment Questionnaire (WPAI)

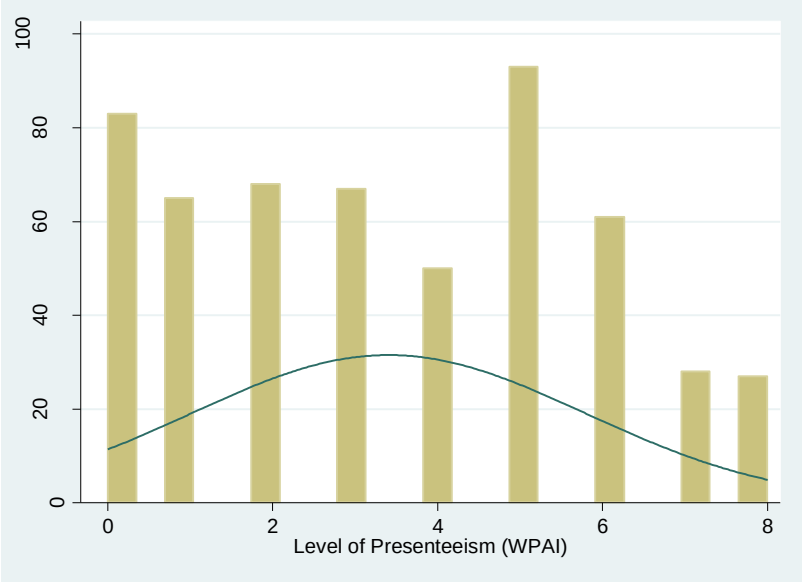
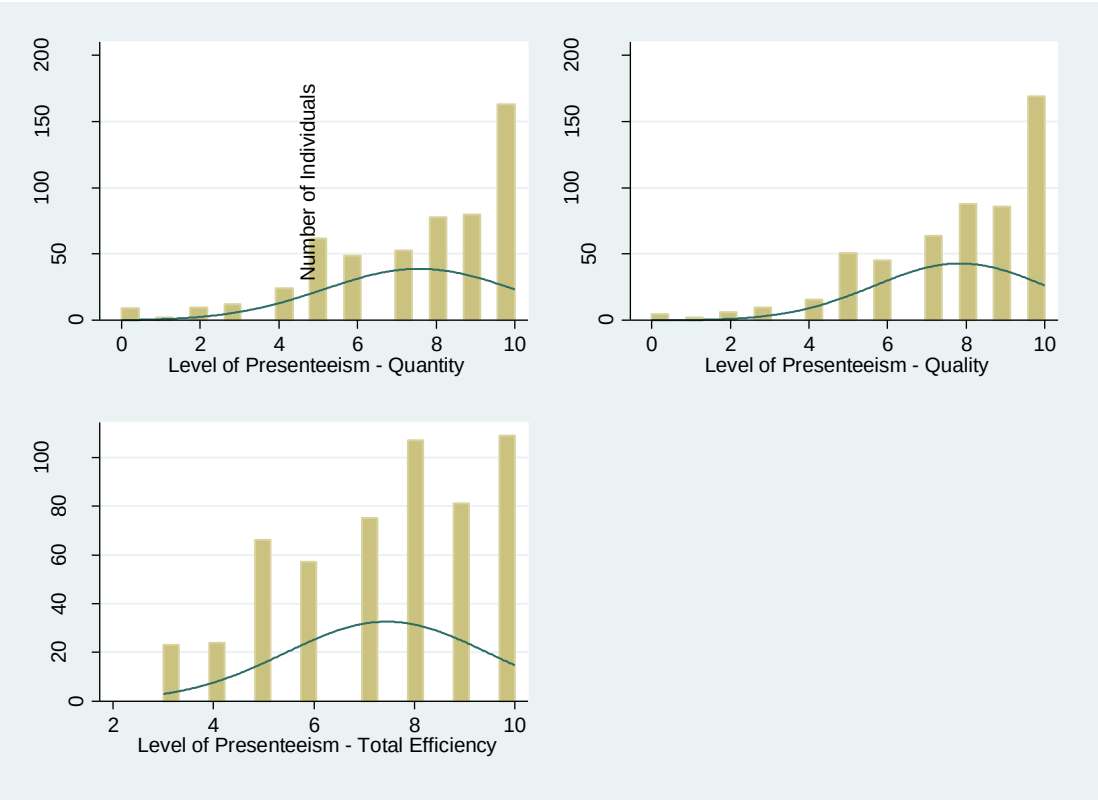


Figure 6.7 shows the three distributions of the QQ method; quantity of work completed, quality of work completed, and total efficiency.

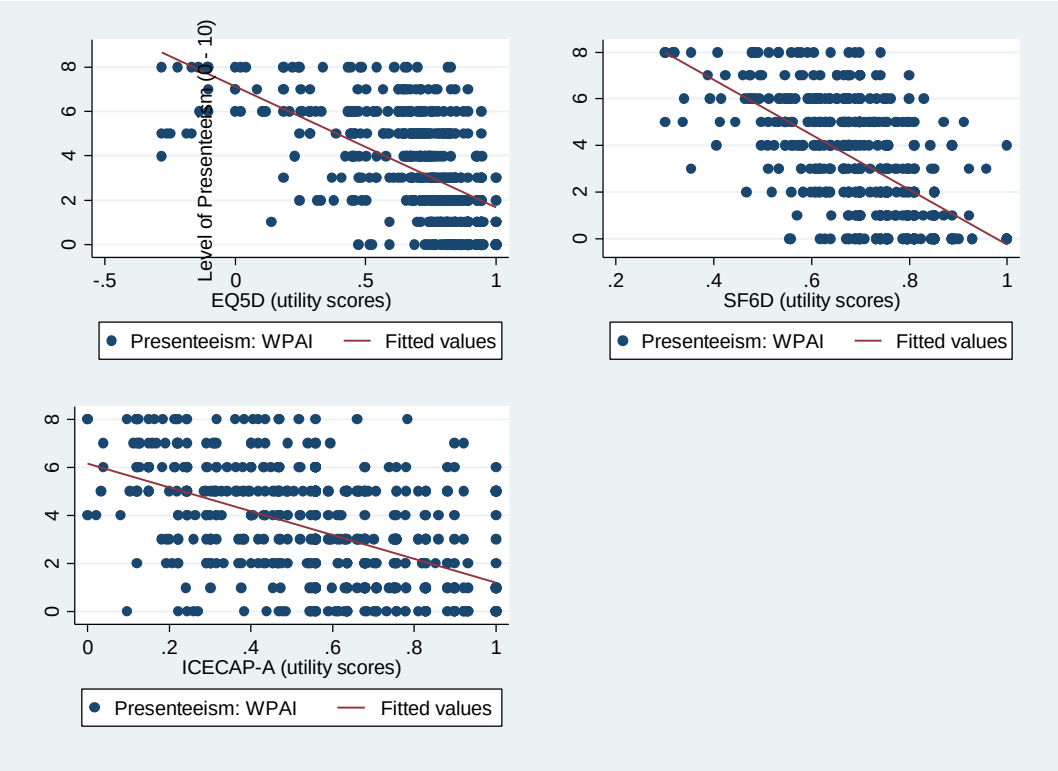
Figure 6.7 Distribution of the Quantity and Quality Method (QQ Method)



6.5.1.2 Conceptual Overlap

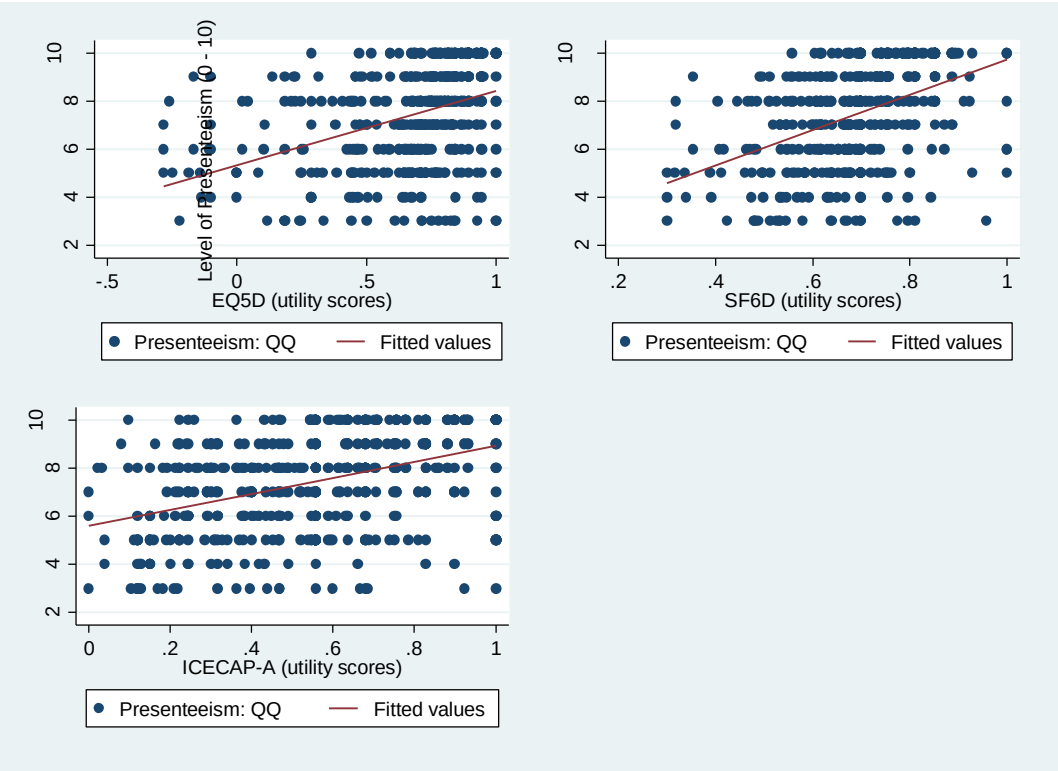
Figures 6.8 and 6.9 present visual plots of observed levels of presenteeism and HRQoL measured by the EQ5D, SF6D and ICECAP-A. Figure 6.8 shows a negative relationship between HRQoL and presenteeism measured by the WPAI and Figure 6.9 shows a positive relationship between HRQoL and presenteeism measured by the QQ method (measured using total efficiency, see figure 6.1); both scatterplots exhibited relationships in the direction expected *a priori*. Visual scatterplots of the SF6D and the two measures of presenteeism displayed stronger linear relationships relative to the EQ5D and ICECAP-A. The EQ5D exhibits a skewed relationship with many participants reporting high levels of health status. The ICECAP-A shows almost no visual relationship with observed levels of presenteeism measured by the QQ method and WPAI.

Figure 6.8 Health-related quality of life and presenteeism measured using the Work Productivity Activity Impairment Questionnaire of the total sample



EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults

Figure 6.9 Health-related quality of life and presenteeism measured by the Quantity and Quality method



EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults; RA: Rheumatoid Arthritis; AS: Ankylosing Spondylitis; QQ: QQ method, specifically the third question regarding efficiency level of work completed

Table 6.12 shows the calculated Spearman's rank correlation coefficients between the total index scores of the EQ5D, SF6D and ICECAP-A with levels of presenteeism measured using the WPAI and the QQ method.

Table 6.12 Spearman's Rank Correlations[4] between HRQOL and Presenteeism

HRQoL	**Presenteeism**			
	WPAI	Quantity[1]	Quality[2]	Total Efficiency[3]
EQ5D-5L	-0.64	0.43	0.48	0.45
SF6D	-0.60	0.43	0.44	0.48
ICECAP-A	-0.52	0.37	0.41	0.43

EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults; [1]Quantity refers to the amount of work the individual was able to produce; [2]Quality refers to the quality of work the individual was able to produce; [3]Total efficiency refers to an individual's ability to work efficiently ; [4]categories of the strength of the correlations: very weak (0 to 0.19); weak (0.2 to 0.39); moderate (0.40 to 0.59); strong (0.6 to 0.79); and very strong (0.8 to 1)

Strong negative correlations were observed between the EQ5D and WPAI (r = -0.64) and SF6D and WPAI (r = -0.60) and a moderate negative correlations was observed for the ICECAP-A (r = -0.52). Moderate and positive correlations were identified for the HRQoL measures and the QQ method with the SF6D having the highest correlation (r =

0.475). The ICECAP-A was less correlated with presenteeism when compared to the EQ5D and SF6D.

From this point forward, the study uses the WPAI as the observed measure of presenteeism. The QQ method was not used to develop a prediction model based on the correlations reported in table 6.12. The degree of conceptual overlap of HRQoL measures with the WPAI was stronger than with the QQ method providing an initial indication that the WPAI would be the better measure for presenteeism compared to the QQ method.

6.5.2 Model specifications

Table 6.13 reports the number of EPV for the domains of each HRQoL measure that are below 20. The low number of EPV may cause biased prediction, therefore prediction models using domain level independent variables (see section 6.3.4.1) were omitted from the analysis.

Table 6.13 Events per variable, HRQoL domains

HRQOL	Level	Same Size
EQ5D Mobility	5	5
EQ5D Self Care	5	7
EQ5D Usual Activities	5	8
EQ5D Pain	5	13
EQ5D Anxiety and Depression	5	16
SF6D Social Functioning	5	14
SF6D Mental Health	5	14
ICECAP-A Love, Friendship and Support	4	10
ICECAP-A Independence	4	13
ICECAP-A Enjoyment and Pleasure	4	14

Heteroscedasticity for OLS models was assessed using the Breusch-Pagan test (Wooldridge 2016). The OLS models satisfied the assumption of homoscedasticity at the 5% significance level (see Table 6.14).

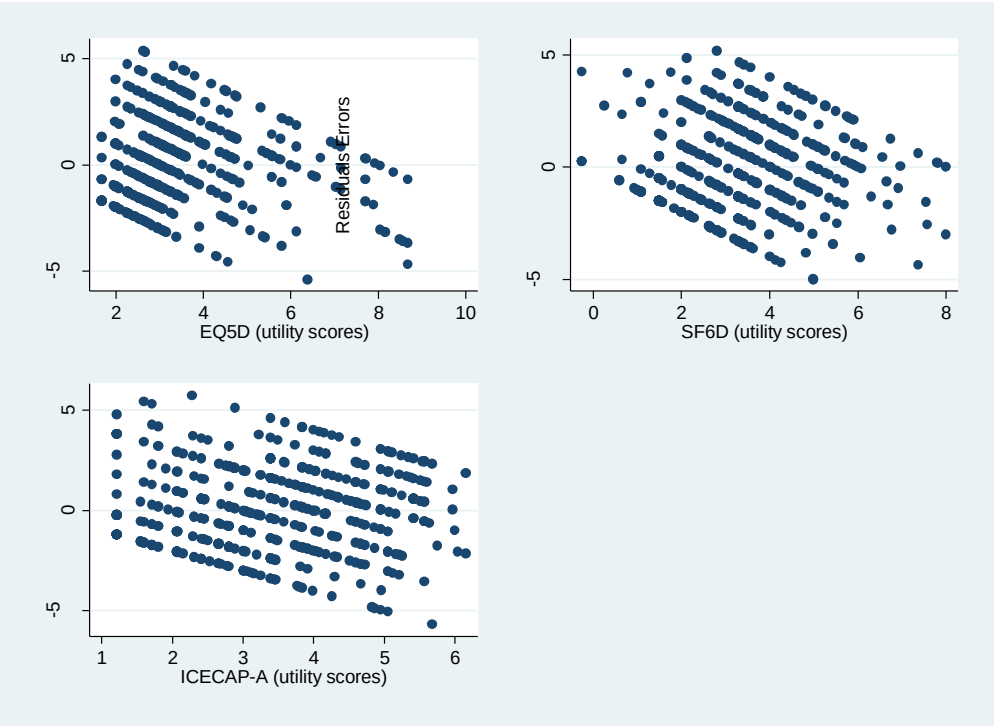
Table 6.14 Test for heteroscedasticity using the Breusch-Pagan test

Model Specification	**Breusch-Pagan test**	
	chi2	p > chi2
EQ5D	0.67	0.4130
EQ5D Age Gender	1.06	0.3030
EQ5D Age, Gender, Age*Gender	1.15	0.2829
SF6D	0.08	0.7724
SF6D Age Gender	0.48	0.4870
SF6D Age, Gender, Age*Gender	1.05	0.3064
ICECAP-A	0.00	0.9456
ICECAP-A Age Gender	0.10	0.7535
ICECAP-A Age, Gender, Age*Gender	0.04	0.8432

EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults

Whilst, all OLS models passed the formal heteroscedasticity test, it was possible that some systematic bias may have been present. Figure 6.10 show the potential presence of heteroscedasticity where plots of the residuals from the three OLS prediction models for the EQ5D, SF6D and ICECAP-A revealed a systematic pattern. Therefore, the OLS and tobit models were ran on the basis that the Breusch-pagan tests for heteroscedasticity were correct. However, because figure 6.10 graphically indicates the potential for heteroscedasticity the CLAD model was run because it is less sensitive to the presence of heteroscedasticity (see section 6.3.6.2).

Figure 6.10 Plot of Residuals: Heteroscedasticity



EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions: ICECAP-A: ICEpop CAPability measure for Adults

6.5.2.2 K-Fold Validation

The K-Fold validation method was used to split the sample and test the performance of each prediction model. The sample was split 10 times forming 10 development datasets and 10 validation datasets. The development datasets were used to develop the prediction model and the validation datasets were used to test the performance of each model. Table 6.15 lists the sample sizes for each development and validation dataset after the 10 K-fold split. The results of the 10 K-fold spilt make up the final results discussed in section 6.5.2.3.

Table 6.15 Sample sizes after 10-Fold

10 - Fold	Development (n=542)	Validation (n=542)
1	494	48
2	479	63
3	480	62
4	502	40
5	501	41
6	487	55
7	492	50
8	473	69
9	482	60
10	488	54

6.5.2.3 Model Performance Results

Tables 6.16, 6.17 and 6.18 present the performance results for each model specification using the OLS, Tobit, and CLAD, respectively. The error results were derived from the 10 K-fold splits performed for each model. The errors that are underlined highlight the smallest observed errors. The AEs ranged from 0.2144 to 0.4206, MAEs from 1.5468 to 1.7793, and RMSEs from 1.9028 to 2.1802. As a percentage, AEs showed a percentage error of up to 5.3%, MAEs up to 22.2%, and RMSE up to 27.3%.

The OLS model with the lowest AE was model seven (ICECAP-A only) and the OLS model with the lowest MAE and RMSE was model five (SF6D Age Gender). The Tobit model with the lowest AE was model 15 (ICECAP Age Gender Age*Gender) and the model with the lowest MAE and RMSE was model 11 (SF6D, Age, Gender). The CLAD model with the lowest AE was model 26 (ICECAP, Age, Gender); the model with the lowest MAE and RMSE was model 22 (SF6D only). No single model had the lowest AE, MAE, and RMSE scores; therefore, the models were ranked based on their AE, MAE and RMSE scores (see Table 6.19). Using this ranking process, the best performing model was model five which included the SF6D, age and gender.

Table 6.16 OLS model performance

		Observed Level of Presenteeism				Predicted Level of Presenteeism				Model Performance		
Model	**Model Specification**	*No. of Obs.	Mean	Min.	Max.	*No. of Obs.	Mean	Min.	Max.	**AE**	**MAE**	**RMSE**
1	EQ5D	488	3.4228	0.0000	7.8000	54	3.3920	1.7376	8.0677	0.3246	1.5886	1.9301
2	EQ5D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.3919	1.6111	8.1344	0.3279	1.5824	1.9314
3	EQ5D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.3957	1.3974	8.0440	0.3191	1.5843	1.9590
4	SF6D	488	3.4228	0.0000	7.8000	54	3.4031	0.4245	7.1975	0.2705	1.5659	1.9083
5	SF6D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.4028	0.3651	7.2710	0.2758	1.5468	1.9028
6	SF6D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.4079	0.3581	7.3570	0.2642	1.5678	1.9204
7	ICECAP	488	3.4228	0.0000	7.8000	54	3.4123	1.2104	5.8833	0.2209	1.6576	2.0226
8	ICECAP, Age, Gender	488	3.4228	0.0000	7.8000	54	3.4131	1.1393	6.0123	0.2232	1.6523	2.0206
9	ICECAP, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.4160	0.8608	6.0799	0.2280	1.6714	2.0416

*No. of Obs. = Number of Observations; EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults; AE: Absolute Error; MAE: Mean Absolute Error; RMSE: Root Mean Square Error; Min: Minimum; Max: Maximum

Table 6.17 Tobit model performance

		Observed				Predicted				Model Performance		
Model	**Model Specification**	*No. of Obs.	Mean	Min.	Max.	*No. of Obs.	Mean	Min.	Max.	**AE**	**MAE**	**RMSE**
10	EQ5D	488	3.4228	0.0000	7.8000	54	3.2629	1.2397	8.9850	0.3433	1.6292	1.9639
11	EQ5D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.2624	1.0754	9.0716	0.3481	1.6320	1.9680
12	EQ5D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.2663	0.8700	8.9545	0.3367	1.6427	1.9901
13	SF6D	488	3.4228	0.0000	7.8000	54	3.2530	-0.5086	8.0448	0.3051	1.6130	1.9491
14	SF6D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.2532	-0.5974	8.1476	0.3071	1.6030	1.9477
15	SF6D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.2594	-0.5485	8.2423	0.2970	1.6163	1.9624
16	ICECAP	488	3.4228	0.0000	7.8000	54	3.2567	0.5784	6.2631	0.2623	1.6517	2.0349
17	ICECAP, Age, Gender	488	3.4228	0.0000	7.8000	54	3.2576	0.4836	6.4302	0.2652	1.6481	2.0339
18	ICECAP, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.2621	0.2207	6.4584	0.2611	1.6847	2.0535

*No. of Obs. = Number of Observations; EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults; AE: Absolute Error; MAE: Mean Absolute Error; RMSE: Root Mean Square Error; Min: Minimum; Max: Maximum

Table 6.18 CLAD model performance

		Observed Level of Presenteeism				Predicted Level of Presenteeism				Model Performance		
Model	**Model Specification**	*No. of Obs.	Mean	Min.	Max.	*No. of Obs.	Mean	Min.	Max.	**AE**	**MAE**	**RMSE**
19	EQ5D	488	3.4228	0.0000	7.8000	54	3.3748	0.8319	10.5077	0.3887	1.7024	2.0179
20	EQ5D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.4782	0.5194	11.0621	0.4206	1.7289	2.1802
21	EQ5D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.4825	0.0548	9.5797	0.4024	1.7793	2.1567
22	SF6D	488	3.4228	0.0000	7.8000	54	3.2439	-0.2299	7.6405	0.3058	1.6103	1.9568
23	SF6D, Age, Gender	488	3.4228	0.0000	7.8000	54	3.2478	-0.3734	7.8453	0.2875	1.6149	2.0090
24	SF6D, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.4034	-0.5265	7.8171	0.2633	1.6858	2.0759
25	ICECAP	488	3.4228	0.0000	7.8000	54	3.2596	0.3649	6.5092	0.2738	1.6558	2.1294
26	ICECAP, Age, Gender	488	3.4228	0.0000	7.8000	54	3.3185	0.1451	6.8586	0.2144	1.6391	2.1130
27	ICECAP, Age, Gender, Age*Gender	488	3.4228	0.0000	7.8000	54	3.3450	-0.0951	7.2225	0.2418	1.7267	2.1765

*No. of Obs. = Number of Observations; EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults; AE: Absolute Error; MAE: Mean Absolute Error; RMSE: Root Mean Square Error; Min: Minimum; Max: Maximum

Table 6.19 Models ranked based on size of errors from smallest to largest

Model number	Model Spec	Model Type	AE	AE Rank	MAE	MAE Rank	RMSE	RMSE Rank	Total Rank
5	SF6D	OLS	0.2705	11	1.5659	2	1.9083	2	15
4	SF6D Age and Gender	OLS	0.2758	13	1.5468	1	1.9028	1	15
6	SF6D Age Gender Age*Gender	OLS	0.2642	9	1.5678	3	1.9204	3	15
2	EQ5D	OLS	0.3246	20	1.5886	6	1.9301	4	30
1	EQ5D Age Gender	OLS	0.3279	21	1.5824	4	1.9314	5	30
14	SF6D Age and Gender	Tobit	0.3071	18	1.6030	7	1.9477	6	31
3	SF6D	Tobit	0.3051	16	1.6130	9	1.9491	7	32
13	EQ5D Age Gender Age*Gender	OLS	0.3191	19	1.5843	5	1.9590	9	33
22	SF6D	CLAD	0.3058	17	1.6103	8	1.9568	8	33
15	SF6D Age Gender Age*Gender	Tobit	0.2970	15	1.6163	11	1.9624	10	36
10	ICECAP Age and Gender	OLS	0.2232	3	1.6523	18	2.0206	16	37
23	SF6D Age and Gender	CLAD	0.2875	14	1.6149	10	2.0090	14	38
11	ICECAP Age and Gender	CLAD	0.2144	1	1.6391	14	2.1130	23	38
12	ICECAP	OLS	0.2209	2	1.6576	20	2.0226	17	39
8	ICECAP	Tobit	0.2623	7	1.6517	17	2.0349	19	43
17	ICECAP Age and Gender	Tobit	0.2652	10	1.6481	16	2.0339	18	44
16	ICECAP Age Gender Age*Gender	OLS	0.2280	4	1.6714	21	2.0416	20	45
7	EQ5D	Tobit	0.3433	23	1.6292	12	1.9639	11	46
26	EQ5D Age Gender	Tobit	0.3481	24	1.6320	13	1.9680	12	49
19	ICECAP Age Gender Age*Gender	Tobit	0.2611	6	1.6847	22	2.0535	21	49
9	EQ5D Age Gender Age*Gender	Tobit	0.3367	22	1.6427	15	1.9901	13	50
18	SF6D Age Gender Age*Gender	CLAD	0.2633	8	1.6858	23	2.0759	22	53
25	ICECAP	CLAD	0.2738	12	1.6558	19	2.1294	24	55
24	ICECAP Age Gender Age*Gender	CLAD	0.2418	5	1.7267	25	2.1765	26	56
27	EQ5D	CLAD	0.3887	25	1.7024	24	2.0179	15	64
21	EQ5D Age Gender Age*Gender	CLAD	0.4024	26	1.7793	27	2.1567	25	78
20	EQ5D Age Gender	CLAD	0.4206	27	1.7289	26	2.1802	27	80

EQ5D-5L: EuroQol Five Dimensions Five Levels; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop Capability measure for Adults; AE: Absolute Error; MAE: Mean Absolute Error; RMSE: Root Mean Square Error; Min: Minimum; Max: Maximum

Three OLS models using the SF6D were identified with a total lowest ranking score of 15. Table 6.20 shows the top three performing models each with a ranking score of 15. Model five had the lowest RMSE score and was therefore selected as the best prediction model for presenteeism using SF6D data. The percentage error for model five's AE, MAE and RMSE were 3.4%, 19.3% and 23.7%, respectively.

Table 6.20 Model selection using the root mean square error (RMSE)

Model number	Model Specification	Model Type	RMSE
5	SF6D Age and Gender	OLS	1.9028
4	SF6D	OLS	1.9083
6	SF6D Age Gender Age*Gender	OLS	1.9204

SF6D: Short Form Six Dimensions: OLS; Ordinary Least Squares

Table 6.21 shows the AE, MAE and RMSE for model five over a range of quartile WPAI values. The model performs better when predicting WPAI values at scores of 0 to 1 and 4 to 5. The model's predictions become more dispersed when attempting to predict high WPAI values (6, 7, 8) which is unsurprising given that the sample sizes for a value of seven and eight were relatively small when compared with the number of individuals who reported values between zero and six.

Table 6.21 Errors across the quartile range of levels of presenteeism

Quartile	Level of presenteeism (WPAI)	Number	AE	MAE	% MAE	RMSE	% RMSE
1	0 or 1	148	0.000	0.4709	5.5	0.4886	6.1
2	2 or 3	135	0.000	0.4925	6.2	0.5000	6.3
3	4 or 5	143	0.000	0.4484	5.6	0.4769	6.0
4	6,7 or 8	116	0.000	0.6696	8.4	0.7667	10.0

AE: Absolute Error; MAE: Mean Absolute Error; RMSE: Root Mean Square Error; N: number of observations; WPAI: Work Productivity Activity Impairment questionnaire

6.5.3 Subgroup Analysis

Subgroups of the total sample, defined *a priori*, were examined to see if, and by how much, predictions changed for the 'best' model (model 5) identified in the main analysis of the total sample. The selected prediction model was run on specific subgroups of the sample collected. The purpose was to explore the extent to which the predictions made by model 5 were applicable for specific subgroups within the collected sample. Table 6.22 shows the ME, MAE and RSME for each model with an added variable compared with the errors of the baseline regression model (model 5). The percentage MAE and percentage RMSE represent the difference in errors compared with the baseline model (model 5).

Table 6.22 Subgroup analysis: results

Subgroups	Sample Size	MAE	% difference MAE	RMSE	% difference RMSE
Baseline*	542	1.575	0.00	1.916	0.00
Education:					
Alevel and Degree	354	1.592	1.1	1.921	0.3
GCSE and below	188	1.540	2.2	1.903	0.7
Job Type					
Non-manual	419	1.598	1.5	1.948	1.7
Manual	123	1.377	12.6	1.718	10.3
Hours worked					
Full-time	372	1.603	1.8	1.938	1.2
Part-time	170	1.490	5.4	1.863	2.8
Disease Severity (RAPID3)					
High	269	1.430	9.2	1.808	5.6
Medium	169	1.437	8.8	1.772	7.5
Low/Remission	104	0.878	44.3	1.179	38.5
Medication Effective					
Yes	132	1.562	0.8	1.958	2.2
Fair	245	1.648	4.6	1.990	3.9
No	64	1.311	16.8	1.640	14.4
Biologics					
Yes	111	1.731	9.9	2.065	7.8
No	431	1.530	2.9	1.868	2.5
csDMARDs					
Yes	151	1.604	1.8	1.949	1.7
No	321	1.544	2.0	1.884	1.7
Type of condition					
RA	472	1.576	0.1	1.911	0.3
AS	70	1.554	1.3	1.914	0.1
Supportive employer; if condition disclosed					
Yes	339	1.635	3.8	1.964	2.5
No	203	1.470	6.7	1.832	4.4

**Original results from the best performing model (five).* RAPID3: Routine Assessment of Patient Index Data Three; csDMARDS: Disease Modifying Anti-Rheumatic Drugs; RA: Rheumatoid Arthritis

Table 6.22 reports the results of the MAE and RMSE after having run the selected prediction model on specified subgroups of the population. The model estimated more accurate levels of presenteeism for manual workers, individuals who work part-time, individuals who have low disease activity or are in remission, individuals that self-reported being on non-effective medication, and individuals on biologics.

6.5.4 Summary of key results

The analysis presented in this chapter sought to quantify the degree of association (link) between measures of ‘health’ and ‘presenteeism’ to understand whether levels of presenteeism can be accurately predicted using standardised measures such as EQ5D, SF6D or ICECAP-A. The results indicated that the best prediction model for

presenteeism was an OLS model that included the SF6D, age and gender (termed model 5; see equation 12):

$$Predicted\ Presenteeism\ (WPAI) = 11.73 - 11.03\ SF6D - 0.097age - 0.367gender \quad \text{(Equation 12)}$$

Where SF6D is the SF6D index term, age is the age of the individual with RA or AS, and gender is the sex of the patient (1 = female, 0 = male). The variables in this prediction model and the respective coefficients, standard errors, and p-values are presented in Table 6.23.

Table 6.23 Summary of values and statistics from the best performing prediction model (model 5)

Variable	Description	Coefficient	Standard Error	P-value
SF6D	SF6D index term	-11.738	0.694	0.000
Age	Age of the employed individual with RA or AS	-0.103	0.042	0.015
Gender	The sex of the employed individual with RA or AS	0.098	0.172	0.569
Constant	Constant term	11.889	0.538	0.000

SF6D: Short Form Six Dimensions; RA: Rheumatoid Arthritis; AS: Ankylosing Spondylitis

6.6 Discussion

The aim of this study was to quantify the relationship between measures of HRQoL (EQ5D; SF6D; ICECAP-A) and presenteeism (QQ; WPAI). The purpose of quantifying the link between measures of HRQoL and presenteeism was to develop a model that could be used to predict levels of presenteeism for use in economic evaluations of WPIs for populations with RA or AS.

The relationship was initially assessed by comparing the degree of conceptual overlap of three HRQoL measures and presenteeism. Overall correlations for HRQoL and the WPAI and QQ method, respectively, were identified as moderate to strong; a result that has not been reported in previous studies (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014). The observed correlations between the WPAI and the three measures of HRQoL were stronger compared with correlations for the QQ method. A possible reason for the difference in conceptual overlap may be because the WPAI uses a broader concept of presenteeism compared with the QQ method. The QQ method splits the concept of presenteeism into three parts: 1) quantity; 2) quality; and 3) efficiency of work. The QQ method, therefore, asks respondents to think about one attribute of their level of

presenteeism rather than presenteeism as a whole. The WPAI, in comparison, asks respondents to rate their level of work performance allowing the respondent to think about everything they feel is important to regarding how well they are able to work.

This was the first study to explore the extent to which three distinct measures of HRQoL were able to predict levels of presenteeism. Previous studies that have developed and assessed prediction models for presenteeism used the EQ5D-3L only (Lamers et al. 2005; Krol et al. 2014). The results from this study suggest that the SF6D is better predictor of presenteeism compared with the EQ5D-5L for populations with RA or AS. Therefore, the results may suggest the performance of HRQoL measures and their ability to predict levels of presenteeism may differ depending on the health condition being studied. Further investigation is needed to address whether the predictive performance of different HRQoL measures is dependent on the health condition being studied.

The SF6D was identified as the HRQoL measure that was best able to predict levels of presenteeism. The result is particularly well suited for populations with RA or AS for two reasons. The WPAI has been validated as a measure of presenteeism suitable for use in patients with rheumatic conditions (Beaton et al. 2016). The SF-36, the original survey from which the SF6D was derived, is a validated measure (Ruta et al. 1998; Revicki et al. 2011) that has been widely used to capture health status in patients with RA and AS (Matcham et al. 2014; Yang et al. 2016). Existing studies of RA and AS populations that have been conducted in the past or are on-going currently are likely to be collecting SF-36 data meaning that it would be possible to convert SF-36 data into SF6D data and, then potentially, predict levels of presenteeism. The ability to predict presenteeism using data that have already been collected, or are currently being collected, may enable further studies of presenteeism to be conducted without having to set-up a completely new study.

Unlike previous studies that have developed prediction models for presenteeism, the selection of HRQoL measures and self-report measures of presenteeism used to develop the prediction model in this study was done systematically using pre-specified criteria. The prediction models by Krol et al., (2014) and Lamers et al., (2005) justified the selection of the EQ5D-3L on the basis that the measure is recommended to capture HRQoL for economic evaluations of healthcare interventions. The two studies (Lamers et al. 2005; Krol et al. 2014) seemed to have arbitrarily selected the QQ method as the

measure of presenteeism without considering other potential measures. This study, in comparison to previous studies, considered many relevant measures of HRQoL and presenteeism that could have been used to develop the prediction model. The advantage of making the selection of HRQoL and presenteeism measures explicit and transparent is that other researchers working on developing methods for quantifying presenteeism will be to understand and debate the justifications for the selections made.

The subgroup analysis showed that the model was able to more accurately predict levels of presenteeism for subgroups that indicate that the individual may be working with greater disease activity, for example individuals who work part-time may do so because their condition has forced them to, a finding that was identified in chapter 5. However, the most surprising result was that model was most accurate in the subgroup of employees who reported low levels of disease activity or are now in remission. A potential explanation for this may be that these individuals have been left with severe deformities that have limited their ability to work but report that they are now in remission or have low disease activity.

The prediction model may be used to retrospectively predict levels of presenteeism in datasets where the SF6D, or Short-Form 36 (SF-36) (Ware and Sherbourne 1992), age and gender are routinely collected in RA or AS datasets. The SF6D is a particularly well-suited measure of HRQoL for populations with RA or AS because it is frequently used to capture health status in patients with RA and AS (Matcham et al. 2014; Yang et al. 2016). The advantage of retrospectively populating existing RA or AS datasets that do not currently have presenteeism data is that it allows researchers to conduct research on the impact RA or AS has on levels of presenteeism using relatively quick and cost-effective methods compared with setting up a new trial. Studies that use the prediction model developed in this thesis to predict levels of presenteeism within RA or AS populations would also provide evidence of the external validity of the model.

The prediction model may also be used to help employers quantify levels of presenteeism of their workforce. A challenge employers have when attempting to assess productivity levels of their workforce using self-report measure of presenteeism is that employees may be unlikely to provide accurate estimations; employees are, for obvious reasons, incentivised to appear as highly productive members of staff. A potential way employers could assess levels of presenteeism would be to ask their workforce to provide information about their current health status using the SF6D. Information

collected via the SF6D would provide the employer with valuable insight regarding the health of their workforce and, by applying the prediction model, understand the likely impact health status has on levels of presenteeism.

In terms of economic evaluation of WPIs, conducted within the extra-welfarist framework (see chapter 2, section 2.2.2) and from an employer's perspective, productivity, as a non-monetary benefit, could be selected as the appropriate maximand.

Limitations

There were seven limitations of this study, three of which apply to the study sample, one regarding the data, two regarding the measures used and one in relation to the methods used to develop the prediction model.

The sample

The first limitation related to the sample was caused by the recruitment strategy adopted by ResearchNow. ResearchNow identified employed individuals with RA by asking them to self-report themselves as employed and whether they 'suffer' from RA. The problem with this recruitment strategy is that individuals who self-identify may do so wrongly. There is a chance that, due to the confusion of the word 'arthritis' in rheumatoid arthritis, individuals may have identified themselves as having RA but they may potentially have meant osteoarthritis (OA). ResearchNow provided reassurance that individuals are offered two separate categories for RA and OA. However, the number of individuals taking either biologics or DMARDs for their RA was significantly lower than would be expected given that NICE recommend early aggressive treatment of RA with DMARDs and/or biologics. Measures were taken to reduce this affect by asking respondents of the survey to confirm whether they had received a diagnosis of RA from a rheumatologist. Perhaps, it would have been better to incorporate a simple explanation of the difference between 'arthritis' and RA. The results of this study may be influenced by a non-inflammatory arthritic condition such as OA; however, it not expected that the results would be significantly different because RA and OA share many of the same symptoms of pain and stiffness or the joints.

The second limitation of the sample is that, and not surprisingly given the sampling strategy, some groups were under-represented. The number of males with AS who completed the survey was far lower what would have been expected from the total population with AS. Also, the AS sample had comparatively higher income and

education levels compared to the national UK average. This may have been caused by recruiting individuals using the support group NASS; it is understood that individuals from a higher socioeconomic status are more likely to engage with support offered from charities and groups compared with those from a lower socioeconomic background.

The third limitation is related to the generalisability of the results. This study focussed on an RA and AS employed population meaning that the results are not necessarily generalisable to disease areas beyond the two conditions.

The Data

The results from this study inform questions about the associated relationship between HRQoL and presenteeism. A panel dataset would need to be collected to inform questions about causality; to what extent does HRQoL cause presenteeism, for example.

The Measures

Self-reported data is perceived to be a second best solution in situations where objective level data are lacking. This study relies on self-reported data for HRQoL as well as levels of presenteeism. The potential for bias reporting is increased when objective data are unavailable and results are therefore susceptible to over or underestimations of the true value of health status and levels of presenteeism.

HRQoL measures use recall periods that differ compared to the WPAI, which uses a seven day recall period. The EQ5D used a recall period of one day whereas the SF6D used a recall period of four weeks. The ICECAP-A does not explicitly state a specific recall period. It is difficult to say how the various recall periods may have affected the results but it would be conservative to suggest that the results may not be as accurate as they might be had one common recall period been used. It is understood in the literature about self-reported presenteeism measures that different recall periods are partially responsible for causing highly varied estimated levels of presenteeism (Brooks et al. 2010) (see chapter two, section 2.3.3). Therefore, the difference in recall periods of the EQ5D, SF6D, ICECAP and WPAI may have had an effect on the results, but it is not possible to suggest the extent to which the results may have been affected.

The Methods

Ideally the performance of the prediction model would have been assessed using an external dataset, as was done by Krol et al., (2014). Unfortunately, it was not possible to externally validate the performance of the prediction model because, and to our

knowledge, there exists no RA or AS dataset where all three HRQoL measures are also included. In the absence of an external dataset, the K-fold technique was applied to assess model performance. Whilst the result may be slightly different using an external dataset, there is no great reason to doubt the performance results derived using the K-fold method. With that said, there may be potential, in the future, to externally validate the best performing prediction model in a dataset where the SF6D, age and gender are included. A dataset by The National Rheumatoid Arthritis Society (NRAS) is currently collecting data on the SF6D and WPAI. There may be potential to apply the prediction model identified in this study and assess its ability to predict presenteeism using a different RA sample.

6.7 Conclusion

The results of this study provide evidence of a quantifiable relationship between a measure of health status (SF6D) and a measure of presenteeism (WPAI) when the covariates age and gender are included. The strength of the correlations between HRQoL and presenteeism indicated that data from a measure of health status may be used to predict levels of presenteeism in RA or AS populations. This study also developed and tested the performance of multiple prediction models for presenteeism using HRQoL data. The SF6D, using an OLS regression technique, was identified as the HRQoL measure that best predicted levels of presenteeism. However, caution must be taken when applying the model across populations with different levels of disease severity; the average predictions made by the prediction model developed in this study is unlikely to produce results that are useful across individuals with low, medium, or high levels of disease severity.

Chapter 7 Discussion and Conclusions

This chapter discusses the overall findings of the review and empirical studies conducted as part of this thesis (7.1). This chapter also discusses the specific contributions (7.2), the potential implications for different stakeholders (7.3) and areas for future research (7.3). The chapter ends with concluding remarks (7.5).

7.1 Overview of the Thesis

The overall aim of this thesis was to identify and appraise current methods available to quantify presenteeism and, if appropriate, to develop methods suitable for use in economic evaluations of workplace interventions (WPIs) that assume the perspective of the employer. The research was conducted within the context of rheumatoid arthritis (RA), ankylosing spondylitis (AS), and psoriatic arthritis (PsA). The thesis began by explaining the context in which the research was carried out. In England, the National Institute for Health and Care Excellence (NICE) adopts an extra-welfarist perspective where the healthcare sector is considered to be the main decision-maker based on the premise that a healthcare budget is used to fund healthcare interventions (Drummond et al. 2015). NICE explicitly recommends excluding productivity from economic evaluations of healthcare interventions (NICE 2013). Excluding productivity, although potentially understandable from an ethical and philosophical perspective for healthcare interventions (Olsen and Richardson 1999), is problematic when conducting economic evaluations from the perspective of the employer. WPIs are typically funded by the employer and to be consistent with extra-welfarist views, the employers' perspective should be adopted for economic evaluations of WPIs. It is difficult, therefore, to exclude the impact WPIs have on levels of productivity (see results from chapter three, section 3.3.4).

NICE guidelines do not exist to inform economic evaluations conducted from the perspective of the employer; however, NICE public health guidelines do provide alternative recommendations when assessing interventions that need to consider productivity. NICE public health guidelines have proposed two methods, cost-consequence analyses (CCA) and cost-benefit analysis (CBA), which may allow for the inclusion of productivity (NICE 2012b). CCA is difficult to use to inform a single investment decision because it presents the decision-maker with a list of possible outcomes for a given cost (see chapter 2, section 2.1). CBA generally adopts a societal

perspective allowing productivity to be included as a factor in a CBA economic evaluation; it is also, however, a method that converts all costs and benefits into monetary units using wages (see chapter two, section 2.1). Productivity valued using wages is problematic when included in economic evaluations of WPIs because of its potential to influence cost-effectiveness results in favour of those subgroups that earn high salaries (Olsen and Richardson 1999). Employers may, as a result of using CBA methods, make decisions in favour of their most highly paid members of staff; an issue of equity that needs to be avoided. Therefore, neither CCA nor CBA methods are suitable to inform employers of the costs and benefits of WPIs.

The decision was taken to develop a method that quantifies presenteeism, suitable for economic evaluations of WPIs from the perspective of the employer, within the extra-welfarist paradigm for a number of reasons. First, the methods are consistent with the perspective adopted by NICE where health is selected as the appropriate maximand; from the employers' perspective, productivity is selected as the appropriate maximand. Second, extra-welfarism elects a single decision maker, in the context of health this is the health care sector because it is from a healthcare budget that interventions are funded; however, WPIs are frequently funded by employers' budgets therefore an employer's perspective ought to be adopted. The greatest advantage of remaining within an extra-welfarist framework is that the method may be easily translated for use in economic evaluations that also adopt an extra-welfarist perspective. For example, NICE adopts an extra-welfarist perspective, however; NICE is currently exploring the possibility of widening the perspective adopted by economic evaluations to include non-health benefits, for example the project 'extending the QALY' currently being carried out by health economists at The University of Sheffield (see section 2.1.4).Therefore, the method developed to quantify presenteeism in this thesis may be useful, especially in the interim, to provide a way to include presenteeism in economic evaluations that is already consistent with extra-welfarist ideals.

This thesis addressed four research questions: 1) the extent to which self-reported measures of presenteeism are underpinned by economic theory; 2) if, and how, the impact of productivity was included in economic evaluations of WPIs designed for musculoskeletal conditions (MSDs); 3) whether aspects of RA, AS or PsA that decrease an employee's ability to work could be captured using health-related quality of life (HRQoL) questionnaires; and 4) whether HRQoL measures could be used to predict levels of presenteeism suitable for use in economic evaluations of WPIs. Each question

was answered using specific methods including systematic reviews, qualitative thematic methods, and quantitative econometric methods, all of which were presented across four chapters (chapters three, four, five and six). The four chapters, together as a whole, advance the knowledge of methods of presenteeism within economic evaluations of WPIs, funded by employers, in the specific context of people with RA, AS or PsA.

7.2 Applications of the approach

The prediction model was developed to predict an individual's level of presenteeism given their HRQoL, age and gender. The prediction model provides employers with a practical way of monitoring or assessing presenteeism caused by ill-health, in this case specifically RA and AS. Employees have strong incentives not to accurately report their level of presenteeism and may tend to overestimate their ability to work for fear of losing their job. Employers can, instead, ask their employees to fill out the SF6D and link the results to presenteeism using the algorithm developed. The SF6D and presenteeism levels data can then be used to inform economic evaluations from a healthcare and/or employers perspective and given the NICE's interest in widening their perspective to include non-health benefits, the algorithm could be used as a way to inform an economic evaluation of healthcare interventions also.

As mentioned in chapter six, the prediction model may be useful to retrospectively populate already existing RA or AS datasets that have collected health data using the SF6D, SF-36 or SF-12. The SF-36 and SF-12 scores can be mapped to the SF6D (Hanmer 2009) and SF6D results can then be linked to presenteeism. The algorithm may provide a short-term solution to access new data that would have been collected through expensive trials. The prediction model for presenteeism can provide data that can be used to explore and inform research questions about presenteeism in a cost-effective way.

7.3 Contributions to knowledge

The research presented in this thesis makes ten contributions to the evidence base which are summarised here. A strategy to publish and disseminate the findings from this thesis is detailed in Appendix 7.1 and 7.2.

The first contribution to knowledge was reported in chapter three, a systematic review of self-reported measures of presenteeism developed to identify and measure

presenteeism caused by RA, AS and PsA. Numerous systematic reviews of self-report measures of presenteeism focussed on assessing the specific methods used within each measure (recall period, types of questions asked and so on) (Mattke et al. 2007; Brooks et al. 2010) and/or the psychometric properties of each measure (Tang. et al 2015; Ospina. et al 2016). The systematic review presented in Chapter Three assessed the extent to which published measures of presenteeism were developed using economic or any other type of theory. All but one measure of presenteeism was found to be underpinned by economic theory. The findings generated further questions regarding whether, and if so how, presenteeism was identified, measured and valued in economic evaluations of WPIs. Chapter three also identified the need for guidelines to help researchers develop self-reported measures of presenteeism and to improve reporting standards of how the measure was developed; a finding that was not highlighted by previously published systematic reviews of self-reported measures of presenteeism (Mattke et al., 2007; Brooks et al., 2010; Ospina et al., 2015). Chapter three suggested an initial list of items that may need to be considered when developing guidelines for the development and reporting of self-reported measures of presenteeism.

Chapter four used systematic review methods to identify economic evaluations of WPIs to understand whether and how productivity was quantified and included in economic evaluations of WPIs designed for musculoskeletal conditions (MSDs). Absenteeism was included in all identified economic evaluations of WPIs (n=20); presenteeism was included in four only. The results suggested that the concept and methods used to quantify absenteeism were far better understood and simpler compared to the concept of and methods used to quantify presenteeism. Of the four studies that included presenteeism, the methods used to quantify it varied extensively indicating a distinct lack of clarity regarding presenteeism and appropriate methods to quantify it. The results from the systematic review highlighted the need for guidelines to improve the reporting of how productivity was identified, measured, valued and included in the economic evaluation and to avoid double-counting; again, a finding that was not highlighted in previous systematic reviews of presenteeism in economic evaluations (Kigozi et al. 2017; Uegaki et al. 2011). Chapter four also presents a list of potential items that would need to be addressed by guidelines used to improve the reporting of productivity in economic evaluations.

At this stage of the thesis, it became clear that there was a need to develop a simple method for presenteeism that could be incorporated easily and systematically in cost-

effectiveness analyses. A key study, identified in Chapter Four, carried out an analysis exploring the association between health status measured by the EQ5D, short-form 12 (SF-12) and the general health questionnaire (GHQ) (Phillips et al. 2012). The authors found physical health status was particularly important towards understanding and explaining presenteeism. The association between health status and presenteeism was not the sole focus of the study and, therefore, very little of how the association was tested was reported. This, however, led towards the idea of using HRQoL measures to predict levels of presenteeism. The development of a prediction model formed the basis for the two subsequent empirical chapters.

The fifth contribution to knowledge, reported in chapter five, was the identification of the extent to which for three HRQoL measures were able to capture presenteeism caused by three types of inflammatory arthritis (IA). Studies that have quantitatively examined the relationship between HRQoL measures and presenteeism (Brouwer et al. 2005; Lamers et al. 2005; Krol, Stolk, and Brouwer 2014) have not also explored the face validity of HRQoL measures and their ability to capture presenteeism. The results from chapter five indicated that HRQoL measures had the potential to predict levels of presenteeism of employees with RA, AS or PsA. This is the first study to have tested the face validity of HRQoL measures and their ability to capture those important aspects that impact levels of presenteeism.

The sixth contribution, also reported in chapter five, was the degree to which the three HRQoL measures conceptually overlapped. Previous studies have investigated the overlap between the ICECAP-A and other HRQoL measures, including the EQ5D-3L, using quantitative methods (Davis et al. 2005; Keeley et al. 2016; Engel et al. 2017). The results in chapter five provide qualitative evidence that agrees with the results from those previous studies (Davis et al. 2005; Keeley et al. 2016; Engel et al. 2017) which suggest that the ICECAP-A is a complimentary measure to the health status that does not measure the same concept as health status specific measures such as the EQ5D and SF6D.

The seventh contribution comes from chapter six; the development of a prediction model for presenteeism using three measures of HRQoL. One study had previously developed a prediction model specifically for presenteeism using HRQoL data (Krol, Stolk, and Brouwer 2014). The study by Krol et al., used the EQ5D-3L to predict levels of presenteeism measured by the QQ method. The prediction models tested in chapter

six question whether it was appropriate to predict levels of presenteeism using EQ5D-3L data only. Therefore, the SF6D and ICECAP-A were included as alternative measures of HRQoL that could be used to predict levels of presenteeism. The results showed that HRQoL measures other than the EQ5D-5L were potentially more suitable at predicting levels of presenteeism within specific populations; in this case working individuals with IA.

The eighth contribution, reported in chapter six, was an exploration of the relationship between HRQoL and presenteeism. Previous studies have quantitatively explored the relationship using one measure of HRQoL (EQ5D-3L) and one measure of presenteeism including the QQ method (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014) and the Short Form Health and Labour Questionnaire (SF-HLQ) (Bouwmans et al. 2014). The study reported in chapter six, uniquely explored the relationship using three HRQoL measures and two measures of presenteeism including the QQ method and the Work Activity Productivity Impairment (WPAI) questionnaire.

The ninth contribution, reported in chapter six, reports how individuals with different characteristics (subgroups) affect the model's ability to accurately predict presenteeism. The prediction model by Krol et al., (2014) was developed using a sample of the Dutch general population. The prediction model was externally validated in sample of Dutch workers with back pain. Krol et al., (2014) found their prediction model overestimated levels of presenteeism. The results from the subgroups analysis of this thesis and those identified by Krol et al., indicate the need to identify potential factors of a population that significantly influence a prediction model's ability to produce appropriate estimations. Therefore, it can be recommended that researchers who apply prediction models test various subgroups within their population to understand whether the predictions generated can be applied to the entire sample.

The tenth, and final, contribution of this thesis was the development of a method that can be used to predict levels of presenteeism suitable for economic evaluations of WPIs for employees with types of inflammatory arthritis such as RA or AS. The method was developed to overcome the identified practical challenges of quantifying presenteeism suitable for economic evaluations of WPIs within an extra-welfarist framework. Previously designed prediction models for presenteeism were developed in healthcare systems where productivity is explicitly included in economic evaluations of healthcare interventions (Lamers et al. 2005; Krol, Stolk, and Brouwer 2014). Krol et al., (2014)

prediction model, developed in the Netherlands where productivity is included in economic evaluation, for presenteeism sought to predict levels of presenteeism which could then be used to estimate the cost impact of presenteeism. The model developed in this thesis explicitly avoided quantifying the impact of presenteeism as a cost and instead valued it as a non-monetary outcome; a valuation of presenteeism that is consistent with extra-welfarism where productivity costs are excluded.

7.4 Implications of research

The research presented in this thesis has implications for a range of stakeholders including employers, health economists, the Health and Safety Executive (HSE), a body that produces guidance to promote health and well-being at the workplace (HSE 2017). The following section discusses the implications of this research for each stakeholder.

7.2.1 Employers

Employers are not currently adequately informed of the costs and consequences associated with WPIs. Current economic evaluations, reported in chapter four, illustrated how complicated and unsystematic the current evidence base was in collecting information about the costs and consequences of WPIs, especially with regards to presenteeism. Employers are a key stakeholder because they are likely to fund WPIs. The prediction model, developed in this thesis, provides presenteeism information that is comprehensible and accessible for employers to help them understand the benefits of WPIs.

In the context of the United States, the prediction model for presenteeism may be used to help convince employers of the need to invest in their workforce's health in order for their employees to be more productive and healthy. Productivity and its link to health are potentially very important for US employers since many fund their employee's health insurance policies. WPIs that are proven to have an impact on health and an individual's ability to work may provide employers with two benefits: 1) more productive workforce and 2) reduced insurance fees.

7.2.2 Health Economists

The research conducted in this thesis can be a useful resource for health economists in two ways. The findings support the premise that it is possible to argue that in some exceptional cases productivity ought to be included in economic evaluations. In the UK, productivity is explicitly excluded from economic evaluations of healthcare

interventions. Productivity is excluded for two reasons: 1) the potential to make healthcare interventions more favourable in cost-effectiveness analyses for those who are 'highly productive' (earn the most money); and 2) the healthcare sector funds interventions therefore only those costs and benefits that directly impact the healthcare sector should be included. The first argument is addressed in this thesis by valuing presenteeism as a non-monetary outcome. The second argument provides a logical rationale for the inclusion of productivity based on who funds the intervention. In economic evaluations of healthcare interventions NICE recommend a healthcare sector is adopted because a healthcare budget is used to fund interventions. It can, therefore, be argued that in circumstances where an employer's budget is used to fund WPIs an employer perspective should be adopted taking into account those costs and benefits that fall on the employer including the impact on presenteeism.

7.2.3 Health and Safety Executive

The Health and Safety Executive (HSE) is a body that provides employers with advice and guidance to promote health and safety within the workplace (HSE 2017). One of the key focusses of the HSE is to boost productivity by improving health of employees at the workplace. The prediction model developed in this thesis provides the HSE with a method that can be used to assess the health and potential benefits of WPIs. The results estimated by the prediction model may be used to help inform employers via the HSE. Also, quantifying the link between HRQoL and presenteeism provides further evidence of the need to keep employees healthy at the workplace.

7.5 Limitations

Specific limitations of the individual studies are reported in each chapter and are not repeated here. This section discusses the broader limitations of the research conducted in this thesis. There are three limitations related to: 1) the reliance on self-reported methods of HRQoL and presenteeism; 2) the decision rule regarding the cost-effectiveness of WPIs in relation to productivity is not explicit; and 3) how the evidence should be used to inform positive issues regarding presenteeism and not normative arguments regarding how presenteeism ought to be quantified.

7.3.1 Self-reported measures of HRQoL and presenteeism

There is a lack of objective measures that can be used to quantify presenteeism, especially in service type employment sectors where output cannot be measured using

simple counts of goods produced. Self-reported measures of HRQoL and presenteeism require individuals to report their level of health and levels of presenteeism; both of which are highly susceptible to bias (Viswanathan et al. 2008). Therefore, the results of this study are based on 'second-best' solutions to the measurement of HRQoL and presenteeism. Further research is needed to thoroughly assess the psychometric properties of self-report HRQoL and presenteeism measures in order to gauge the reliability of estimates of HRQoL and presenteeism.

7.3.2 Decision Rule

In economic evaluations of healthcare interventions the cost per quality adjusted life year (QALY) (see chapter two, section 2.1.4) is often compared with a cost-effectiveness threshold (Drummond et al. 2015; McCabe, Claxton, and Culyer. 2008). The NICE cost-effectiveness threshold for technology appraisals is set between £20,000 and £30,000 per QALY and represents the healthcare sector's willingness to pay for one extra QALY and, therefore, which can be interpreted as the opportunity cost of health forgone when a decision is made to invest (McCabe, Claxton, and Culyer 2008). Interventions that generate QALYs for less than £20,000 to £30,000 are said to be cost-effective. A limitation of the method proposed to quantify presenteeism using HRQoL is that the decision rule to invest in a WPI is not explicit. The decision rule to invest in WPIs is dependent on the budget constraints of the individual employer. Unlike the National Health Service in England where all interventions are funded from one single budget, employers are their own individual entities that range from large multi-national corporations to small and medium size enterprises (SMEs) each facing different budget constraints. Further research is needed to explore whether it is useful and, if so, possible to define a decision rule, for example a threshold, for investment in WPIs. At present, the method developed in this thesis leaves the employer with the decision of whether to invest or not and is a pragmatic solution because employers understand their own business and whether it can afford to install a WPI for a given improvement in levels of presenteeism.

7.3.3 Potential method for quantifying presenteeism

The research presented in this thesis cannot address or be used to inform questions about how presenteeism ought to be quantified in economic evaluations of WPIs. Value judgements are needed to inform normative arguments regarding the methods that ought to be used to quantify the impact of presenteeism. The results from this thesis do not attempt to make value judgements and therefore address normative issues.

7.4 Future Research

The overall aim of this thesis was to develop a method that quantifies the impact of presenteeism suitable for economic evaluations of WPIs. The research conducted as part of this thesis has identified a number of areas for future research. Areas for future research are summarised below:

1. To test the validity of the prediction model using: 1) an external dataset that contains the WPAI, SF6D (or SF-36), age and gender; or 2) collect another sample using the questionnaire developed and test the performance of the model;

2. This thesis made the assumption that employers want to see the costs and benefits of WPIs presented to them using economic evaluations methods. Currently, there is a paucity of evidence regarding how employers would want to be informed of the costs and benefits of WPIs. Qualitative interviews could be conducted to ask employers more about what information they need and how they would like to have it presented to them to aid them with their decision making;

3. Chapter three found a distinct need for guidelines concerning the development and reporting of self-reported instruments and Chapter four highlighted the need for reporting guidelines for productivity in economic evaluations of WPIs. Moher et al. (2010) provide guidance for the development of health research reporting guidelines. Moher et al., state that the most important first step to developing a guideline is to identify the need for it. Once the need is identified Moher et al., explain the multiple steps required towards developing a guideline which include: developing guidelines using Delphi methods; running a pilot test status; and performing extensive evaluations of the impact and limitations of the guidelines developed;

4. To predict levels of presenteeism related to unpaid work. Presenteeism data related to unpaid work was also collected by the WPAI. The study could be re-run to predict levels of presenteeism related to unpaid work. A comparison of results found in this study and a study looking at unpaid

work would provide further evidence of the relationship between HRQoL and presenteeism;

5. Re-run the analysis using the QQ method rather than the WPAI as the measure of presenteeism. The QQ method includes three measures of presenteeism: 1) the quantity of work completed; 2) the quality of work completed; and 3) degree of efficiency of work completed. All three measures can be used to assess the relationship between HRQoL and presenteeism;

6. To modify the questionnaire for alternative health conditions that are understood to have an impact on presenteeism. For example, the systematic review by Schultz, Chen, and Edington (2009) highlight conditions where presenteeism is understood to be linked. Alternative conditions include asthma, cancer, depression/anxiety, migraine, and chronic pain, to name a few;

7. To investigate the associated relationship between HRQoL and presenteeism using explanatory variables included within the dataset collected. The study would involve selecting a regression model based on presenteeism data and including a list of covariates that explain the associated relationship;

8. A previous study by Mukuria et al. (2017) focussed developing a prediction model for absenteeism related lost productivity using HRQoL data (EQ5D). The dataset collected for the study in chapter six also includes a measure of absenteeism caused directly because of RA or AS. There is scope to develop a prediction model for absenteeism using all three HRQoL measures. A fourth study may involve examining the relationship between HRQoL and presenteeism (WPAI and QQ method) using covariates to identify significant factors that are associated with levels of presenteeism;

9. Informed by the results in chapter five where measures of health status and capability are complements rather than substitutes, prediction models

that use one measure of health status and one measure of capability could be developed and tested for their level of performance. A comparison with the results presented in chapter six may provide evidence of the potential importance of capturing factors beyond health to predict levels of presenteeism;

10. To conduct qualitative interviews with employers to explore more about what they understand about individuals working with chronic conditions such as inflammatory arthritis;

11. To perform think-aloud studies to identify what individuals consider when completing HRQoL measures such as the EQ5D. It would be useful to identify whether individuals take into account their ability to work when filling out health states. If the ability to work is considered, then there may be potential for double-counting of productivity in economic evaluations that have used cost-utility analysis methods.

7.5 Conclusion

Presenteeism is often, appropriately, an ignored factor in economic evaluations of healthcare interventions in England. However, when assuming the perspective of the employer it is appropriate to include presenteeism when assessing the costs and benefits of WPIs. Existing methods used to identify, measure, and value presenteeism do not exist to adequately inform the employer of the productive benefits of WPIs. Current methods that quantify presenteeism vary extensively and produce results that cannot be compared across studies. Therefore, a method suitable for economic evaluations of WPIs from the perspective of the employer was needed.

The thesis aimed to appraise current methods used to identify, measure, and value presenteeism and, if appropriate, develop methods to quantify the impact of presenteeism suitable for economic evaluations of WPIs conducted from the perspective of the employer. The thesis applied a mixed methods approach and produced ten clear contributions to advance knowledge regarding the quantification of presenteeism.

The key contribution of this thesis was the development and initial validation of a model that can be used to predict presenteeism using HRQoL suitable for economic evaluations of WPIs. The results of the thesis clearly suggest that HRQoL measures, in

general, and the SF6D, specifically, can be used to capture and predict levels of presenteeism caused by RA, AS and PsA.

References

Abma, L. van der Klink, C. B. Terwee, B. C. Amick, and U. Bültmann. 2012. 'Evaluation of the Measurement Properties of Self-Reported Health-Related Work-Functioning Instruments among Workers with Common Mental Disorders'. *Scandinavian Journal of Work, Environment & Health* 38 (1):5–18. https://doi.org/10.5271/sjweh.3190.

Acheson, D. n.d. 'Independent Inquiry into Inequalities in Health Report'. Accessed 6 September 2017. https://www.gov.uk/government/publications/independent-inquiry-into-inequalities-in-health-report.

Amick, B. C., D. Lerner, W. H. Rogers, T. Rooney, and J. N. Katz. 2000. 'A Review of Health-Related Work Outcome Measures and Their Uses, and Recommended Measures'. *Spine* 25 (24):3152–60.

Andersen, J., L. Caplan, J. Yazdany, M. L. Robbins, T. Neogi, K. Michaud, K. G. Saag, J. R. O'Dell, and S. Kazi. 2012. 'Rheumatoid Arthritis Disease Activity Measures: American College of Rheumatology Recommendations for Use in Clinical Practice'. *Arthritis Care & Research* 64 (5):640–47. https://doi.org/10.1002/acr.21649.

Anderson, R. 2010. 'Systematic Reviews of Economic Evaluations: Utility or Futility?'. *Health Economics* 19 (3):350–64. https://doi.org/10.1002/hec.1486.

Anguita, D., L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella. 2012. 'The "K" in K-Fold Cross Validation'. *European Symposium on Artifical Neural Networks, Computational Intelligence and Machine Learning*.

Aronsson, G., K. Gustafsson, and M. Dallner. 2000. 'Sick but yet at Work. An Empirical Study of Sickness Presenteeism'. *Journal of Epidemiology and Community Health* 54 (7):502–9. https://doi.org/10.1136/jech.54.7.502.

Austin, P.C, and Steyerberg, E. W. 2014. 'Events per Variable (EPV) and the Relative Performance of Different Strategies for Estimating the out-of-Sample Validity of Logistic Regression Models'. *Statistical Methods in Medical Research* 26 (2): 796–808. https://doi.org/10.1177/0962280214558972.

Bayoumi, A. M. 2004. 'The Measurement of Contingent Valuation for Health Economics'. *PharmacoEconomics* 22 (11): 691–700.

Beaton, D. E., S. Dyer, A. Boonen, S.M. M. Verstappen, R. Escorpizo, D. V. Lacaille, A. Bosworth, et al. 2016. 'OMERACT Filter Evidence Supporting the Measurement of At-Work Productivity Loss as an Outcome Measure in Rheumatology Research'. *The Journal of Rheumatology* 43 (1):214–22. https://doi.org/10.3899/jrheum.141077.

Becker, G., S. 1964. 'Human Capital'.

Berg, B. L. 2001. *Qualitative Research Methods for the Social Sciences*.

Bernaards, C. M., J. E. Bosmans, V. H. Hildebrandt, M. W. van Tulder, and M. W. Heymans. 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'. *Occupational and Environmental Medicine* 68 (4):265–72. https://doi.org/10.1136/oem.2008.045450.

Birch, S, and A Gafni. 1992. 'Cost Effectiveness/utility Analyses: Do Current Decision Rules Lead Us to Where We Want to Be?'. *Journal of Health Economics* 11 (3): 279–96. https://doi.org/10.1016/0167-6296(92)90004-K.

Birch, S., and C.Donaldson. 2003. 'Valuing the Benefits and Costs of Health Care Programmes: Where's the "extra" in Extra-Welfarism?'. *Social Science & Medicine* 56 (5):1121–33. https://doi.org/10.1016/S0277-9536(02)00101-6.

Black, C. 2008. 'Working for a Healthier Tomorrow'. Government.

Boezeman, E. J., J. K. Sluiter, and K. Nieuwenhuijsen. 2015. 'Measuring Work Functioning: Validity of a Weighted Composite Work Functioning Approach'. *Journal of Occupational Rehabilitation* 25 (3):537–42. https://doi.org/10.1007/s10926-014-9560-3.

Bouwmans, C. a. M., P. Vemer, A. van Straten, S. S. Tan, and L. Hakkaart-van Roijen. 2014. 'Health-Related Quality of Life and Productivity Losses in Patients with Depression and Anxiety Disorders'. *Journal of Occupational and Environmental Medicine* 56 (4):420–24. https://doi.org/10.1097/JOM.0000000000000112.

Bouwmans, C., Marieke K., H.Severens, M. Koopmanschap, W. Brouwer, and L. Hakkaart-van Roijen. n.d. 'The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses'. *Value in Health*. Accessed 7 September 2015. https://doi.org/10.1016/j.jval.2015.05.009.

Braakman-Jansen, L. M. A., E. Taal, I. H. Kuper, and M. A. F. J. van de Laar. 2012. 'Productivity Loss due to Absenteeism and Presenteeism by Different Instruments in Patients with RA and Subjects without RA'. *Rheumatology* 51 (2):354–61. https://doi.org/10.1093/rheumatology/ker371.

Brazier, J. E., and R. Edlin and C.McCabe. 2017. 'Health Economics and Cost Consequences Analysis: A Step Back in Time'. *The BMJ*, November. http://www.bmj.com/rapid-response/2011/10/30/health-economics-and-cost-consequences-analysis-step-back-time.

Brazier, J. E., Y. Yang, A. Tsuchiya, and D. L. Rowen. 2009. 'A Review of Studies Mapping (or Cross Walking) Non-Preference Based Measures of Health to Generic Preference-Based Measures'. *The European Journal of Health Economics* 11 (2):215–25. https://doi.org/10.1007/s10198-009-0168-z.

Brazier, J., J. Roberts, and M. Deverill. 2002. 'The Estimation of a Preference-Based Measure of Health from the SF-36'. *Journal of Health Economics* 21 (2):271–92.

Briggs, A. H., and B. J. O'Brien. 2001. 'The Death of Cost-Minimization Analysis?'. *Health Economics* 10 (2):179–84. https://doi.org/10.1002/hec.584.

Brooks, A., S. E. Hagen, S. Sathyanarayanan, A. B. Schultz, and D. W. Edington. 2010. 'Presenteeism: Critical Issues'. *Journal of Occupational and Environmental Medicine* 52 (11):1055–67. https://doi.org/10.1097/JOM.0b013e3181f475cc.

Brouwer, W. B. F., N. J. A. van Exel, M. A. Koopmanschap, and F. F. H. Rutten. 2002. 'Productivity Costs before and after Absence from Work: As Important as Common?'. *Health Policy (Amsterdam, Netherlands)* 61 (2):173–87.

Brouwer, W. B., M. A. Koopmanschap, and F. F. Rutten. 1999. 'Productivity Losses without Absence: Measurement Validation and Empirical Evidence'. *Health Policy (Amsterdam, Netherlands)* 48 (1):13–27.

Brouwer, Werner B. F., A. J. Culyer, N. Job A. van Exel, and F. F. H. Rutten. 2008. 'Welfarism vs. Extra-Welfarism'. *Journal of Health Economics* 27 (2):325–38. https://doi.org/10.1016/j.jhealeco.2007.07.003.

Brouwer, W. B. F., W. Meerding, L. M. Lamers, and J. L. Severens. 2005. 'The Relationship between Productivity and Health-Related QOL: An Exploration'. *PharmacoEconomics* 23 (3):209–18.

Brown, H. E., N. Burton, N. D. Gilson, and W. Brown. 2014. 'Measuring Presenteeism: Which Questionnaire to Use in Physical Activity Research?'. *Journal of Physical Activity & Health* 11 (2):241–48. https://doi.org/10.1123/jpah.2011-0307.

Brown, S. , and J. Gary. Sessions. 1996. 'The Economics of Absence: Theory and Evidence'. *Journal of Economic Surveys* 10 (1):23–53. https://doi.org/10.1111/j.1467-6419.1996.tb00002.x.

Bruijn, C. de., M. Goossens, R. de Bie, A. Ament, J. Geraets, and G-J. Dinant. 2007. 'Cost-Effectiveness of an Education and Activation Program for Patients with Acute and Subacute Shoulder Complaints Compared to Usual Care'. *International Journal of Technology Assessment in Health Care* 23 (01):80–88. https://doi.org/10.1017/S0266462307051604.

Bültmann, U., D.Sherson, J. Olsen, C. L. Hansen, T. Lund, and J. Kilsgaard. 2009. 'Coordinated and Tailored Work Rehabilitation: A Randomized Controlled Trial with Economic Evaluation Undertaken with Workers on Sick Leave due to Musculoskeletal Disorders'. *Journal of Occupational Rehabilitation* 19 (1):81–93. https://doi.org/10.1007/s10926-009-9162-7.

Burton, W. N., C. Y. Chen, D. J. Conti, A. B. Schultz, and D. W. Edington. 2006. 'The Association Between Health Risk Change and Presenteeism Change': *Journal of Occupational and Environmental Medicine* 48 (3):252–63. https://doi.org/10.1097/01.jom.0000201563.18108.af.

Burton, W. N., D. J. Conti, C. Y. Chen, A. B. Schultz, and D. W. Edington. 1999. 'The Role of Health Risk Factors and Disease on Worker Productivity'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 41 (10):863–77.

Cameron, A. C., and P. K Trivedi. 2010. *Microeconometrics Using STATA*. Revised.

Cancelliere, C. J., D. Cassidy, C. Ammendolia, and P. Côté. 2011. 'Are Workplace Health Promotion Programs Effective at Improving Presenteeism in Workers? A Systematic Review and Best Evidence Synthesis of the Literature'. *BMC Public Health* 11 (1):395. https://doi.org/10.1186/1471-2458-11-395.

Canfield, G. W., and D. G. Soash. 1955. 'Presenteeism-a Constructive View'. *Industrial Medicine & Surgery* 24 (9):417–18.

Carreño, A., I. Fernández, X. Badia, C. Varela, and M. Roset. 2011. 'Using HAQ-DI to Estimate HUI-3 and EQ-5D Utility Values for Patients with Rheumatoid Arthritis in Spain'. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 14 (1):192–200. https://doi.org/10.1016/j.jval.2010.11.001.

Clandinin, D. J.. 2006. 'Narrative Inquiry: A Methodology for Studying Lived Experience'. *Research Studies in Music Education* 27 (1):44–54. https://doi.org/10.1177/1321103X060270010301.

Claxton, K., M.Sculpher, and T. Ades. 2004. 'Cost Consequences: Implicit, Opaque and Anti Scientific. A Response.' *BMJ : British Medical Journal*.

Claxton, Karl, Mark Sculpher, Stephen Palmer, and Anthony J Culyer. 2015. 'Causes for Concern: Is Nice Failing to Uphold Its Responsibilities to All NHS Patients?'. *Health Economics* 24 (1): 1–7. https://doi.org/10.1002/hec.3130.

Claxton, K., S. Walker, S. Palmer, and M.Sculpher. 2010. 'Appropriate Perspectives for Health Care Decisions'. Working Paper 054cherp. Centre for Health Economics, University of York. https://ideas.repec.org/p/chy/respap/54cherp.html.

Coast, J. 1999. 'The Appropriate Uses of Qualitative Methods in Health Economics'. *Health Economics* 8 (4):345–53. https://doi.org/10.1002/(SICI)1099-1050(199906)8:4<345::AID-HEC432>3.0.CO;2-Q.

Coast, J. 2004. 'Is Economic Evaluation in Touch with Society's Health Values?'. *BMJ* 329 (7476):1233–36. https://doi.org/10.1136/bmj.329.7476.1233.

Coast, J., T. N. Flynn, L. Natarajan, K. Sproston, J. Lewis, J. J. Louviere, and T. J. Peters. 2008a. 'Valuing the ICECAP Capability Index for Older People'. *Social Science & Medicine (1982)* 67 (5):874–82. https://doi.org/10.1016/j.socscimed.2008.05.015.

Coast, J., R. Smith, and P. Lorgelly. 2008b. 'Should the Capability Approach Be Applied in Health Economics?'. *Health Economics* 17 (6):667–70. https://doi.org/10.1002/hec.1359.

Collins, J. J., C. M. Baase, C. E. Sharda, R. J. Ozminkowski, S. Nicholson, G. M. Billotti, R. S. Turpin, M. Olson, and M. L. Berger. 2005. 'The Assessment of Chronic Health Conditions on Work Performance, Absence, and Total Economic Impact for Employers': *Journal of Occupational and Environmental Medicine* 47 (6):547–57. https://doi.org/10.1097/01.jom.0000166864.58664.29.

Concato, J, P Peduzzi, T Holford R., and A Feinstein R. 1995. 'Importance of Events per Independent Variable in Proportional Hazards Analysis I. Background, Goals, and General Strategy'. *Journal of Clinical Epidemiology* 48 (12): 1495–1501. https://doi.org/10.1016/0895-4356(95)00510-2.

Connolly, D., C. Fitzpatrick, L. O'Toole, M. Doran, and F. R. O'Shea. 2015. 'Impact of Fatigue in Rheumatic Diseases in the Work Environment: A Qualitative Study'. *International Journal of Environmental Research and Public Health* 12 (11):13807–22. https://doi.org/10.3390/ijerph121113807.

Cooksey, R., M.J. Husain, S. Brophy, H. Davies, M. A. Rahman, M. D. Atkinson, C. J. Phillips, and S. Siebert. 2015. 'The Cost of Ankylosing Spondylitis in the UK Using Linked Routine and Patient-Reported Survey Data'. *PLoS ONE* 10 (7). https://doi.org/10.1371/journal.pone.0126105.

Cookson, R. 2005. 'QALYs and the Capability Approach'. *Health Economics* 14 (8):817–29. https://doi.org/10.1002/hec.975.

Cooper, N. J. 2000. 'Economic Burden of Rheumatoid Arthritis: A Systematic Review'. *Rheumatology* 39 (1):28–33. https://doi.org/10.1093/rheumatology/39.1.28.

CRD. 2009. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. Third. CRD, University of York.

Croon, E. M. de, J. K. Sluiter, T. F. Nijssen, B. a. C. Dijkmans, G. J. Lankhorst, and M. H. W. Frings-Dresen. 2004. 'Predictive Factors of Work Disability in Rheumatoid Arthritis: A Systematic Literature Review'. *Annals of the Rheumatic Diseases* 63 (11):1362–67. https://doi.org/10.1136/ard.2003.020115.

Dakin, H., N. Devlin, Y. Feng, N. Rice, Phill O'Neill, and David Parkin. 2014. 'THE INFLUENCE OF COST-EFFECTIVENESS AND OTHER FACTORS ON NICE DECISIONS'. *Health Economics*, September. https://doi.org/10.1002/hec.3086.

Dakin, H. 2013. 'Review of studies mapping from quality of life of clinical measures to EQ5D: an online database. Health and Quality of Life Outcomes. 11:151. HERC database of mapping studies, Version 5.0 (Last updated 16th May 2016). Available at http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies.

Danve, A., A. Reddy, K. Vakil-Gilani, N. Garg, Al. Dinno, and A. Deodhar. 2015. 'Routine Assessment of Patient Index Data 3 Score (RAPID3) Correlates Well with Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) in the Assessment of Disease Activity and Monitoring Progression of Axial Spondyloarthritis'. *Clinical Rheumatology* 34 (1):117–24. https://doi.org/10.1007/s10067-014-2827-4.

Davis, K, S.R. Collins, M.M. Doty, A. Ho, and A. Holmgren. 2005. 'Health and Productivity among U.S. Workers.' *Issue Brief (Commonw Fund).*, no. 856:1–10.

Davison, N, A. Thompson, A. Turner, and K. Payne. 2017. 'Generating EQ-5D-3L Utility Scores from the Dermatology Life Quality Index: A Mapping Study in Psoriasis Patients'. *Value in Health*.

Dean, L. E., G. T. Jones, A. G. MacDonald, C. Downham, R. D. Sturrock, and G. J. Macfarlane. 2014. 'Global Prevalence of Ankylosing Spondylitis'. *Rheumatology* 53 (4):650–57. https://doi.org/10.1093/rheumatology/ket387.

Despiégel, N., N. Danchenko, C. François, B. Lensberg, and M. F. Drummond. 2012. 'The Use and Performance of Productivity Scales to Evaluate Presenteeism in Mood Disorders'. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 15 (8):1148–61. https://doi.org/10.1016/j.jval.2012.08.2206.

Devlin, N. J., and P. F. M. Krabbe. 2013. 'The Development of New Research Methods for the Valuation of EQ-5D-5L'. *The European Journal of Health Economics* 14 (Suppl 1):1–3. https://doi.org/10.1007/s10198-013-0502-3.

Dew, K, V. Keefe, and K. Small. 2005. '"Choosing" to Work When Sick: Workplace Presenteeism'. *Social Science & Medicine* 60 (10):2273–82. https://doi.org/10.1016/j.socscimed.2004.10.022.

Dillman, D. A. 2006. *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Wiley.

Dolan, P., and C. Gudex. 1995. 'Time Preference, Duration and Health State Valuations'. *Health Economics* 4 (4):289–99.

Dolan, P., C. Gudex, P. Kind, and A. Williams. 1996. 'The Time Trade-off Method: Results from a General Population Study'. *Health Economics* 5 (2):141–54. https://doi.org/10.1002/(SICI)1099-1050(199603)5:2<141::AID-HEC189>3.0.CO;2-N.

Driessen, M., J. Bosmans, K. Proper, J. Anema, P. Bongers, and A. van der Beek. 2012. 'The Economic Evaluation of a Participatory Ergonomics Programme to Prevent Low Back and Neck Pain'. *Work (Reading, Mass.)* 41 Suppl 1:2315–20. https://doi.org/10.3233/WOR-2012-0458-2315.

Drummond, M. F., M. J. Sculpher, K, Claxton, G. L. Stoddart, and G. W. Torrance. 2015. 'Methods for the Economic Evaluation of Health Care Programmes'. 4th Edition. Oxford University Press.

Duarte, G. V., C. Faillace, and J. F. de Carvalho. 2012. 'Psoriatic Arthritis'. *Best Practice & Research Clinical Rheumatology*, New Pearls on CTDs, 26 (1):147–56. https://doi.org/10.1016/j.berh.2012.01.003

Ebringer, A. author. 2013. *Ankylosing Spondylitis and Klebsiella*. London: Springer London. http://dx.doi.org/10.1007/978-1-4471-4300-0.

Endicott, J., and J. Nee. 1997. 'Endicott Work Productivity Scale (EWPS): A New Measure to Assess Treatment Effects'. *Psychopharmacology Bulletin* 33 (1):13–16.

Engel, L., D. Mortimer, S. Bryan, S. A. Lear, and D. G. T. Whitehurst. 2017. 'An Investigation of the Overlap Between the ICECAP-A and Five Preference-Based Health-Related Quality of Life Instruments'. *PharmacoEconomics* 35 (7):741–53. https://doi.org/10.1007/s40273-017-0491-7.

Equality Act. 2010. 'Equality Act 2010'. 2010. https://www.gov.uk/definition-of-disability-under-equality-act-2010.

Escorpizo, R., C. Bombardier, A. Boonen, J. M. W. Hazes, D.Lacaille, V. Strand, and D. Beaton. 2007. 'Worker Productivity Outcome Measures in Arthritis.' *The Journal of Rheumatology* 34 (6):1372–80.

European Commission. 2017. 'What Is an SME? - Growth'. Growth. 2017. /growth/smes/business-friendly-environment/sme-definition_en.

EuroQol Group. 1990. 'EuroQol--a New Facility for the Measurement of Health-Related Quality of Life'. *Health Policy (Amsterdam, Netherlands)* 16 (3):199–208.

Evans, C. J. 2004. ‘Health and Work Productivity Assessment: State of the Art or State of Flux?’. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 46 (6 Suppl):S3–11.

Evers, S., M. Goossens, H. de Vet, M. van Tulder, and A. Ament. 2005. ‘Criteria List for Assessment of Methodological Quality of Economic Evaluations: Consensus on Health Economic Criteria’. *International Journal of Technology Assessment in Health Care* 21 (2):240–45.

Feldstein, M. S. 1964. ‘Net Social Benefit Calculation and the Public Investment Decision’. *Oxford Economic Papers* 16 (1): 114–31.

Filipovic, I., D. Walker, F. Forster, and A. S. Curry. 2011. ‘Quantifying the Economic Burden of Productivity Loss in Rheumatoid Arthritis’. *Rheumatology* 50 (6):1083–90. https://doi.org/10.1093/rheumatology/keq399.

Finckh, A. 2009. ‘Treatment of Very Early Rheumatoid Arthritis With Symptomatic Therapy, Disease-Modifying Antirheumatic Drugs, or Biologic Agents: A Cost-Effectiveness Analysis’. *Annals of Internal Medicine* 151 (9):612. https://doi.org/10.7326/0003-4819-151-9-200911030-00006.

Flynn, T. N., E. Huynh, T. J. Peters, H. Al-Janabi, S. Clemens, A. Moody, and J. Coast. 2015. ‘Scoring the Icecap-a Capability Instrument. Estimation of a UK General Population Tariff’. *Health Economics* 24 (3):258–69. https://doi.org/10.1002/hec.3014.

Fransen, J. and P. L. C. M. van Riel. 2009. ‘Outcome Measures in Inflammatory Rheumatic Diseases’. *Arthritis Research & Therapy* 11 (5):244. https://doi.org/10.1186/ar2745.

Gafni, A. 1997. ‘Willingness to Pay in the Context of an Economic Evaluation of Healthcare Programs: Theory and Practice’. *The American Journal of Managed Care* 3 Suppl (May): S21–32.

Gale, N. K., G. Heath, E. Cameron, S. Rashid, and S. Redwood. 2013. ‘Using the Framework Method for the Analysis of Qualitative Data in Multi-Disciplinary Health Research’. *BMC Medical Research Methodology* 13:117. https://doi.org/10.1186/1471-2288-13-117.

Geraets, J. J. X. R., M.E. J. B. Goossens, C. P. C. de Bruijn, I. J. M. de Groot, A. J. S. Köke, R.A. G. B. Pelt, G. Van der Heijden, G-Jan Dinant, and W. J. A. van den Heuvel. 2006. ‘Cost-Effectiveness of a Graded Exercise Therapy Program for Patients with Chronic Shoulder Complaints’. *International Journal of Technology Assessment in Health Care* 22 (1):76–83.

Gilworth, G., M. A. Chamberlain, A. Harvey, A. Woodhouse, J. Smith, M. G. Smyth, and A. Tennant. 2003. ‘Development of a Work Instability Scale for Rheumatoid Arthritis’. *Arthritis and Rheumatism* 49 (3):349–54. https://doi.org/10.1002/art.11114.

Ginty, A.T. 2013. ‘Psychometric Properties’. In *Encyclopedia of Behavioral Medicine*, edited by Marc D. Gellman and J. Rick Turner, 1563–64. Springer New York. http://link.springer.com/referenceworkentry/10.1007/978-1-4419-1005-9_480.

Goetzel, R. Z., Stacey R. MS Long, Ronald J. Ozminkowski, Kevin Hawkins, Shaohung Wang, and Wendy Lynch. 2004. ‘Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers’. *Journal of Occupational* 46 (4):398–412.

Goetzel, R. Z., R. J. Ozminkowski, and S. R. Long. 2003. ‘Development and Reliability Analysis of the Work Productivity Short Inventory (WPSI) Instrument Measuring Employee Health and Productivity’. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 45 (7):743–62. https://doi.org/10.1097/01.jom.0000079085.95532.32.

Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein, eds. 1996. *Cost-Effectiveness in Health and Medicine*. New York: OUP USA.

Gomersall, J. S., Y. T. Jadotte, Y. Xue, S. Lockwood, D. Riddle, and A. Preda. 2015. 'Conducting Systematic Reviews of Economic Evaluations'. *International Journal of Evidence-Based Healthcare* 13 (3):170–78. https://doi.org/10.1097/XEB.0000000000000063.

Gopalakrishnan, S., and P. Ganeshkumar. 2013. 'Systematic Reviews and Meta-Analysis: Understanding the Best Evidence in Primary Healthcare'. *Journal of Family Medicine and Primary Care* 2 (1):9–14. https://doi.org/10.4103/2249-4863.109934.

Grahn, B. E. M., L. A. Borgquist, and C. S. Ekdahl. 2004. 'Rehabilitation Benefits Highly Motivated Patients: A Six-Year Prospective Cost-Effectiveness Study'. *International Journal of Technology Assessment in Health Care* 20 (2):214–21.

Grewal, I., J. Lewis, T. Flynn, J. Brown, J. Bond, and J. Coast. 2006. 'Developing Attributes for a Generic Quality of Life Measure for Older People: Preferences or Capabilities?'. *Social Science & Medicine* 62 (8):1891–1901. https://doi.org/10.1016/j.socscimed.2005.08.023.

Grootendorst, P., D. Marshall, D. Pericak, N. Bellamy, D. Feeny, and G.W. Torrance. 2007. 'A Model to Estimate Health Utilities Index Mark 3 Utility Scores from WOMAC Index Scores in Patients with Osteoarthritis of the Knee'. *The Journal of Rheumatology* 34 (3):534–42.

Gudex, C. 2005. 'The Descriptive System of the EuroQol Instrument'. In *EQ-5D Concepts and Methods: A Developmental History*, edited by Paul Kind, Richard Brooks, and Rosalind Rabin, 19–27. Springer Netherlands. http://link.springer.com/chapter/10.1007/1-4020-3712-0_2.

Hagen, E. M, A. Grasdal, and H. R. Eriksen. 2003. 'Does Early Intervention with a Light Mobilization Program Reduce Long-Term Sick Leave for Low Back Pain: A 3-Year Follow-up Study'. *Spine* 28 (20):2309–15; discussion 2316. https://doi.org/10.1097/01.BRS.0000085817.33211.3F.

Haldorsen, E., M. Håland, A. L. Grasdal, J. S. Skouen, A. E. Risa, K. Kronholm, and H. Ursin. 2002. 'Is There a Right Treatment for a Particular Patient Group? Comparison of Ordinary Treatment, Light Multidisciplinary Treatment, and Extensive Multidisciplinary Treatment for Long-Term Sick-Listed Employees with Musculoskeletal Pain'. *Pain* 95 (1-2):49–63.

Hanmer, J. 2009. 'Predicting an SF-6D Preference-Based Score Using MCS and PCS Scores from the SF-12 or SF-36'. *Value in Health* 12 (6): 958–66. https://doi.org/10.1111/j.1524-4733.2009.00535.x.

Health and Safety Executive. 2017. 'Health and Safety Executive's story' [Accessed 13/12/2017] http://www.hse.gov.uk/aboutus/hse-story.htm

Health and Safety Executive. 2016. 'Musculoskeletal Disorders (MSDs) in Great Britain'. 2016. http://www.hse.gov.uk/statistics/causdis/musculoskeletal/index.htm.

Hemp, P. 2004. 'Presenteeism: At Work—But Out of It'. Harvard Business Review. 2004. https://hbr.org/2004/10/presenteeism-at-work-but-out-of-it.

Hendrikx, J., M. J. de Jonge, J. Fransen, W. Kievit, and P. LCM van Riel. 2016. 'Systematic Review of Patient-Reported Outcome Measures (PROMs) for Assessing Disease Activity in Rheumatoid Arthritis'. *RMD Open* 2 (2):e000202. https://doi.org/10.1136/rmdopen-2015-000202.

Herdman, M., C. Gudex, A. Lloyd, Mf Janssen, P. Kind, D. Parkin, G. Bonsel, and X. Badia. 2011. 'Development and Preliminary Testing of the New Five-Level Version of EQ-5D (EQ-5D-5L)'. *Quality of Life Research* 20 (10):1727–36. https://doi.org/10.1007/s11136-011-9903-x.

Her, M., and A. Kavanaugh. 2012. 'Critical Analysis of Economic Tools and Economic Measurement Applied to Rheumatoid Arthritis'. *Clinical and Experimental Rheumatology* 30 (4):S107–11.

Heuvel, Swenne G. van den, Goedele A. Geuskens, Wendela E. Hooftman, Lando L. J. Koppes, and Seth N. J. van den Bossche. 2009. 'Productivity Loss at Work; Health-Related and Work-Related Factors'. *Journal of Occupational Rehabilitation* 20 (3):331–39. https://doi.org/10.1007/s10926-009-9219-7.

Hillage, J, J. Rick, H. Pilgrim, C. Carroll, and A Booth. 2008. 'Evidence Review 1:Review of the Effectiveness and Cost Effectiveness of Interventions, Strategies, Programmes and Policies to Reduce the Number of Employees Who Move from Short-Term to Long-Term Sickness Absence and to Help Employees on Long-Term Sickness Absence Return to Work'. http://www.nice.org.uk/nicemedia/pdf/PH19evidenceReview1.pdf.

Hlobil, H., K.Uegaki, J. B. Staal, M. C. de Bruyne, T. Smid, and W. van Mechelen. 2007. 'Substantial Sick-Leave Costs Savings due to a Graded Activity Intervention for Workers with Non-Specific Sub-Acute Low Back Pain'. *European Spine Journal* 16 (7):919–24. https://doi.org/10.1007/s00586-006-0283-9.

H. M. Treasury. 2014. 'Budget 2014'. Open Government Licence.

Hodgson, T. A. 1983. 'The State of the Art of Cost-of-Illness Estimates'. *Advances in Health Economics and Health Services Research* 4:129–64.

Hout, B. van, M. F. Janssen, Y-S. Feng, T. Kohlmann, J. Busschbach, D. Golicki, A. Lloyd, L. Scalone, P. Kind, and A. S. Pickard. 2012. 'Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets'. *Value in Health* 15 (5):708–15. https://doi.org/10.1016/j.jval.2012.02.008.

Hout, W. B. van den. 2010. 'The Value of Productivity: Human-Capital versus Friction-Cost Method'. *Annals of the Rheumatic Diseases* 69 Suppl 1 (January):i89–91. https://doi.org/10.1136/ard.2009.117150.

Hoving, J. L., V. Zwieten, M. C. B, M. van der Meer, J. K. Sluiter, and M. H. W. Frings-Dresen. 2013. 'Work Participation and Arthritis: A Systematic Overview of Challenges, Adaptations and Opportunities for Interventions'. *Rheumatology* 52 (7):1254–64. https://doi.org/10.1093/rheumatology/ket111.

Hummer, J., B. Sherman, and N. Quinn. 2002. 'Present and Unaccounted for'. *Occupational Health & Safety (Waco, Tex.)* 71 (4):40–42, 44, 100.

Humphreys, J. H., S. M. M. Verstappen, K. L. Hyrich, J. R. Chipping, T. Marshall, and D. P. M. Symmons. 2013. 'The Incidence of Rheumatoid Arthritis in the UK: Comparisons Using the 2010 ACR/EULAR Classification Criteria and the 1987 ACR Classification Criteria. Results from the Norfolk Arthritis Register'. *Annals of the Rheumatic Diseases* 72 (8):1315–20. https://doi.org/10.1136/annrheumdis-2012-201960.

Husereau, D., M. Drummond, S. Petrou, C. Carswell, D. Moher, D. Greenberg, F. Augustovski, A. H. Briggs, J. Mauskopf, E. Loder, et al. 2013. 'Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force'. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 16 (2):231–50. https://doi.org/10.1016/j.jval.2013.02.002.

Ida, H., K. Nakagawa, M. Miura, K. Ishikawa, and N. Yakura. 2012. '[Development of the Work Limitations Questionnaire Japanese version (WLQ-J): fundamental examination of the reliability and validity of the WLQ-J]'. *Sangyō Eiseigaku Zasshi = Journal of Occupational Health* 54 (3):101–7.

ICECAP-A. 2017. 'ICECAP-A'. [Online] Available at

https://www.birmingham.ac.uk/research/activity/mds/projects/HaPS/HE/ICECAP/ICECAP-A/index.aspx
Ilmarinen, J., K. Tuomi, and M. Klockars. 1997. 'Changes in the Work Ability of Active Employees over an 11-Year Period'. *Scandinavian Journal of Work, Environment & Health* 23 Suppl 1:49–57.
Ilmarinen, J. 2007. 'The Work Ability Index (WAI)'. *Occupational Medicine* 57 (2):160–160. https://doi.org/10.1093/occmed/kqm008.
Inman, R. D. 2012. '273 - The Spondyloarthropathies'. In *Goldman's Cecil Medicine (Twenty-Fourth Edition)*, edited by Lee GoldmanAndrew I. Schafer, 1690–97. Philadelphia: W.B. Saunders. http://www.sciencedirect.com/science/article/pii/B9781437716047002736.
INVOLVE. 2018. 'What is public involvement in research?' http://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/
ISPOR (International Societal for Pharmacoeconomics and Outcomes Research). 2017. 'Pharmacoeconomic Guidelines Around the World'. Accessed 23 May 2017. https://www.ispor.org/PEguidelines/index.asp.
Jackson, C. 2007. 'The General Health Questionnaire'. *Occupational Medicine* 57 (1):79–79. https://doi.org/10.1093/occmed/kql169.
Jacob-Tacken, K. H. M., M. A. Koopmanschap, W. J. Meerding, and J. L. Severens. 2005. 'Correcting for Compensating Mechanisms Related to Productivity Costs in Economic Evaluations of Health Care Programmes'. *Health Economics* 14 (5):435–43. https://doi.org/10.1002/hec.948.
Al-Janabi, H., T. N. Flynn, and J. Coast. 2012. 'Development of a Self-Report Measure of Capability Wellbeing for Adults: The ICECAP-A'. *Quality of Life Research* 21 (1):167–76. https://doi.org/10.1007/s11136-011-9927-2.
Al-Janabi, H., T. J. Peters, J. Brazier, S. Bryan, T. N. Flynn, S. Clemens, A. Moody, and J.Coast. 2013. 'An Investigation of the Construct Validity of the ICECAP-A Capability Measure'. *Quality of Life Research* 22 (7):1831–40. https://doi.org/10.1007/s11136-012-0293-5.
Jenkinson, C., and R. Layte. 1997. 'Development and Testing of the UK SF-12 (short Form Health Survey)'. *Journal of Health Services Research & Policy* 2 (1):14–18.
Jensen, C., C. V. Nielsen, O. K. Jensen, and K. D. Petersen. 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* 38 (13):1059–67. https://doi.org/10.1097/BRS.0b013e31828ca0af.
Jensen, I. B., H. Busch, L. Bodin, J. Hagberg, A. Nygren, and G. Bergström. 2009. 'Cost Effectiveness of Two Rehabilitation Programmes for Neck and Back Pain Patients: A Seven Year Follow-Up'. *Pain* 142 (3):202–8. https://doi.org/10.1016/j.pain.2008.12.015.
Jensen, I. B., G.Bergström, T. Ljungquist, and L.t Bodin. 2005. 'A 3-Year Follow-up of a Multidisciplinary Rehabilitation Programme for Back and Neck Pain'. *Pain* 115 (3):273–83. https://doi.org/10.1016/j.pain.2005.03.005.
Jette, A. M., A. R. Davies, P. D. Cleary, D. R. Calkins, L. V. Rubenstein, A. Fink, J. Kosecoff, R. T. Young, R. H. Brook, and T. L. Delbanco. 1986. 'The Functional Status Questionnaire: Reliability and Validity When Used in Primary Care'. *Journal of General Internal Medicine* 1 (3):143–49.
Johansson, G. and I. Lundberg. 2004. 'Adjustment Latitude and Attendance Requirements as Determinants of Sickness Absence or Attendance. Empirical

Tests of the Illness Flexibility Model'. *Social Science & Medicine (1982)* 58 (10):1857–68. https://doi.org/10.1016/S0277-9536(03)00407-6.
Johns, G.. 2010. 'Presenteeism in the Workplace: A Review and Research Agenda'. *Journal of Organizational Behavior* 31 (4):519–42. https://doi.org/10.1002/job.630.
Jones, C., K. Payne, B. Gannon, and S. M. M. Verstappen. 2016. 'Economic Theory and Self-Reported Measures of Presenteeism in Musculoskeletal Disease'. *Current Rheumatology Reports* 18. https://doi.org/10.1007/s11926-016-0600-1.
Keeley, T., J. Coast, E. Nicholls, N. E. Foster, S. Jowett, and H. Al-Janabi. 2016. 'An Analysis of the Complementarity of ICECAP-A and EQ-5D-3 L in an Adult Population of Patients with Knee Pain'. *Health and Quality of Life Outcomes* 14 (March):36. https://doi.org/10.1186/s12955-016-0430-x.
Kessler, R. C., C. Barber, A. Beck, P. Berglund, P. D. Cleary, D. McKenas, N. Pronk, et al. 2003. 'The World Health Organization Health and Work Performance Questionnaire (HPQ)'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 45 (2):156–74.
Keyser, F. De. 2011. 'Choice of Biologic Therapy for Patients with Rheumatoid Arthritis: The Infection Perspective'. *Current Rheumatology Reviews* 7 (1):77–87. https://doi.org/10.2174/157339711794474620.
Khan, K. A., S. Petrou, O. Rivero-Arias, S. J. Walters, and S. E. Boyle. 2014. 'Mapping EQ-5D Utility Scores from the PedsQL^{TM} Generic Core Scales'. *PharmacoEconomics* 32 (7):693–706. https://doi.org/10.1007/s40273-014-0153-y.
Khan, K. S, R. Kunz, J. Kleijnen, and G. Antes. 2003. 'Five Steps to Conducting a Systematic Review'. *Journal of the Royal Society of Medicine* 96 (3):118–21.
Kigozi, J., S. Jowett, M. Lewis, P. Barton, and J. Coast.. 2017b. 'The Estimation and Inclusion of Presenteeism Costs in Applied Economic Evaluation: A Systematic Review'. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 20 (3):496–506. https://doi.org/10.1016/j.jval.2016.12.006.
King's Fund. 2014. 'An Alternative Guide to the New NHS in England'. The King's Fund. 2014. http://www.kingsfund.org.uk/projects/nhs-65/alternative-guide-new-nhs-england.
Kivimäki, M., J. Head, J. E. Ferrie, H. Hemingway, M. J. Shipley, J. Vahtera, and M. G. Marmot. 2005. 'Working While Ill as a Risk Factor for Serious Coronary Events: The Whitehall II Study'. *American Journal of Public Health* 95 (1):98–102. https://doi.org/10.2105/AJPH.2003.035873.
Klareskog, L., A. I. Catrina, and S. Paget. 2009. 'Rheumatoid Arthritis'. *The Lancet* 373 (9664):659–72. https://doi.org/10.1016/S0140-6736(09)60008-8.
Koopman, C., K. R. Pelletier, J. F. Murray, C. E. Sharda, M. L. Berger, R. S. Turpin, P. Hackleman, P. Gibson, D. Holmes, and T. Bendel. 2002. 'Stanford Presenteeism Scale: Health Status and Employee Prod... : Journal of Occupational and Environmental Medicine'. LWW. 2002. http://journals.lww.com/joem/Fulltext/2002/01000/Stanford_Presenteeism_Scale__Health_Status_and.4.aspx.
Koopmanschap, M. A. 2005. 'PRODISQ: A Modular Questionnaire on Productivity and Disease for Economic Evaluation Studies.' *Expert Review of Pharmacoeconomics & Outcomes Research* 5 (1):23–28. https://doi.org/10.1586/14737167.5.1.23.
Koopmanschap, M. A., and F. F. Rutten. 1994. 'The Impact of Indirect Costs on Outcomes of Health Care Programs'. *Health Economics* 3 (6):385–93.

Koopmanschap, M. A., F. F. Rutten, B. M. van Ineveld, and L. van Roijen. 1995. 'The Friction Cost Method for Measuring Indirect Costs of Disease'. *Journal of Health Economics* 14 (2):171–89.

Kopec, J. A., and J. M. Esdaile. 1998. 'Occupational Role Performance in Persons with Back Pain'. *Disability and Rehabilitation* 20 (10):373–79.

Krol, M. and W. Brouwer. 2014. 'How to Estimate Productivity Costs in Economic Evaluations'. *PharmacoEconomics* 32 (4):335–44. https://doi.org/10.1007/s40273-014-0132-3.

Krol, M., W. Brouwer, and F. Rutten. 2013. 'Productivity Costs in Economic Evaluations: Past, Present, Future'. *PharmacoEconomics* 31 (7):537–49. https://doi.org/10.1007/s40273-013-0056-3.

Krol, M., E.Stolk, and W. Brouwer. 2014. 'Predicting Productivity Based on EQ-5D: An Explorative Study'. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care* 15 (5):465–75. https://doi.org/10.1007/s10198-013-0487-y.

Kumar, R. N., S. L. Hass, J. Z. Li, D. J. Nickens, C. L. Daenzer, and L. K. Wathen. 2003. 'Validation of the Health-Related Productivity Questionnaire Diary (HRPQ-D) on a Sample of Patients with Infectious Mononucleosis: Results from a Phase 1 Multicenter Clinical Trial'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 45 (8):899–907. https://doi.org/10.1097/01.jom.0000083039.56116.79.

Lachenbruch, P. A., and M. R. Mickey. 1968. 'Estimation of Error Rates in Discriminant Analysis'. *Technometrics* 10 (1):1–11. https://doi.org/10.1080/00401706.1968.10490530.

Lambeek, L. C., J. E. Bosmans, B. J. Van Royen, M. W. Van Tulder, W. Van Mechelen, and J. R. Anema. 2010. 'Effect of Integrated Care for Sick Listed Patients with Chronic Low Back Pain: Economic Evaluation alongside a Randomised Controlled Trial'. *BMJ (Clinical Research Ed.)* 341:c6414.

Lamers, L. M., W. J. Meerding, J. L. Severens, and W. B. F. Brouwer. 2005. 'The Relationship between Productivity and Health-Related Quality of Life: An Empirical Exploration in Persons with Low Back Pain'. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation* 14 (3):805–13.

Lee, S., A. Mendelsohn, and E. Sarnes. 2010. 'The Burden of Psoriatic Arthritis'. *Pharmacy and Therapeutics* 35 (12):680–89.

Leggett, S., A. van der Zee-Neuen, A. Boonen, D. Beaton, M. Bojinca, A. Bosworth, S. Dadoun, et al. 2016. 'Content Validity of Global Measures for at-Work Productivity in Patients with Rheumatic Diseases: An International Qualitative Study.' *Rheumatology (Oxford, England)* 55 (8):1364–73. https://doi.org/10.1093/rheumatology/kev435.

Leggett, S., A. B., C. Hoffstetter, A. Boonen, D. Lacaille, C. Mihai, C. Talli, et al. 2017. 'THU0602 Worker Productivity Loss Remains a Major Issue for Patients with Inflammatory Arthritis and Osteoarthritis: Results from the International Eular-pro Worker Productivity Study'. *Annals of the Rheumatic Diseases* 76 (Suppl 2):433–433. https://doi.org/10.1136/annrheumdis-2017-eular.5041.

Lenney, E. 1977. 'Women's Self-Confidence in Achievement Settings'. *Psychological Bulletin* 84 (1):1–13.

Lensberg, B. R, M. F Drummond, N. Danchenko, N. Despiégel, and C. François. 2013. 'Challenges in Measuring and Valuing Productivity Costs, and Their Relevance in Mood Disorders'. *ClinicoEconomics and Outcomes Research: CEOR* 5 (November):565–73. https://doi.org/10.2147/CEOR.S44866.

Lenssinck, M-L. B., A. Burdorf, A. Boonen, M. A. Gignac, J. M. W. Hazes, and J. J. Luime. 2013. 'Consequences of Inflammatory Arthritis for Workplace Productivity Loss and Sick Leave: A Systematic Review'. *Annals of the Rheumatic Diseases* 72 (4):493–505. https://doi.org/10.1136/annrheumdis-2012-201998.

Leon, L., J.A. Jover, G. Candelas, C. Lajas, C.Vadillo, M. Blanco, E. Loza, M. Angel Perez, M. Redondo, and L. Abasolo. 2009. 'Effectiveness of an Early Cognitive–behavioral Treatment in Patients with Work Disability due to Musculoskeletal Disorders'. *Arthritis Care & Research* 61 (7):996–1003. https://doi.org/10.1002/art.24609.

Lerner, D., B. C. Amick, W. H. Rogers, S. Malspeis, K. Bungay, and D. Cynn. 2001. 'The Work Limitations Questionnaire'. *Medical Care* 39 (1):72–85.

Loeppke, R., P. A. Hymel, J. H. Lofland, L. T. H. Pizzi, D. L. Konicki, G. W. Anstadt, C. Baase, J. Fortuna, and T. Scharf. 2003. 'Health-Related Workplace Productivity Measurement: General and Migraine-Specific Recommendations from the ACOEM Expert Panel'. *Journal of Occupational and Environmental Medicine* 45 (4):349–59.

Loeppke, R., M. Taitel, D.Richling, T. Parry, R. C. Kessler, P. Hymel, and D. Konicki. 2007. 'Health and Productivity as a Business Strategy': *Journal of Occupational and Environmental Medicine* 49 (7):712–21. https://doi.org/10.1097/JOM.0b013e318133a4be.

Lofland, J.H., L. Pizzi, and K. D. Frick. 2004. 'A Review of Health-Related Workplace Productivity Loss Instruments'. *PharmacoEconomics* 22 (3):165–84.

Loisel, P., J. Lemaire, S. Poitras, M.-J. Durand, F. Champagne, S. Stock, B. Diallo, and C. Tremblay. 2002. 'Cost-Benefit and Cost-Effectiveness Analysis of a Disability Prevention Model for Back Pain Management: A Six Year Follow up Study'. *Occupational and Environmental Medicine* 59 (12):807–15.

Long, S.. J, and J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. Second. Stata Press.

Longworth, L. and D. Rowen. 2013. 'Mapping to Obtain EQ-5D Utility Values for Use in NICE Health Technology Assessments'. *Value in Health* 16 (1):202–10. https://doi.org/10.1016/j.jval.2012.10.010.

Lorgelly, P. K. 2015. 'Choice of Outcome Measure in an Economic Evaluation: A Potential Role for the Capability Approach'. *PharmacoEconomics* 33 (8):849–55. https://doi.org/10.1007/s40273-015-0275-x.

Luce, B. R., and K. Simpson. 1995. 'Methods of Cost-Effectiveness Analysis: Areas of Consensus and Debate'. *Clinical Therapeutics* 17 (1): 109–25. https://doi.org/10.1016/0149-2918(95)80012-3.

Lynch, W., W.M. Mercer, and John E. Riedel, eds. 2001. 'Measuring Employee Productivity: A Guide to Self-Assessment Tools'. William M Mercer & The Institute for Health and Productivity Management.

Maddala, G.S. 1999. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press.

Maini, M. Al., F. Adelowo, J. Al Saleh, Y. Al Weshahi, G-R. Burmester, M. Cutolo, Jo. Flood, et al. 2015. 'The Global Challenges and Opportunities in the Practice of Rheumatology: White Paper by the World Forum on Rheumatic and Musculoskeletal Diseases'. *Clinical Rheumatology* 34 (5):819–29. https://doi.org/10.1007/s10067-014-2841-6.

Manning, V. L., B. Kaambwa, J. Ratcliffe, D. L. Scott, E. Choy, M. V. Hurley, and L. M. Bearne. 2015. 'Economic Evaluation of a Brief Education, Self-Management and Upper Limb Exercise Training in People with Rheumatoid Arthritis

(EXTRA) Programme: A Trial-Based Analysis'. *Rheumatology* 54 (2):302–9. https://doi.org/10.1093/rheumatology/keu319.

Martindale, J., R. Shukla, and J. Goodacre. 2015. 'The Impact of Ankylosing Spondylitis/axial Spondyloarthritis on Work Productivity'. *Best Practice & Research Clinical Rheumatology*, Occupation and Musculoskeletal Disorders, 29 (3):512–23. https://doi.org/10.1016/j.berh.2015.04.002.

Martz, E. 2013. 'Handling Multicollinearity in Regression Analysis. Enough Is Enough!'. 2013. http://blog.minitab.com/blog/understanding-statistics/handling-multicollinearity-in-regression-analysis.

Mason, R.L, Gunst, and Hess. 2003. *Statistical Design and Analysis of Experiments: With Applications to Engineering and Science*. John Wiley & Sons.

Matcham, F., I. C. Scott, L. Rayner, M. Hotopf, G. H. Kingsley, S. Norton, D. L. Scott, and S. Steer. 2014. 'The Impact of Rheumatoid Arthritis on Quality-of-Life Assessed Using the SF-36: A Systematic Review and Meta-Analysis'. *Seminars in Arthritis and Rheumatism* 44 (2):123–30. https://doi.org/10.1016/j.semarthrit.2014.05.001.

Mattke, S., A. Balakrishnan, G. Bergamo, and S. J. Newberry. 2007. 'A Review of Methods to Measure Health-Related Productivity Loss'. *The American Journal of Managed Care* 13 (4):211–17.

McCabe, C., K. Claxton, and A. J. Culyer. 2008. 'The NICE Cost-Effectiveness Threshold: What It Is and What That Means'. *PharmacoEconomics* 26 (9):733.

McInnes, I. B., and G. Schett. 2011. 'The Pathogenesis of Rheumatoid Arthritis'. *New England Journal of Medicine* 365 (23):2205–19. https://doi.org/10.1056/NEJMra1004965.

Mease, P. J., and P. S. Helliwell. 2008. *Atlas of Psoriatic Arthritis*. Springer Science & Business Media.

Meijer, E. M., J. K. Sluiter, A. Heyma, K. Sadiraj, and M. H. W. Frings-Dresen. 2006. 'Cost-Effectiveness of Multidisciplinary Treatment in Sick-Listed Patients with Upper Extremity Musculoskeletal Disorders: A Randomized, Controlled Trial with One-Year Follow-Up'. *International Archives of Occupational and Environmental Health* 79 (8):654–64. https://doi.org/10.1007/s00420-006-0098-3.

Menard, S. 1995. *Applied Logistic Regression Analysis: Sage University Series on Quantitative Applications in the Social Sciences.* Thousand Oaks, CA: Sage.

Moher, D., A. Liberati, J. Tetzlaff, and D. G. Altman. 2009. 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement'. *BMJ* 339 (July):b2535. https://doi.org/10.1136/bmj.b2535.

Moher, D., K. F. Schulz, I. Simera, and D. G. Altman. 2010. 'Guidance for Developers of Health Research Reporting Guidelines'. *PLOS Medicine* 7 (2):e1000217. https://doi.org/10.1371/journal.pmed.1000217.

Mokkink, L. B., C. B. Terwee, D. L. Patrick, J. Alonso, P. W. Stratford, D. L. Knol, L. M. Bouter, and H. C. W. de Vet. 2010. 'The COSMIN Checklist for Assessing the Methodological Quality of Studies on Measurement Properties of Health Status Measurement Instruments: An International Delphi Study'. *Quality of Life Research* 19 (4):539–49. https://doi.org/10.1007/s11136-010-9606-8.

Morris, S., N.Devlin, D. Parkin, and A.Spencer. 2012. *Economic Analysis in Health Care*. John Wiley & Sons.

Mukuria, C., D. Rowen, M. Hernández-Alava, S. Dixon, and R. Ara. 2017. 'Predicting Productivity Losses from Health-Related Quality of Life Using Patient Data'. *Applied Health Economics and Health Policy*, March. https://doi.org/10.1007/s40258-017-0326-x.

National Centre for Biotechnology, U. S. National Library of Medicine 8600 Rockville Pike, Bethesda MD, and 20894 Usa. 2017. *PubMed Help*. National Center for Biotechnology Information (US). https://www.ncbi.nlm.nih.gov/books/NBK3827/.

NICE (National Institute for Health and Care Excellence). 2012a. 'Specifying a Service for People Who Need Biologic Drugs for the Treatment of Inflammatory Disease in Rheumatology, Dermatology, and Gasteroenterology'. https://www.nice.org.uk/proxy/?sourceurl=http://www.nice.org.uk/usingguidance/commissioningguides/biologicaltherapies/specifyingserviceforpeoplerequiringbiologicdugs.jsp.

NICE (National Institute for Health and Care Excellence). 2012b. 'Methods for the Development of NICE Public Health Guidance'.

NICE (National Institute for Health and Care Excellence). 2009. 'Rheumatoid Arthritis; Guidance and Guidelines'. https://www.nice.org.uk/guidance/cg79. https://www.nice.org.uk/process/pmg4/chapter/introduction

NICE (National Institute for Health and Care Excellence). 2013. 'Guide to the Methods of Technology Appraisal; Guidance and Guidelines'. http://www.nice.org.uk/article/pmg9/chapter/1-introduction.

NICE (National Institute for Health and Care Excellence). 2014. 'NICE Technology Appraisal Guidance'. https://www.nice.org.uk/About/What-we-do/Our-Programmes/NICE-guidance/NICE-technology-appraisal-guidance.

NICE (National Institute for Health and Care Excellence). 2017.'Methotrexate'. https://bnf.nice.org.uk/medicinal-forms/methotrexate.html

NICE (National Institute for Health and Care Excellence). 2017b. 'NICE Position Statement on the EQ5D-5L'.

NICE (National Institute for Health and Care Excellence). 2006. 'The Public Health Guidance Development Process: An Overview for Stakeholders Including Public Health Practitioners, Policy Makers and the Public'.

NICE (National Institute for Health and Care Excellence). 2010. 'Tocilizumab for the Treatment of Rheumatoid Arthritis' https://www.nice.org.uk/guidance/ta247/chapter/2-the-technology

NICE (National Institute for Health and Care Excellence). 2013. 'Guide to the Methods of Technology Appraisal 2013'. 2013. https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#framework-for-estimating-clinical-and-cost-effectiveness.

Nicholson, S., M. V. Pauly, D. Polsky, C. Sharda, H. Szrek, and M. L. Berger. 2006. 'Measuring the Effects of Work Loss on Productivity with Team Production'. *Health Economics* 15 (2):111–23. https://doi.org/10.1002/hec.1052.

Niemistö, L., P. Rissanen, S. Sarna, T. Lahtinen-Suopanki, K-A. Lindgren, and H. Hurri. 2005. 'Cost-Effectiveness of Combined Manipulation, Stabilizing Exercises, and Physician Consultation Compared to Physician Consultation Alone for Chronic Low Back Pain: A Prospective Randomized Trial with 2-Year Follow-Up'. *Spine* 30 (10):1109–15.

Nieuwenhuijsen, K., R-L. Franche, and F. J. H. van Dijk. 2010. 'Work Functioning Measurement: Tools for Occupational Mental Health Research'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 52 (8):778–90. https://doi.org/10.1097/JOM.0b013e3181ec7cd3.

Noben, C. YG., S. M. A. A. Evers, F. J. Nijhuis, and A. E. de Rijk. 2014. 'Quality Appraisal of Generic Self-Reported Instruments Measuring Health-Related Productivity Changes: A Systematic Review'. *BMC Public Health* 14 (1):115. https://doi.org/10.1186/1471-2458-14-115.

Nussbaum, M. 2003. ‘Capabilities as Fundamental Entitlements: Sen and Social Justice’. *Feminist Economics* 9 (2-3):33–59. https://doi.org/10.1080/1354570022000077926.

Ogdie, A., S. Langan, T. Love, K. Haynes, D. Shin, N. Seminara, N. N. Mehta, A. Troxel, H. Choi, and J. M. Gelfand. 2013. ‘Prevalence and Treatment Patterns of Psoriatic Arthritis in the UK’. *Rheumatology* 52 (3):568–75. https://doi.org/10.1093/rheumatology/kes324.

OHE (Office of Health Economics). 2016. ‘Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England’.

Olsen, J. A., and J. Richardson. 1999. ‘Production Gains from Health Care: What Should Be Included in Cost-Effectiveness Analyses?’. *Social Science & Medicine (1982)* 49 (1):17–26.

ONS (Office for Nationa Statistics). 2015. ‘UK Health Accounts: 2015’

ONS (Office for National Statistics). 2013. ‘Full Report - Graduates in the UK Labour Market 2013’. Government.

ONS (Office for National Statistics). 2010. ‘Standard Occupational Classification’.

ONS (Office for National Statistics). 2007. ‘UK Standard Industrial Classification of Economic Activities’.

Oostenbrink, Jan B., M. A. Koopmanschap, and F. F. H. Rutten. 2002. ‘Standardisation of Costs: The Dutch Manual for Costing in Economic Evaluations’. *PharmacoEconomics* 20 (7):443–54.

Ospina, M. B., L. D., A. Waye, P. Jacobs, and A. H. Thompson. 2015. ‘A Systematic Review of Measurement Properties of Instruments Assessing Presenteeism’. *The American Journal of Managed Care* 21 (2):e171–85.

Osterhaus, J. T., O.Purcaru, and L. Richard. 2009. ‘Discriminant Validity, Responsiveness and Reliability of the Rheumatoid Arthritis-Specific Work Productivity Survey (WPS-RA)’. *Arthritis Research & Therapy* 11:R73. https://doi.org/10.1186/ar2702.

Osterhaus, J. T., D. L. Gutterman, and J. R. Plachetka. 1992. ‘Healthcare Resource and Lost Labour Costs of Migraine Headache in the US’. *PharmacoEconomics* 2 (1):67–76.

Palmer, K. T., E. C. Harris, C. Linaker, M. Barker, W. Lawrence, C. Cooper, and D. Coggon. 2012. ‘Effectiveness of Community- and Workplace-Based Interventions to Manage Musculoskeletal-Related Sickness Absence and Job Loss: A Systematic Review’. *Rheumatology* 51 (2):230–42. https://doi.org/10.1093/rheumatology/ker086.

Pasta, D. 2009. ‘Learning When to Be Discrete: Continuous vs Categorical Predictors’.

Pauly, M. V., S. Nicholson, D. Polsky, M. L. Berger, and C. Sharda. 2008. ‘Valuing Reductions in on-the-Job Illness: “presenteeism” from Managerial and Economic Perspectives’. *Health Economics* 17 (4):469–85. https://doi.org/10.1002/hec.1266.

Pauly, M. V., S. Nicholson, J. Xu, D. Polsky, P. M. Danzon, J. F. Murray, and M. L. Berger. 2002. ‘A General Model of the Impact of Absenteeism on Employers and Employees’. *Health Economics* 11 (3):221–31.

Peduzzi, P., J. Concato, E. Kemper, T. R. Holford, and A. R. Feinstein. 1996. ‘A Simulation Study of the Number of Events per Variable in Logistic Regression Analysis’. *Journal of Clinical Epidemiology* 49 (12): 1373–79.

Philips, Z., L. G., M. Sculpher, K. Claxton, S. Golder, R. Riemsma, N. Woolacott, and J. Glanville. 2004. ‘Review of Guidelines for Good Practice in Decision-Analytic Modelling in Health Technology Assessment’. http://www.ncbi.nlm.nih.gov/books/NBK62255/.

Phillips, C. J., R. Phillips (nee Buck), C. J. Main, P. J. Watson, S. Davies, A. Farr, C. Harper, et al. 2012. 'The Cost Effectiveness of NHS Physiotherapy Support for Occupational Health (OH) Services'. *BMC Musculoskeletal Disorders* 13 (1):29. https://doi.org/10.1186/1471-2474-13-29.

Pincus, T., Y.Yazici, and M.J. Bergman. 2009. 'RAPID3, an Index to Assess and Monitor Patients with Rheumatoid Arthritis, Without Formal Joint Counts: Similar Results to DAS28 and CDAI in Clinical Trials and Clinical Care'. *Rheumatic Disease Clinics* 35 (4):773–78. https://doi.org/10.1016/j.rdc.2009.10.008.

Powell, J. L. 1984. 'Least Absolute Deviations Estimation for the Censored Regression Model'. *Journal of Econometrics* 25 (3):303–25. https://doi.org/10.1016/0304-4076(84)90004-6.

Prasad, M., P. W., R. S., and Y-Chen. T. Shih. 2004a. 'A Review of Self-Report Instruments Measuring Health-Related Work Productivity: A Patient-Reported Outcomes Perspective'. *PharmacoEconomics* 22 (4):225–44.

PriceWaterhouseCooper. 2017. 'Industries'. 2017. https://www.pwc.co.uk/industries.html.

Prochaska, J. O., K. E. Evers, J. L. Johnson, P.H. Castle, J. M. Prochaska, L. E. Sears, E.Y. Rula, and J. E. Pope. 2011. 'The Well-Being Assessment for Productivity: A Well-Being Approach to Presenteeism'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 53 (7):735–42. https://doi.org/10.1097/JOM.0b013e318222af48.

Public Health England. 2014. 'Local Action on Health Inequalities: Workplace Interventions to Improve Health and Well-Being'. PHE publications. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/355773/Briefing5a_Workplace_interventions_health_inequalities.pdf.

Pullenayegum, E.M., J-E. Tarride, F.Xie, R. Goeree, H. C. Gerstein, and D. O'Reilly. 2010. 'Analysis of Health Utility Data When Some Subjects Attain the Upper Bound of 1: Are Tobit and CLAD Models Appropriate?'. *Value in Health* 13 (4):487–94. https://doi.org/10.1111/j.1524-4733.2010.00695.x.

Rafia, R., R. Ara, J. Packham, K. Haywood, and E. Healey. 2012. 'Healthcare Costs and Productivity Losses Directly Attributable to Ankylosing Spondylitis'. *Clinical and Experimental Rheumatology* 30 (2):246–53.

Refaeilzadeh, P., L.Tang, and H. Liu. 2009. 'Cross-Validation'. In *Encyclopedia of Database Systems*, edited by LING LIU and M. TAMER ÖZSU, 532–38. Springer US. http://link.springer.com/referenceworkentry/10.1007/978-0-387-39940-9_565.

Reilly, M. C., A. S. Zbrozek, and E. M. Dukes. 2012. 'The Validity and Reproducibility of a Work Productivity and Activity Impairment Instrument'. *PharmacoEconomics* 4 (5):353–65. https://doi.org/10.2165/00019053-199304050-00006.

ResearchNow. (2017). 'ResearchNow: Online Marketing Company. [online] Available at: https://www.researchnow.com/?lang=gb

Revicki, D. A, A. M. Rentz, M. P. Luo, and R. L Wong. 2011. 'Psychometric Characteristics of the Short Form 36 Health Survey and Functional Assessment of Chronic Illness Therapy-Fatigue Subscale for Patients with Ankylosing Spondylitis'. *Health and Quality of Life Outcomes* 9 (May):36. https://doi.org/10.1186/1477-7525-9-36.

Roelofs, P. D. D. M., S. M. A. Bierma-Zeinstra, M. N. M. van Poppel, W. van Mechelen, B. W. Koes, and M. W. van Tulder. 2010. 'Cost-Effectiveness of Lumbar Supports for Home Care Workers with Recurrent Low Back Pain: An

Economic Evaluation alongside a Randomized-Controlled Trial'. *Spine* 35 (26):E1619–26. https://doi.org/10.1097/BRS.0b013e3181cf7244.

Rogerson, M. D., R. J. Gatchel, and S. M. Bierner. 2010. 'A Cost Utility Analysis of Interdisciplinary Early Intervention versus Treatment as Usual for High-Risk Acute Low Back Pain Patients'. *Pain Practice: The Official Journal of World Institute of Pain* 10 (5):382–95. https://doi.org/10.1111/j.1533-2500.2009.00344.x.

Roijen, L. Van, M-L. Essink-Bot, M. A. Koopmanschap, G. Bonsel, and F. F. H. Rutten. 1996. 'Labor and Health Status in Economic Evaluation of Health Care: The Health and Labor Questionnaire'. *International Journal of Technology Assessment in Health Care* 12 (03):405–15. https://doi.org/10.1017/S0266462300009764.

Roijen, L. van. 2010. 'Handleiding Short Form- Health and Labour Questionnaire (SF-HLQ)'. https://www.scienceopen.com/document?vid=5c474b18-e4f7-4f67-9d39-7dbd2c7d2f55.

Roijen, L. van, M. L. Essink-Bot, M. A. Koopmanschap, G. Bonsel, and F. F. Rutten. 1996. 'Labor and Health Status in Economic Evaluation of Health Care. The Health and Labor Questionnaire'. *International Journal of Technology Assessment in Health Care* 12 (3):405–15.

Roy, J-S, F. Desmeules, and J. C. MacDermid. 2011. 'Psychometric Properties of Presenteeism Scales for Musculoskeletal Disorders: A Systematic Review'. *Journal of Rehabilitation Medicine* 43 (1):23–31. https://doi.org/10.2340/16501977-0643.

RUG (Research User Group). 2017. 'Research User Group (RUG)'. [Online] Available at http://research.bmh.manchester.ac.uk/Musculoskeletal/rug/

Ruta, D. A., N. P. Hurst, P. Kind, M. Hunter, and A. Stubbings. 1998. 'Measuring Health Status in British Patients with Rheumatoid Arthritis: Reliability, Validity and Responsiveness of the Short Form 36-Item Health Survey (SF-36)'. *British Journal of Rheumatology* 37 (4):425–36.

Sanders, G. D., P. J. Neumann, A. Basu, D. W. Brock, D. Feeny, M. Krahn, K. M. Kuntz, et al. 2016. 'Recommendations for Conduct, Methodological Practices, and Reporting of Cost-Effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine'. *JAMA* 316 (10):1093–1103. https://doi.org/10.1001/jama.2016.12195.

Sanghera, S., E. Frew, and T. Roberts. 2015. 'Adapting the CHEERS Statement for Reporting Cost-Benefit Analysis'. *PharmacoEconomics* 33 (5):533–34. https://doi.org/10.1007/s40273-015-0265-z.

Sauerland, S., and C. M. Seiler. 2005. 'Role of Systematic Reviews and Meta-Analysis in Evidence-Based Medicine'. *World Journal of Surgery* 29 (5):582–87. https://doi.org/10.1007/s00268-005-7917-7

Sawtooth Software. 2017. Sawtooth Software & Conjoint Analysis. [online] Available at: http://www.sawtoothsoftware.com/

Schultz, A. B., C-Y. Chen, and D. W. Edington. 2009. 'The Cost and Impact of Health Conditions on Presenteeism to Employers: A Review of the Literature'. *PharmacoEconomics* 27 (5):365–78.

Schultz, A. B., and D. W. Edington. 2007. 'Employee Health and Presenteeism: A Systematic Review'. *Journal of Occupational Rehabilitation* 17 (3):547–79. https://doi.org/10.1007/s10926-007-9096-x.

Schweikert, B., E. Jacobi, R. Seitz, R. Cziske, A. Ehlert, J. Knab, and R. Leidl. 2006. 'Effectiveness and Cost-Effectiveness of Adding a Cognitive Behavioral Treatment to the Rehabilitation of Chronic Low Back Pain'. *The Journal of Rheumatology* 33 (12):2519–26.

Scott, D. L., and G. H. Kingsley. 2007. *Inflammatory Arthritis in Clinical Practice*. Springer Science & Business Media.
Scott, D. L, F. Wolfe, and T. W. J Huizinga. 2010. 'Rheumatoid Arthritis'. *The Lancet* 376 (9746):1094–1108. https://doi.org/10.1016/S0140-6736(10)60826-4.
Sheridan, A. 2004. 'Chronic Presenteeism: The Multiple Dimensions to Men's Absence from Part-Time Work'. *Gender, Work & Organization* 11 (2):207–25. https://doi.org/10.1111/j.1468-0432.2004.00229.x.
Shikiar, R., M. T. Halpern, A.M. Rentz, and Z. M. Khan. 2004. 'Development of the Health and Work Questionnaire (HWQ): An Instrument for Assessing Workplace Productivity in Relation to Worker Health'. *Work (Reading, Mass.)* 22 (3):219–29.
Sieper, J., J. Braun, M. Rudwaleit, A. Boonen, and A. Zink. 2011. 'Overview of Ankyloisng Spondylitis'. Anyklosing Spondylitis in Clinical Practice. Springer.
Simpson, R. 1998. 'Presenteeism, Power and Organizational Change: Long Hours as a Career Barrier and the Impact on the Working Lives of Women Managers'. *British Journal of Management* 9 (s1):37–50. https://doi.org/10.1111/1467-8551.9.s1.5.
Sloman, John, Alison Wride, and Dean Garratt. 2012. *Economics*. 8th ed. Pearson.
Smith, D. J. 1970. 'Absenteeism and "Presenteeism" in Industry'. *Archives of Environmental Health: An International Journal* 21 (5):670–77. https://doi.org/10.1080/00039896.1970.10667313.
Smith, R.D., and T. H. Sach. 2009. 'Contingent valuation: (still) on the road to nowhere?' *Health Economics* 18 (8): 863-866. doi:10.1002/hec.1527
Smith, S. (1994). *Labour economics*. London: Routledge
Smolen, J.S., R.t Landewé, F. C. Breedveld, M. Buch, G. Burmester, M. Dougados, P. Emery, et al. 2013. 'EULAR Recommendations for the Management of Rheumatoid Arthritis with Synthetic and Biological Disease-Modifying Antirheumatic Drugs: 2013 Update'. *Annals of the Rheumatic Diseases*, October, annrheumdis – 2013–204573. https://doi.org/10.1136/annrheumdis-2013-204573.
Speklé, E. M., J. Heinrich, M. J. M. Hoozemans, B. M. Blatter, A. J. van der Beek, J. H. van Dieën, and M. W. van Tulder. 2010. 'The Cost-Effectiveness of the RSI QuickScan Intervention Programme for Computer Workers: Results of an Economic Evaluation alongside a Randomised Controlled Trial'. *BMC Musculoskeletal Disorders* 11 (1):259. https://doi.org/10.1186/1471-2474-11-259.
Squires, H., J. Rick, C. Carroll, and J. Hillage. 2012. 'Cost-Effectiveness of Interventions to Return Employees to Work Following Long-Term Sickness Absence due to Musculoskeletal Disorders'. *Journal of Public Health (Oxford, England)* 34 (1):115–24. https://doi.org/10.1093/pubmed/fdr057.
SSA (Social Security Administration). 2017. 'Code of Federal Regulations'. 2017. https://www.ssa.gov/OP_Home/cfr20/404/404-1567.htm.
Steenstra, I. A., J. R. Anema, M. W. van Tulder, P. M. Bongers, H. C. W. de Vet, and Willem van Mechelen. 2006. 'Economic Evaluation of a Multi-Stage Return to Work Program for Workers on Sick-Leave due to Low Back Pain'. *Journal of Occupational Rehabilitation* 16 (4):557–78. https://doi.org/10.1007/s10926-006-9053-0.
Stewart, W. F., Judith A. R., C. Leotta, and E. Chee. 2004. 'Validation of the Work and Health Interview'. *PharmacoEconomics* 22 (17):1127–40.
Stolz, R. L. 1993. 'Reducing Turnover through Incentive Programs'. *Cornell Hotel and Restaurant Administration Quarterly* 34 (1):79.

Sugden, R. and A.Williams. 1978. *The Principles of Practical Cost-Benefit Analysis*. Oxford University Press.

Swinscow, T. D..V. 1997. *Statistics at Square One*. 9th ed. BMJ.

Symmons, D. P. M., and Sherine E. Gabriel. 2011. 'Epidemiology of CVD in Rheumatic Disease, with a Focus on RA and SLE'. *Nature Reviews. Rheumatology* 7 (7):399–408. https://doi.org/10.1038/nrrheum.2011.75.

Symmons, D., G. Turner, R. Webb, P. Asten, E. Barrett, M. Lunt, D. Scott, and A. Silman. 2002. 'The Prevalence of Rheumatoid Arthritis in the United Kingdom: New Estimates for a New Century'. *Rheumatology* 41 (7):793–800. https://doi.org/10.1093/rheumatology/41.7.793.

Tang, K. 2015. 'Estimating Productivity Costs in Health Economic Evaluations: A Review of Instruments and Psychometric Evidence'. *PharmacoEconomics* 33 (1):31–48. https://doi.org/10.1007/s40273-014-0209-z.

Teijlingen, E. van, and V. Hundley. 2002. 'The Importance of Pilot Studies'. *Nursing Standard (Royal College of Nursing (Great Britain): 1987)* 16 (40):33–36. https://doi.org/10.7748/ns2002.06.16.40.33.c3214.

The University of Sheffield. School of Health and Related Research (ScHARR). 2017. 'Extending the QALY'. Accessed [online] https://scharr.dept.shef.ac.uk/e-qaly/

Tillett, W., C. de-Vries, and N. J. McHugh. 2012. 'Work Disability in Psoriatic Arthritis: A Systematic Review'. *Rheumatology* 51 (2):275–83. https://doi.org/10.1093/rheumatology/ker216.

Tompa, E., R. Dolinschi, and J. Natale. 2013. 'Economic Evaluation of a Participatory Ergonomics Intervention in a Textile Plant'. *Applied Ergonomics* 44 (3):480–87. https://doi.org/10.1016/j.apergo.2012.10.019.

Torrance, G. W., W. H. Thomas, and D. L. Sackett. 1972. 'A Utility Maximization Model for Evaluation of Health Care Programs'. *Health Services Research* 7 (2):118–33.

Torstensen, T. A., A. E. Ljunggren, H. D. Meen, E. Odland, P. Mowinckel, and S. Geijerstam. 1998. 'Efficiency and Costs of Medical Exercise Therapy, Conventional Physiotherapy, and Self-Exercise in Patients with Chronic Low Back Pain. A Pragmatic, Randomized, Single-Blinded, Controlled Trial with 1-Year Follow-Up'. *Spine* 23 (23):2616–24.

Tugwell, P., M. Boers, P. Brooks, L. Simon, V. Strand, and L. Idzerda. 2007. 'OMERACT: An international initiative to improve outcome measures in rheumatology'. *American Journal of Managed Care*. 14 (4): 234

Tugwell, P., L. Idzerda and G. A. Wells. 2008. 'Generic quality-of-life assessment in rheumatoid arthritis'

Turpin, R. S., Ronald J. Ozminkowski, C. E. Sharda, J. J. Collins, M. L. Berger, G. M. Billotti, C. M. Baase, M. J. Olson, and S. Nicholson. 2004. 'Reliability and Validity of the Stanford Presenteeism Scale'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 46 (11):1123–33.

Tu, Y-K., M. Kellett, V. Clerehugh, and M. S. Gilthorpe. 2005. 'Problems of Correlations between Explanatory Variables in Multiple Regression Analyses in the Dental Literature'. *British Dental Journal* 199 (7):4812743. https://doi.org/10.1038/sj.bdj.4812743.

Uegaki, K., M. C. de Bruijne, A. J. van der Beek, W. van Mechelen, and M. W. van Tulder. 2011. 'Economic Evaluations of Occupational Health Interventions from a Company's Perspective: A Systematic Review of Methods to Estimate the Cost of Health-Related Productivity Loss'. *Journal of Occupational Rehabilitation* 21 (1):90–99. https://doi.org/10.1007/s10926-010-9258-0.

Vass, C. M., D. Rigby, and K. Payne. 2017. 'Investigating the Heterogeneity in Women's Preferences for Breast Screening: Does the Communication of Risk Matter?'. *Value in Health* 0 (0). https://doi.org/10.1016/j.jval.2017.07.010.

Verstappen, S. M. M., J. W. J. Bijlsma, H. Verkleij, E. Buskens, A. a. M. Blaauw, E. J. ter Borg, J. W. G. Jacobs, and Utrecht Rheumatoid Arthritis Cohort Study Group. 2004. 'Overview of Work Disability in Rheumatoid Arthritis Patients as Observed in Cross-Sectional and Longitudinal Surveys'. *Arthritis and Rheumatism* 51 (3):488–97. https://doi.org/10.1002/art.20419.

Verstappen, S. M. M. 2015. 'Rheumatoid Arthritis and Work: The Impact of Rheumatoid Arthritis on Absenteeism and Presenteeism'. *Best Practice & Research Clinical Rheumatology*, Occupation and Musculoskeletal Disorders, 29 (3):495–511. https://doi.org/10.1016/j.berh.2015.06.001.

Verstappen, S. M. M., B. Fautrel, S. Dadoun, D. P. M. Symmons, and A.Boonen. 2012. 'Methodological Issues When Measuring Paid Productivity Loss in Patients with Arthritis Using Biologic Therapies: An Overview of the Literature'. *Rheumatology (Oxford, England)* 51 (2):216–29. https://doi.org/10.1093/rheumatology/ker363.

Vilsteren, M. van, C. R. L. Boot, D. L. Knol, D. van Schaardenburg, A. E. Voskuyl, R. Steenbeek, and J. R. Anema. 2015. 'Productivity at Work and Quality of Life in Patients with Rheumatoid Arthritis'. *BMC Musculoskeletal Disorders* 16:107. https://doi.org/10.1186/s12891-015-0562-x.

Visser, B. A., M. C. Ashton, and P. A. Vernon. 2008. 'What Makes You Think You're so Smart?'. *Journal of Individual Differences* 29 (1):35–44. https://doi.org/10.1027/1614-0001.29.1.35.

Ware, J. E., Jr., and C. D. Sherbourne. 1992. 'The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection'. *Medical Care* 30 (6):473–83.

Whitehouse, D. 2005. 'Workplace Presenteeism: How Behavioral Professionals Can Make a Difference'. *Behavioral Healthcare Tomorrow* 14 (1):32–35.

WHO Europe (World Health Organisation Europe). 2017. 'Public Health Services'. 6 September 2017. http://www.euro.who.int/en/health-topics/Health-systems/public-health-services.

Williams, A. 1992. 'Cost-Effectiveness Analysis: Is It Ethical?'. *Journal of Medical Ethics* 18 (1):7–11. https://doi.org/10.1136/jme.18.1.7.

Williams, R. 2017. 'Ordinal Independent Variables'. University of Notre Dame.

Williams, R. M., G. Schmuck, S. Allwood, M. Sanchez, R. Shea, and G. Wark. 2007. 'Psychometric Evaluation of Health-Related Work Outcome Measures for Musculoskeletal Disorders: A Systematic Review'. *Journal of Occupational Rehabilitation* 17 (3):504–21. https://doi.org/10.1007/s10926-007-9093-0.

Wooldridge, J. M. 2016. *Introductory Econometrics: A Modern Approach*. Sixth Edition.

Woolf, A. D. and B. Pfleger. 2003. 'Burden of Major Musculoskeletal Conditions'. *Bulletin of the World Health Organization* 81 (9):646–56.

Worrall, L., C. L. Cooper and F. Campbell. 2000. 'The New Reality for UK Managers: Perpetual Change and Employment Instability'. *Work Employment and Society - WORK EMPLOY SOC* 14 (4):647–68. https://doi.org/10.1177/09500170022118662.

Xie, F., E. M. Pullenayegum, S-C. Li, R. Hopkins, J. Thumboo, and N-N. Lo. 2010. 'Use of a Disease-Specific Instrument in Economic Evaluations: Mapping WOMAC onto the EQ-5D Utility Index'. *Value in Health* 13 (8):873–78. https://doi.org/10.1111/j.1524-4733.2010.00770.x.

Yang, X., D. Fan, Q. Xia, M. Wang, X. Zhang, X. Li, G. Cai, et al. 2016. 'The Health-Related Quality of Life of Ankylosing Spondylitis Patients Assessed by SF-36: A Systematic Review and Meta-Analysis'. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation* 25 (11):2711–23. https://doi.org/10.1007/s11136-016-1345-z.

Zar, J. H. 2014. 'Spearman Rank Correlation: Overview'. In *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat05964/abstract.

Zhang, W., N.Bansback, and A. H. Anis. 2011. 'Measuring and Valuing Productivity Loss due to Poor Health: A Critical Review'. *Social Science & Medicine* 72 (2):185–92. https://doi.org/10.1016/j.socscimed.2010.10.026.

Zhang, W., Nick Bansback, A.Boonen, J. L. Severens, and A. H. Anis. 2012. 'Development of a Composite Questionnaire, the Valuation of Lost Productivity, to Value Productivity Losses: Application in Rheumatoid Arthritis'. *Value in Health* 15 (1):46–54. https://doi.org/10.1016/j.jval.2011.07.009.

Zhang, W., M.A. M. Gignac, D. Beaton, K. Tang, and A. H. Anis. 2010. 'Productivity Loss Due to Presenteeism Among Patients with Arthritis: Estimates from 4 Instruments'. *Journal of Rheumatology* 37 (9):1805–14. https://doi.org/10.3899/jrheum.100123.

Zheltoukhova, K. 2013. 'Musculoskeletal Disorders and Work: Results of a Survey of Individuals Living with Musculoskeletal Disorders in Six European Countries'. The Work Founcation (Lancaster University).

Appendix 1.1

Curr Rheumatol Rep (2016) 18:53
DOI 10.1007/s11926-016-0600-1

HEALTH ECONOMICS AND QUALITY OF LIFE (M HARRISON, SECTION EDITOR)

Economic Theory and Self-Reported Measures of Presenteeism in Musculoskeletal Disease

Cheryl Jones[1,2] · Katherine Payne[1] · Brenda Gannon[1] · Suzanne Verstappen[3,4]

Abstract This study had two objectives: to describe the historical development of self-reported presenteeism instruments that can be used to identify and measure presenteeism as a result of musculoskeletal disease (MSD) and to identify if, and how many of these, presenteeism instruments are underpinned by economic theory. Systematic search methods were applied to identify self-report instruments used to quantify presenteeism caused by MSD. A total of 24 self-reported presenteeism instruments were identified; 24 were designed for use in general health, and 1 was specifically designed for use in rheumatoid arthritis. One generic self-reported presenteeism instrument was explicitly reported to be underpinned by economic theory. Overtime, self-reported presenteeism instruments have become more differentiated and complex by incorporating many different contextual factors that may impact levels of presenteeism. Researchers are encouraged to further develop presenteeism instruments that are underpinned by relevant economic theory and informed by robust empirical research.

Keywords Systematic review · Self-report presenteeism instruments · Musculoskeletal diseases · Economic theory

This article is a part of Topical Collection on *Health Economics and Quality of Life*

✉ Suzanne Verstappen
suzanne.verstappen@manchester.ac.uk

1 Manchester Centre for Health Economics, The University of Manchester, 4th Floor, Jean McFarlane Building, Oxford Road, M13 9PL Manchester, UK

2 Arthritis Research UK/MRC Centre for Musculoskeletal Health and Work, University of Southampton, Southampton, UK

3 Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, Stopford Building, Oxford Road, M13 9PT Manchester, UK

4 NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester, UK

Introduction

Since 1990, the global burden of musculoskeletal diseases (MSDs), including chronic rheumatological conditions, as measured by disability-adjusted life years (DALYs), has been shown to increase dramatically [1]. For many chronic rheumatological conditions, such as rheumatoid arthritis (RA), psoriatic arthritis (PsA) and ankylosing spondylitis (AS), disease onset can occur at any age; however, peak incidence rates have been found to occur between the ages of 40 and 65 years [2]. Previous reviews have estimated that within 13 years after onset RA, the probability of becoming work-disabled is 50 % [3, 4]. Similarly, after 5 years of the onset of AS, up to 13 % of adults are likely to lose their jobs [5]. The reduction in health status in people with these diseases will not only affect their daily functioning, and cause early mortality, but may also have a major impact on their productivity at work (productivity loss). In the last two decades, there have been some advances in the treatment of MSD, but these strategies are not curative and many people still experience productivity loss [6–8].

Productivity can be viewed, within the context of the employment environment, as a measure of technical efficiency that examines how inputs, such as labour and capital (technology), are used to produce outputs of sufficient volume and quality [9]. The productive rate of any individual may be affected by a variety of factors such as job demands, levels of support, working hours, job satisfaction and, perhaps most importantly, poor health. The impact of a reduction in health

Published online: 11 July 2016

status has been directly linked to productivity loss through absenteeism and presenteeism. Absenteeism refers to the time (hours, days, weeks) spent *away* from work because of illness [10–13], and presenteeism refers to the reduction in 'working performance whilst at work due to ill health' [14].

It is relatively easy to objectively quantify absenteeism using simple counts of days away from work. Quantifying the impact of presenteeism is much more challenging, involving two stages: (1) identifying and measuring the volume of unproductive time and (2) valuing the impact of that unproductive time. The lack of available objective measures that can be used to identify and measure presenteeism has led to the development of instruments that rely on self-reports from the individual affected by an adverse health condition, such as a MSD.

A number of reviews have systematically identified a number of instruments, both general health and disease-specific, that are available to self-report presenteeism. Some of these systematic reviews have focused on how to identify and measure presenteeism and found that the available instruments differed extensively and lead to vastly different estimations of the volume of presenteeism [11, 15, 16]. In contrast, the methods used to value the impact of presenteeism has largely focused on using two, similar, methods that centre on using cost as the unit of measurement: the human capital approach (HCA) and the friction cost approach (FCA) [17]. The HCA and FCA value the amount of productivity loss by multiplying the amount of time an individual is unproductive during a working week by the wage rate. The two methods differ in terms of the perspective they take. The HCA calculates the cost of lost productivity from the perspective of the patient/employee. Therefore, the cost of lost productivity continues until that individual employee/patient has found another job. The FCA takes an employer perspective and calculates the cost of lost productivity based on the amount of time it takes to replace the sick employee; this period of time is known as the 'friction' period. Once the sick employee has been replaced, the FCA assumes that initial production levels are restored [17]. The HCA and the FCA are grounded in economic theory that assumes that productivity is equal to the market wage which represents the marginal revenue product of labour of an employee working for an employer in the context of a perfectly competitive market [18]. Two studies by Pauly et al. [19] and Zhang et al. [12] criticise the economic theories that are used to underpin methods that value the cost of lost productivity. Pauly et al. suggested that the allowances for sick days and protection against fluctuating wages will mean that employees will accept a wage rate that is lower than the value of the marginal productivity of the worker. Therefore, the value of productivity loss will exceed the value of the wage because the wage is lower than marginal productivity. Similarly, Zhang et al. argued that the cost of productivity loss will exceed the value of the wage rate if a job involved team-based work, unavailability of substitutes, and produces highly time-sensitive output.

It is important to understand the theoretical underpinning of an approach to quantify the impact of a subjective construct, such as presenteeism, to enable the development of a robust approach to its identification, measurement and valuation. Economic theories provide a common framework from which to develop methods that can be used to identify and measure the impact of presenteeism. To our knowledge, it is currently unknown which, if any, self-reported presenteeism instrument used to identify and measure presenteeism is underpinned by economic theory. This study had two objectives: (1) to describe the historical development of self-reported instruments that can be used to identify and measure presenteeism as a result of a MSD and (2) to identify if, and how many of these, self-reported presenteeism instruments are underpinned by economic theory.

Methods

A systematic review was carried out to identify all published studies that describe the development of self-reported presenteeism instruments that can be used in MSD. The search was run up until November 2015. The systematic review was conducted, in line with advice and guidelines published by the Centre for Reviews and Dissemination (CRD) [20] and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Search Strategy

The search for relevant studies involved updating a recent systematic search conducted by Ospina et al. in 2012 [21•]. Ospina et al. conducted a systematic search that identified all general health and disease-specific presenteeism instruments. The electronic search strategies used by Ospina et al. (see Appendix) were retrieved and re-ran in eight electronic databases including Medline (1946 to September week 3 2015), Embase (1980 to week 40 2015), Cochrane Central Register of Controlled Trials (CENTRAL) (August 2015), PsychINFO (1806 to September week 4 2015), Web of Science (1900 to November 6, 2015), CINAHL (1937 to November 6, 2015), Business Source Complete (1886 to November 6, 2015) and ABI inform (1970 to November 6, 2015). The electronic search strategies comprised of the specific names of presenteeism instruments, such as the 'Endicott work productivity scale' and more generic terms such as 'productivity' and 'presenteeism'. The new search was constrained to run between 1st January, 2012 to 6th November, 2015.

Study Selection Process

All titles and abstracts identified were double screened for inclusion by two independent reviewers (CJ and either KP, BG or SV) and accepted if the study met the inclusion criteria specified in Table 1.

Data Extraction and Synthesis

One reviewer (CJ) extracted the data from each study using a bespoke data collection form to extract author, year and country that the study was completed; name of the presenteeism instrument; aims of the instrument; whether the instrument also measured absenteeism; the structure of the instrument; recall period used; whether estimations of presenteeism using that specific instrument could be converted into monetary values using the HCA or FCA; and clear reporting of the economic theories used to underpin the presenteeism instrument developed. The results were tabulated and summarised as part of a narrative synthesis.

Results

In total, 24 studies that described the development of presenteeism instruments for use in MSD were identified. Of these, the Work Productivity Survey for Rheumatoid Arthritis (WPS-RA) [22] was the only one designed to specifically measure presenteeism associated with a MSD (RA). The remaining 24 presenteeism instruments were developed for use across a wide range of health conditions, including MSD. Figure 1 illustrates the identification and inclusion of relevant studies. A summary of the identified presenteeism instruments are presented in Table 2.

Self-Reported Presenteeism Instruments: a History

The earliest identified measure of presenteeism, the Work Performance Scale (WPS), was designed by Jette et al. in 1986 [23]. The WPS asks the respondent to rate their ability to function physically, mentally and socially. The measure is simple and originally designed for clinical use. In 1993, Reilly et al. (1993) developed the Work Productivity Activity Index (WPAI) which differs substantially to the WPS. The WPAI asks the respondent to state the number of days missed from work and the number of days they found work difficult. The instrument also asks about productivity loss when doing unpaid work.

By the late 90's and early 2000's, presenteeism instruments were being designed to collect additional information regarding the contextual factors of an individual's occupation. For example, the Occupational Role Performance Questionnaire (ORQ) developed by Kopec and Esdaile in 1998 [27] collects information about the individual's job satisfaction, job security and the quality of the relationships they have with their colleagues. The Work Instability Scale (WIS) developed by Gilworth et al. in 2003 [34] asks questions about the respondent's work situation and physical work factors.

In 2004, Stewart et al. [39] developed the Work Health Interview (WHI). The WHI is a telephone interview designed to collect information that can be used to estimate the cost of productivity loss. The interview introduces questions about the type of work tasks individuals are expected to complete as part of their job. The WHI is one of the first instruments that explicitly take into account how job characteristics may affect levels of presenteeism. In 2012, Zhang and colleagues developed the Valuation of Lost Productivity (VOLP) questionnaire [43•], a presenteeism instrument that explicitly takes into account how factors such as team dynamics, availability of perfect substitutes and time sensitivity of outputs either compensate or multiply levels of productivity loss caused by health conditions. Since 2015, presenteeism instruments, including the Composite Work Functioning Approach [44] and the

Table 1 Inclusion criteria

	Inclusion criteria	Exclusion criteria
Study type	Development of method that quantifies presenteeism	Studies that apply the developed method, for example, in economic evaluations
		Studies that test methods of presenteeism in terms of their psychometric properties *and* do not discuss the development of the instrument
Focus	Methods developed for assessing health-related presenteeism	Methods developed for assessing other forms of productivity loss, e.g. shirking
	Methods developed for assessing generic health or musculoskeletal conditions	Methods developed that focus on disease-specific areas except musculoskeletal conditions, e.g. mental health
	Original development of presenteeism methods	Adaptations of methods for use in other countries, e.g. WLQ-J
		Adaptations of methods for use in specific disease areas if the original was developed for general health
Publication type	English language	Foreign languages

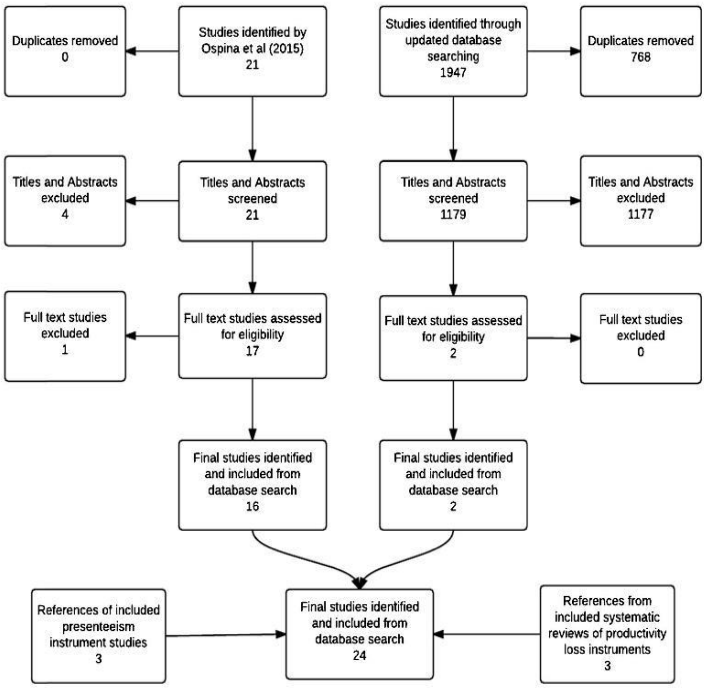


Fig. 1 Flow diagram of study selection process

iMTA Productivity Cost Questionnaire (iPCQ) [45], have been developed using questions from pre-existing measures including the Work Limitations Questionnaire (WLQ) [31], the Health Limitations Questionnaire (HLQ) [25] and the Productivity and Disease Questionnaire (PRODISQ) [46] rather than developing another completely new presenteeism instrument.

Economic Theory Underpinning Presenteeism Instruments

Of the 24 studies that report the development of presenteeism instruments, only one study by Zhang et al. (2012) [43•] discussed how economic theory was used to underpin the design of their presenteeism instrument. They explain the economic rationale behind the development of their presenteeism instrument, the VOLP. The VOLP was designed to identify and measure productivity loss associated with various chronic health conditions and was validated using a sample of employees working with rheumatoid arthritis. Zhang et al. state that the concept of productivity is based on the theory of the production function, where output is a function of inputs including labour, capital and technology. Based on the economic theory of the production function, the authors define productivity loss due to ill health as the output loss associated with reduced labour input. Zhang et al. highlighted that no other existing presenteeism instrument captures both time input loss and information about workplace/job characteristics. The aim of the VOLP is to capture this information so that it can be used to measure productivity loss in terms of output loss associated with reduced labour input caused by ill health. Zhang et al. also critiqued the economics of valuing presenteeism using wage rates. Economic theory states that wages are assumed to be equal to the marginal productivity of workers. However, Zhang et al. argued that wages are often not an accurate reflection of the true value of productivity at the margin because of various other factors such as team production, availability of perfect substitutes and time sensitivity of outputs. Zhang et al. argue that these workplace and job characteristics need to be taken into account explicitly when attempting to measure productivity loss caused by ill health.

The remaining 23 studies stated that the motivation for the development of their presenteeism instruments was based on (1) the need to estimate the impact of presenteeism suitable for economic evaluations of healthcare and workplace interventions (18 studies: [22, 24, 25, 28–33, 35–42, 43•, 45]) and (2) the need to estimate individuals' ability to function at work (5 studies: [23, 26, 27, 34, 44]). No formal theoretical framework of presenteeism from an economic or other relevant discipline, for example psychology, was used to underpin the methods used by the remaining 24 presenteeism instruments.

Table 2 Summary of presenteeism instruments

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
Jette et al., 1986, USA [23]	The Functional Status Questionnaire/ Work Performance Scale (WPS)	Screen disability and monitor clinically meaningful change in function	General	Yes	Four domain scale: 1. Physical functioning 2. Psychological functioning 3. Social/role functioning 4. Six single-item questions Respondents answer a set of statements and assign a grade, for example, 1 = usually did not complete and 4 = no difficulty.	1 month	Not reported	No
Reilly et al., 1993, USA [24]	Work Productivity and Activity Impairment Instrument (WPAI)	To measure the effect of general health and symptom severity on work productivity and regular activities. The WPAI uses function related end-points to allow a measure of the economic impact of relative differences of therapeutic interventions	General	Yes	Questionnaire asks for: 1. Number of days and hours missed from work 2. Days and hours worked 3. Number of days work was difficult 4. Extent to which poor health was attributable to work loss 5. Parallel set of questions (1 to 4) about regular activities of unpaid work Respondent asked to rate their own working performance. Overall work productivity calculated as a %.	7 days	HCA	No
Van Roijen et al., 1996, The Netherlands [25]	Health and Labour Questionnaire (HLQ)	Collect data on relationship between illness, treatment and work performance	General	Yes	Four modules: 1. Absenteesim (paid work) 2. Reduced productivity at work (paid) 3) Unpaid work 4) Impediments to paid and unpaid labour Absenteeism: Respondents asked to state whether they performed paid work or not during past 2 weeks. If large sample mean number of work days lost can be calculated by multiplying by 26 Reduced productivity: 7 items. Respondents asked to rate 1 = never to 4 = always. Score equals the sum of all items. Unpaid production: Respondents asked how many hours spent doing unpaid work. These hours are compared to those of the general population or control group. The mean annual hours of unpaid production calculated by multiplying by 26	2 weeks	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
					Impediments to paid and unpaid labour: Respondents asked to state level of impediment experienced whilst performing the job: 0 = no impediment and 3 = a lot of impediment			
Endicott et al., 1997, USA [26]	Endicott Work Productivity Scale (EWPS)	To assess the extent to which a health condition affects the individual's ability to function at work	General	Yes	25-item questionnaire. Domains: 1. Employment status including self-employed 2. Absenteeism 3. Presenteeism Respondents are asked to rate their performance using a five-point scale: 0 = never and 4 = almost always	1 week	Not reported	No
Kopec and Esdaile, 1998, Canada [27]	Occupational Role Performance (ORQ)	Develop a back pain instrument to measure individual's ability to perform job	General	Not clear	16 items grouped. Six domains: 1. Amount of time working 2. Productivity/efficiency 3. Quality of work 4. Job satisfaction 5. Job security 6. Relations with co-workers Respondents were asked to choose the response they most associated with; no, a little, somewhat and a lot. Overall score calculated as a sum of item scores and converted to a 0 to 100 range.	2 weeks	Not reported	No
Brouwer, Koopmanschap and Rutten, 1999, The Netherlands [28]	The Quality and Quantity Method (QQ)	Measure the quality and quantity of work performed	General	No	Two domains: 1. Quantity of work 2. Quality of work Respondents asked to rate the quantity and quality of work completed using a scale from 1 to 10.	Last working day	HCA	No
Burton et al., 1999, USA [29]	Work Productivity Index (WPI)	To measure decreased productivity associated with health condition by using an objective measure of productivity	General	Yes	Two domains: 1. Time away from job 2. Time lost to maintain the productivity standard Threshold of productivity established: employee score over 0.5 indicates meeting standard productivity = 0 h lost. Scores less than 0.5 evaluated as a proportion of 0.5 and subtracted from 100 %.	1 week	HCA	No
			General	Not clear	Five modules:	4 weeks	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
Amick, et al., 2000 USA [30]	Work Role Functioning Questionnaire (WRFQ)	Through use of literature review, the aim to was discuss advantages and disadvantages of current instruments and to develop the WRFQ			1. Work scheduling 2. Physical demands of jobs 3. Mental demands of jobs 4. Social demands 5. Output demands Respondents asked to state the amount of time they have had difficulty meeting the demands of their work. 100 % all the time, 50 % half of the time and 0 % none			
Lerner et al., 2001, USA [31]	The Work Limitations Questionnaire (WLQ)	To measure on-the-job impact of chronic conditions and treatment on work productivity	General	Yes	25-item questionnaire grouped in four modules: 1. Time management 2. Physical demands 3. Mental/interpersonal demands 4. Output demands Respondents asked to rate the level of difficulty or ability to perform 25 specific job demands.	2 weeks	Not reported	No
Pelletier and Koopman, 2001, USA [32]	Stanford/American Health Association Presenteeism Scale (SAHAPS)	Measures the ability to concentrate on work among employees with health problems	General	No	42 items grouped in two modules: 1. Demographics 2. Presenteeism Respondents are asked to compare their usual performance and rate their performance using 5-point Likert scales and 0–100 % scales	1 month	Not specified but reported as suitable for hourly or salaried occupations	No
Koopman et al., 2002, USA [33]	Stanford Presenteeism Scale-6 (SPS-6)	Developed from the SPS-32 designed to assess the relationship between presenteeism and health problems.	General	No	Six-item questionnaire each designed to capture specific aspects related to presenteeism; 1. Cognitive 2. Emotional 3. Behavioural Respondents are asked to rate their health by stating the degree to which they agree on a scale of 1 to 5: 1 = strongly disagree, 5 = strongly agree.	1 month	Not reported	No
Gilworth et al., 2003, UK [34]	Work Instability Scale (WIS)	Develop a tool that can be used to indicate the level of risk of work disability	General	No	23-item questionnaire capturing information regarding the following: 1. Health 2. Work situation 3. Physical work factors 4. Hobbies The overall score indicates low, medium or at high risk of work disability.	Not Stated	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
					0 = No problems at work, 4 = majority of job is unsuitable and individual is unlikely to cope.			
Goetzel et al., 2003, USA [35]	Work Productivity Short Inventory (WPSI)	Developed to gather information about absenteeism and presenteeism. Also gathers information about productivity loss if acting as the primary caregiver.	General	Yes	Questionnaire asks for the following: 1. Demographics 2. Employment status 3. Absenteeism 4. Presenteeism 5. Productivity loss associated with being a caregiver 6. The questionnaire asks about productivity loss associated with 15 health conditions. The respondent is asked to state number of days with the health condition, number of hours unproductive because of health condition and number of days missed from work. The same questions apply to those who are caregivers.	Three versions and vary only by recall period: 12 months 3 months 2 weeks	HCA	No
Kessler et al., 2003, USA [36]	The World Health Organisation Work Performance Questionnaire (HPQ)	Monetise the workplace costs of illness or cost savings of an intervention.	General	Yes	Three domains: 1. Work performance 2. Absenteeism 3. Job-related accidents Respondents asked to rate their overall work performance using a 0 to 10 scale: 0 = worst possible work performance, 10 = best possible work performance	Various; 1 week 4 weeks	Valuation of lost productive time is discussed. Authors do not recommend one method over another	No
Kumar et al., 2003, USA [37]	Health-Related Productivity Questionnaire-Diary (HRPQ-D)	Developed as a brief, self-administered instrument to be used within clinical trials and survey data collection.	General	Yes	Questionnaire asks for the following: 1. Premature retirement or reduction to part-time work 2. Absenteeism 3. Presenteeism Questionnaire designed in a diary format. Respondents are required to answer all questions for every day of the week. Some responses require the respondent to estimate their unproductive time other require respondent to choose from list of pre-specified %	1 day over 1 week	Not reported	No
Shikar et al., 2004, USA [38]	Health and Work Questionnaire (HWQ)	To assess various aspects of productivity without relying on only self-	General	Yes	24-item questionnaire assessing the following 1. Work quality 2. Work quantity 3. Impatience	1 week	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
		reported estimations			4. Concentration/focus 5. Work satisfaction 6. Non-work satisfaction Respondents are asked to rate their work performance using a 1 to 10 scale: 1 = worst and 10 = best			
Stewart et al., 2004, USA [39]	Work Health Interview (WHI)	To estimate the cost of illness of both absenteeism and presenteeism	General	Yes	Telephone interview. Six modules: 1. Informed consent 2. Employment status 3. Health conditions 4. Tasks and activities performed at work 5. Lost productive time (LPT): absenteeism and presenteeism 6. Demographics Respondents asked to choose the response that most applies, e.g. all of the time, some of the time, half of the time, none of the time. These responses are then converted into %	2 weeks	HCA	No
Turpin et al., 2004, USA [40]	Stanford Presenteeism Scale-13 (SPS-13)	Developed to assess presenteeism on (1) knowledge based and production jobs and (2) provide information on the health condition most likely to affect productivity	General	Yes	13-item questionnaire. Respondent states their primary health condition and is asked to base all answers given this health state. Presenteeism measured using the work impairment score which is the sum of responses to 10 Likert type questions. The final result is presented as a percentage of lost productivity.	4 weeks	Not reported	No
Ilmarinen et al., 2007, Finland [41]	Work Ability Index (WAI)	Assess work ability during health examinations and in workplace surveys	General	Yes	Seven-item questionnaire capturing information regarding the following: 1. Presenteeism 2. Health conditions 3. Absenteeism 4. Mental health Respondents rate their ability to work using various scales (1–10, 1–4 etc.). The index is calculated by summing the ratings given by the respondent.	Varies: 1 year 2 years Lifetime	Not reported	No
Osterhaus, Purcaru and Richard, 2009, USA [22]	Work Productivity Survey for Rheumatoid Arthritis (WPS-RA)	Estimate the productivity limitations associated with RA in paid jobs and unpaid work	MSD	Yes	Three-item questionnaire: 1. Employment status and occupation type	1 month	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
					2. Absenteeism and presenteeism related to paid work 3. Absenteeism and presenteeism related to unpaid work Respondent is asked to rate the extent to which arthritis has interfered with their ability to work using a scale of 0–10: 0 = no interference to 10 = complete interference			
Prochaska et al., 2011, USA [42]	Well-Being Assessment for Productivity (WBA-P)	To create a measure of productivity based on well-being.	General		12 items assess reduced functioning related to personal and work well-being domains: 1. Personal: health, caring for others, financial, personal issues, depressed/stressed 2. Work: lack of resources, issues with co-workers, not enough time, issues with supervisors, lack of training, technical issues Respondents asked to choose response they most associate: not at all, some, a lot. A single number is estimated comprising of 11 items. The score ranges from 0 (not at all) to 100 (a lot for all 11 reasons).	4 weeks	Not reported	No
Zhang et al., 2012, Canada [43•]	The Valuation of Lost Productivity (VOLP)	Explicitly takes into account workplace characteristics necessary for valuing output loss resulting from input loss. The VOLP also used to calculate multipliers and compensation mechanisms.	General	Yes	Five modules: 1. Employment status 2. Absenteeism 3. Presenteeism 4. Unpaid work activity loss 5. Job and workplace characteristics Questionnaire also identifies information: 1. Team dynamics 2. Substitutability of work 3. Time sensitivity of output 4. Compensation 5. Availability of substitutes	7 days	HCA with multiplier Recommended method: Time multiplied by value of time (value of time = wage multiplied by relevant multiplier)	Yes
Boezeman et al., 2015, The Netherlands [44]	Composite Work Functioning Approach	Questionnaire that considers the relative importance of different aspects of work using weights	General	No	Questions based on WLQ and the Tilburg Psychological Contract Questionnaire. 1. Capacity to work 2. Quantity of work 3. Quality of work performance	Not stated	Not reported	No

Table 2 (continued)

Author, year country[a]	Name	Aims	General health or MSD	Absenteeism also measured?	Structure of instrument	Recall period	Monetise productivity loss?	Economic theory
					Respondents are asked to rate their ability to work using a 0 to 4 scale. Weights attached to different aspects of work functioning to reflect relative importance			
Bouwmans et al., 2015, The Netherlands [45]	iMTA Productivity Cost Questionnaire (iPCQ)	To enhance the generalisability and comparability of outcomes for economic evaluations	General	Yes	Questions based on the short form HLQ and the PRODISQ. 18-item questionnaire. 1. Demographics 2. Absenteeism 3. Presenteeism 4. Unpaid work	4 weeks	HCA and FCA	No

[a] Listed in order of development

Discussion

This systematic review has identified a substantial number of self-report presenteeism instruments, and one of these was specifically designed for use in an MSD-related condition (RA). With one exception, the development of the existing instruments was not underpinned by economic theory. Currently, the majority of self-reported instruments are atheoretical, which is problematic because the rationale, construct and development of the instruments cannot be linked. Zhang et al. [43•] was the only study that described how the economic theory of productivity was used to inform the design of the VOLP. The advantage of underpinning the design of the VOLP with economic theory of productivity is that it is clear what the rationale of the VOLP was, how it was designed and how it should be interpreted and used. Zhang et al. [43•] argue that most presenteeism instruments in the literature focus on measuring an individual's labour input by measuring the time spent not working rather than the output lost from reduced labour. It is clear that in the absence of an economic theory of presenteeism, the interpretation of presenteeism from an economic viewpoint is contentious.

The absence of economic theory used to support the instrument to identify and measure presenteeism may have also contributed to the way in which researchers have approached the development of presenteeism instruments. The lack of a theoretical model for presenteeism means that researchers do not have a common framework from which to begin their research and develop their ideas. Therefore, as research into the measurement of presenteeism has grown, the instruments developed to quantify presenteeism have become more differentiated and more complex. The presenteeism instruments developed in the 1980s and 1990s are relatively simple where the respondent is asked to give information about their perceived level of absenteeism and presenteeism based on their health condition. In comparison, those presenteeism instruments developed in the late 2000s ask the respondent to consider a wide range of factors including job characteristics, team dynamics, time-sensitive output, job satisfaction, job security and relationships with colleagues, as well as the direct impact on presenteeism caused by their health condition. Ospina et al. [21•] and Noben et al. [47] recommended that the development of more self-reported presenteeism instruments is not needed and instead the literature should focus their effort on improving the ones that already exist. To some extent, this is happening where the two latest presenteeism instruments, the composite work functioning approach and the iPCQ, use questions from pre-existing presenteeism instruments. However, it is not yet clear which instruments are the most appropriate for measuring presenteeism in the context of health conditions in

general, and MSD, specifically. The OMERACT group are currently working towards recommending which of the available measures is best used in the context of rheumatoid arthritis [48]. However, taking into account economic theory suggests the need to define the best measure in terms of clearly specifying the three constituent parts (identification, measurement, valuation) to quantify the impact of presenteeism.

Limitations

The studies that were used to identify whether or not presenteeism instruments were developed using economic theory did not provide extensive detail. Many studies provided limited information that described how the presenteeism instrument was created. In an area where the quantification of a concept is subjective, such as presenteeism, it should be encouraged that researchers publish information about the conceptualisation and development of their presenteeism instrument. Such information would help inform the correct application and interpretation of their instrument, especially in the absence of applying an economic framework from which to underpin the instrument.

Conclusions

This review has systematically identified all self-reported presenteeism instruments, providing a historical context and whether presenteeism instruments are underpinned by economic theory. With the exception of the VOLP, none of the instruments are explicitly underpinned by economic theory. One key area for further research is to take account of the need to understand how to identify, quantify and value the impact of presenteeism, while underpinning these stages with relevant economic theory for each constituent part of this process. Economic theory would aid the correct interpretation and application of the self-report presenteeism instrument and valuation approach. It is also vital that further development of presenteeism instruments are informed by robust empirical studies that take account of the context in which the final instrument will be used.

Acknowledgments This work was supported by Arthritis Research UK and the Medical Research Council [20665].

Compliance with Ethical Standards

Conflicts of Interest CJ, KP, BG and SV declare that they have no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Appendix. Electronic search strategies

OVID databases:
Medline (1946 to September week 3 2015),
Embase (1980 to week 40 2015),
CENTRAL (August 2015)
PsycINFO (1806 to September week 4 2015)

1. American Productivity Audit.tw.
2. "angina-related limitations at work question?aire".tw.
3. endicott work productivity scale.tw.
4. "health and labo?r question?aire".tw.
5. "health and performance question?aire".tw.
6. "health and productivity question?aire".tw.
7. "health and work question?aire".tw.
8. health-related productivity question?aire*.tw.
9. "lam employment absence and productivity scale".tw.
10. (migraine disability assessment adj2 (question?aire or survey or scale or score*)).tw.
11. MIDAS.tw. and migraine*.mp.
12. 10 or 11
13. (productiv* or presenteeism or absenteeism or work* or employ*).mp.
14. 12 and 13
15. "migraine work and productivity loss question?aire".tw.
16. (osterhaus and (work* or productivity or presenteeism)).tw.
17. (osterhaus adj3 technique).tw.
18. "productivity and disease question?aire".tw.
19. PRODISQ.tw.
20. (quantity adj2 quality adj (method or instrument)).tw.
21. (Stanford* adj5 Presenteeism Scale).tw.
22. "work and health interview".tw.
23. work performance scale.tw.
24. "work productivity and activity impairment*".tw.
25. WPAI*.tw.
26. (US National Health and Wellness Survey).tw. and (productiv* or presenteeism or absenteeism or work*).mp.
27. "health and work performance question?aire".tw.
28. work productivity short inventory.tw.
29. wellness inventory.tw.
30. work role functioning question?aire.tw.
31. (or/1–9) or (or/14–30)
32. limit 31 to english language
33. (work productivity survey or WPS-RA).tw.
34. valuation of lost productivity.tw.
35. work limitations question?aire.tw.
36. (33 or 34) not 32
37. limit 36 to english language

38. 35 not 32
39. limit 38 to english language
40. 32 or 37
41. 32 or 37 or 39

ISI platform databases:
Web of Science (1900–November 6, 2015)

1 TS="angina-related limitations at work questionnaire"
2 ((TS="endicott work productivity scale")) AND Language=(English)
3 (TS=("health and labor questionnaire" OR "health and labour questionnaire")) AND Language=(English)
4 ((TS="health and productivity questionnaire")) AND Language=(English)
5. ((TS="health and work questionnaire")) AND Language=(English)
6. ((TS="health-related productivity questionnaire")) AND Language=(English)
7. ((TS="lam employment absence and productivity scale")) AND Language=(English)
8. 8 ((TS="migraine work and productivity loss questionnaire")) AND Language=(English)
9. ((TS="Stanford Presenteeism Scale")) AND Language=(English)
10. ((TS="health and work performance questionnaire")) AND Language=(English)
11. ((TS="work and health interview")) AND Language=(English)
12. ((TS="work performance scale")) AND Language=(English)
13. ((TS="work productivity short inventory")) AND Language=(English)
14. ((TS="American Productivity Audit")) AND Language=(English)
15. ((TS=(osterhaus and (work or productivity)))) AND Language=(English)
16. TS=("work productivity and activity impairment*")
17. TS=WPAI*
18. TS="wellness inventory"
19. TS=("work productivity survey")
20. TS=("Work role functioning*" SAME limitations)
21. TS=(osterhaus SAME technique)
22. TS=("productivity and disease questionnaire")
23. (TS=(migraine disability assessment SAME (score* OR scale OR questionnaire OR survey))) AND Language=(English)
24. (TS=(MIDAS AND migraine)) AND Language=(English)
25. (TS=Osterhaus) AND Language=(English)
26. (#23 OR #24 OR #25) AND Language=(English)
27. (TS = (work* or productivity or performance or presenteeism or absenteeism or employ*)) AND Language=(English)
28. (#26 AND #27) AND Language=(English)
29. (TS=("work role functioning questionnaire")) AND Language=(English)
30. (TS="health and performance questionnaire") AND Language=(English)
31. TS=("valuation of lost productivity")
32. TS=("work limitations questionnaire")
33. (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #28 OR #29 OR #30 OR #31 OR #32) AND Language=(English)

EBSCO databases:
CINAHL (1937–November 6, 2015)
Business Source Complete (1886–November 6, 2015)

Search ID#	Search terms
S28	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25
S25	"work limitations questionnaire"
S24	"valuation of lost productivity"
S23	"work productivity survey"
S22	"health and performance questionnaire"
S21	"productivity and disease questionnaire"
S20	osterhaus AND (technique or presenteeism or absenteeism or productivity or work* or employ*)
S19	"Work role functioning questionnaire"
S18	"wellness inventory"
S17	WPAI*
S16	"work productivity and activity impairment*"
S15	"American Productivity Audit"
S14	"work productivity short inventory"
S13	"work performance scale"
S12	"work and health interview"
S11	"health and work performance questionnaire"
S10	"Stanford Presenteeism Scale"
S9	"migraine work and productivity loss questionnaire"
S8	"lam employment absence and productivity scale"
S7	(("migraine disability assessment" n2 (score* or scale or survey or questionnaire)) OR (MIDAS AND migraine)) AND (work* OR productivity OR presenteeism OR absenteeism OR employ*)
S6	"health-related productivity questionnaire"
S5	"health and work questionnaire"
S4	"health and productivity questionnaire"
S3	("health and labor questionnaire") OR ("health and labour questionnaire")
S2	"endicott work productivity scale"
S1	"angina-related limitations at work questionnaire"

Proquest databases:
ABI Inform (1970–November 6, 2015)

ALL (("American Productivity Audit") OR ("angina-related limitations at work questionnaire") OR ("endicott work productivity scale") OR ("health and labor questionnaire") OR ("health and labour questionnaire") OR ("health and productivity questionnaire") OR ("health and work questionnaire") OR ("health and performance questionnaire") OR ("health-related productivity questionnaire") OR ("lam employment absence and productivity scale") OR ("migraine work and productivity loss questionnaire") OR ("Stanford

Presenteeism Scale") OR ("work performance scale") OR ("work and health interview") OR ("health and work performance questionnaire") OR ("productivity and disease questionnaire") OR WPAI* OR ("work productivity and activity impairment*") OR ("work productivity short inventory") OR ("wellness inventory") OR ("work role functioning questionnaire") OR ("valuation of lost productivity") OR ("work productivity survey") OR ("work limitations questionnaire") OR ((("migraine disability assessment questionnaire") OR (MIDAS AND migraine) OR Osterhaus) AND (productivity or presenteeism or work* or employ*))

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance

1. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2197–223.
2. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. Arthritis Res Ther. 2009;11(3):229.
3. Burton W, Morrison A, MacLean R, Ruderman E. Systematic review of studies of productivity loss due to rheumatoid arthritis. Occup Med (Lond). 2006;56(1):18–27.
4. Verstappen SM, Bijlsma JW, Verkleij H, Buskens E, Blaauw AA, ter Borg EJ, et al. Overview of work disability in rheumatoid arthritis patients as observed in cross-sectional and longitudinal surveys. Arthritis Rheum. 2004;51(3):488–97.
5. Boonen A, Chorus A, Miedema H, van der Heijde D, van der Tempel H, van der Linden S. Employment, work disability, and work days lost in patients with ankylosing spondylitis: a cross sectional study of Dutch patients. Ann Rheum Dis. 2001;60(4):353–8.
6. Verstappen SM. Rheumatoid arthritis and work: the impact of rheumatoid arthritis on absenteeism and presenteeism. Best Pract Res Clin Rheumatol. 2015;29(3):495–511.
7. ter Wee MM, Lems WF, Usan H, Gulpen A, Boonen A. The effect of biological agents on work participation in rheumatoid arthritis patients: a systematic review. Ann Rheum Dis. 2012;71(2):161–71.
8. Tillett W, de-Vries C, McHugh NJ. Work disability in psoriatic arthritis: a systematic review. Rheumatology (Oxford). 2012;51(2):275–83.
9. Sloman J, Wride A, Garrat D. "Introducing Economics" in Economics. 8th ed. Harlow, New York: Pearson; 2012.
10. Prasad M, Wahlqvist P, Shikiar R, Shih YC. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. Pharmacoeconomics. 2004;22(4):225–44.
11. Mattke S, Balakrishnan A, Bergamo G, Newberry SJ. A review of methods to measure health-related productivity loss. Am J Manag Care. 2007;13(4):211–7.
12. Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: a critical review. Soc Sci Med. 2011;72(2):185–92.
13. Krol M, Brouwer W. How to estimate productivity costs in economic evaluations. Pharmacoeconomics. 2014;32(4):335–44.
14. Escorpizo R, Bombardier C, Boonen A, Hazes JM, Lacaille D, Strand V, et al. Worker productivity outcome measures in arthritis. J Rheumatol. 2007;34(6):1372–80.
15. Brooks A, Hagen SE, Sathyanarayanan S, Schultz AB, Edington DW. Presenteeism: critical issues. J Occup Environ Med. 2010;52(11):1055–67.
16. Verstappen SM, Fautrel B, Dadoun S, Symmons DP, Boonen A. Methodological issues when measuring paid productivity loss in patients with arthritis using biologic therapies: an overview of the literature. Rheumatology (Oxford). 2012;51(2):216–29.
17. Koopmanschap MA, van Ineveld BM. Towards a new approach for estimating indirect costs of disease. Soc Sci Med. 1992;34(9):1005–10.
18. Lensberg BR, Drummond MF, Danchenko N, Despiegel N, Francois C. Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. Clinicoecon Outcome Res. 2013;5:565–73.
19. Pauly MV, Nicholson S, Polsky D, Berger ML, Sharda C. Valuing reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. Health Econ. 2008;17(4):469–85.
20. CRD. Systematic Reviews: CRD's guidance for undertaking reviews in health care. Third ed. University of York; 2009.
21.• Ospina MB, Dennett L, Waye A, Jacobs P, Thompson AH. A systematic review of measurement properties of instruments assessing presenteeism. Am J Manag Care. 2015;21(2):e171–85. **This review focuses on identifying all measures of presenteeism, general or disease-specific, and evaluates the psychometric properties of each presenteeism instrument**.
22. Osterhaus JT, Purcaru O, Richard L. Discriminant validity, responsiveness and reliability of the rheumatoid arthritis-specific Work Productivity Survey (WPS-RA). Arthritis Res Ther. 2009;11(3):R73.
23. Jette AM, Davies AR, Cleary PD, Calkins DR, Rubenstein LV, Fink A, et al. The functional status questionnaire: reliability and validity when used in primary care. J Gen Intern Med. 1986;1(3):143–9.
24. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993;4(5):353–65.
25. van Roijen L, Essink-Bot ML, Koopmanschap MA, Bonsel G, Rutten FF. Labor and health status in economic evaluation of health care. The Health and Labor Questionnaire. Int J Technol Assess Health Care. 1996;12(3):405–15.
26. Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. Psychopharmacol Bull. 1997;33(1):13–6.
27. Kopec JA, Esdaile JM. Occupational role performance in persons with back pain. Disabil Rehabil. 1998;20(10):373–9.
28. Brouwer WB, Koopmanschap MA, Rutten FF. Productivity losses without absence: measurement validation and empirical evidence. Health Policy. 1999;48(1):13–27.
29. Burton WN, Conti DJ, Chen CY, Schultz AB, Edington DW. The role of health risk factors and disease on worker productivity. J Occup Environ Med. 1999;41(10):863–77.
30. Amick III BC, Lerner D, Rogers WH, Rooney T, Katz JN. A review of health-related work outcome measures and their uses, and recommended measures. Spine (Phila Pa 1976). 2000;25(24):3152–60.
31. Lerner D, Amick III BC, Rogers WH, Malspeis S, Bungay K, Cynn D. The work limitations questionnaire. Med Care. 2001;39(1):72–85.
32. Pelletier KR, Koopman C. Standford/American health association presenteeism scale (SAHAPS). In: Lynch W, Riedel J, editors. Measuring employee productivity: a guide to self-assessment tools. Scottsdale, AZ: Institute for Health and Productivity Management 67; 2001.

33. Koopman C, Pelletier KR, Murray JF, Sharda CE, Berger ML, Turpin RS, et al. Stanford presenteeism scale: health status and employee productivity. J Occup Environ Med. 2002;44(1):14–20.
34. Gilworth G, Chamberlain MA, Harvey A, Woodhouse A, Smith J, Smyth MG, et al. Development of a work instability scale for rheumatoid arthritis. Arthritis Rheum. 2003;49(3):349–54.
35. Goetzel RZ, Ozminkowski RJ, Long SR. Development and reliability analysis of the Work Productivity Short Inventory (WPSI) instrument measuring employee health and productivity. J Occup Environ Med. 2003;45(7):743–62.
36. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The world health organization health and work performance questionnaire (HPQ). J Occup Environ Med. 2003;45(2):156–74.
37. Kumar RN, Hass SL, Li JZ, Nickens DJ, Daenzer CL, Wathen LK. Validation of the Health-Related Productivity Questionnaire Diary (HRPQ-D) on a sample of patients with infectious mononucleosis: results from a phase 1 multicenter clinical trial. J Occup Environ Med. 2003;45(8):899–907.
38. Shikiar R, Halpern MT, Rentz AM, Khan ZM. Development of the Health and Work Questionnaire (HWQ): an instrument for assessing workplace productivity in relation to worker health. Work. 2004;22(3):219–29.
39. Stewart WF, Ricci JA, Leotta C, Chee E. Validation of the work and health interview. Pharmacoeconomics. 2004;22(17):1127–40.
40. Turpin RS, Ozminkowski RJ, Sharda CE, Collins JJ, Berger ML, Billotti GM, et al. Reliability and validity of the Stanford Presenteeism Scale. J Occup Environ Med. 2004;46(11):1123–33.
41. Ilmarinen J. The Work Ability Index (WAI). J Occup Environ Med Am Coll Occup Environ Med. 2007;57(2):160.
42. Prochaska JO, Evers KE, Johnson JL, Castle PH, Prochaska JM, Sears LE, et al. The well-being assessment for productivity: a well-being approach to presenteeism. J Occup Environ Med. 2011;53(7):735–42.
43.• Zhang W, Bansback N, Boonen A, Severens JL, Anis AH. Development of a composite questionnaire, the valuation of lost productivity, to value productivity losses: application in rheumatoid arthritis. Value Health. 2012;15(1):46–54. **This study describes the development of the valuation of lost productivity (VOLP) questionnaire. It is the only identified presenteeism instrument that explicitly discussed the economic theory of productivity and how it was used to inform the design of the VOLP.**
44. Boezeman EJ, Sluiter JK, Nieuwenhuijsen K. Measuring work functioning: validity of a weighted composite work functioning approach. J Occup Rehabil. 2015;25(3):537–42.
45. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van RL. The iMTA productivity cost questionnaire: a standardized instrument for measuring and valuing health-related productivity losses. Value Health. 2015;18(6):753–8.
46. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. Expert Rev Pharmacoecon Outcome Res. 2005;5(1):23–8.
47. Noben CY, Evers SM, Nijhuis FJ, de Rijk AE. Quality appraisal of generic self-reported instruments measuring health-related productivity changes: a systematic review. BMC Public Health. 2014;14:115.
48. Tang K, Boonen A, Verstappen SM, Escorpizo R, Luime JJ, Lacaille D, et al. Worker productivity outcome measures: OMERACT filter evidence and agenda for future research. J Rheumatol. 2014;41(1):165–76.

Appendix 3.1

What methodologies have been used to quantify presenteeism? A systematic review protocol

Background

The impact of poor health has been directly linked to lost productivity through absenteeism and presenteeism (Johns 2010). Absenteeism and presenteeism are related but distinct concepts. Absenteeism refers to the time (hours, days, weeks) spent away from work because of illness (Prasad et al. 2004; Mattke et al. 2007; Zhang, Bansback, and Anis 2011; Krol and Brouwer 2014). Presenteeism 'characterises the behaviour of an employee who goes to work even though his health status would justify sick leave' (Huver, Richard, and Vaneecloo 2012) potentially leading to a 'reduced level of productivity whilst at work'(Schultz and Edington 2007; Zhang, Bansback, and Anis 2011; Krol and Brouwer 2014). There is growing evidence that a number of health conditions, both mental and physical, can negatively affect productivity (Goetzel et al. 2004). The current evidence base has focussed on theoretical and empirical research to quantify absenteeism, but far less has been done to understand if, and how, presenteeism can be identified, either as a cost or as a (dis-)benefit (Heuvel et al. 2009). There is some evidence that presenteeism accounts for the majority of the total cost associated with health conditions exceeding both medical costs and the costs associated with absenteeism (Tillett, de-Vries, and McHugh 2012; Johns 2010; Pauly et al. 2008; Collins et al. 2005). There is ongoing debate on if, and how, to capture the impact and quantify presenteeism in a way that the can be used in empirical studies.

Numerous systematic reviews have been published on the general topic of quantifying the impact of presenteeism (Mattke et al. 2007; Schultz, Chen, and Edington 2009; Brooks et al. 2010; Johns 2010; Zhang, Bansback, and Anis 2011; Her and Kavanaugh 2012; Krol, Brouwer, and Rutten 2013). The reviews discuss the limitations and challenges of measuring and subsequently valuing productive time lost associated with presenteeism.

Measuring and Valuing Presenteeism

The challenge of measuring presenteeism stems from the lack of available objective data that can be used to measure productivity. In a study by Burton et al. (2001), the number of telephone calls made by a sample of customer service representatives with

allergic disorders was used as an objective measure of presenteeism; this, however, is a rare example of having access to am objective measure of presenteeism (Schultz, Chen, and Edington 2009). Objective data are largely unavailable across the majority of jobs and industries within the United Kingdon's economy. The limited availability of objective data has led to the development of numerous self-reported questionnaires designed to subjectively measure the volume of lost productivity. Many of the self-reported questionnaires are limited in terms of their validity and reliability and produce estimates of presenteeism that vary substantially (Johns 2010). The wide variation of the impact of presenteeism estimated is largely due to the various methods used by the self-reported questionnaire (Mattke et al. 2007; Brooks et al. 2010; Zhang, Bansback, and Anis 2011; Krol, Brouwer, and Rutten 2013).

Mattke et al. (2007) found several inconsistencies across self-reported instruments of presenteeism. The first concerns how respondents were asked to estimate their level of lost productivity when ill at work. The most popular approach asked respondents to assess how much their illness affects them whilst at work. Two other distinct approaches included asking the respondent to make a comparison of their work performance to themselves when in a good state of health or with other employees that work in a similar job role. The length of recall period used was also identified to differ across self-report instruments of presenteeism. Recall periods are the amount of time a respondent is asked to think back on and provide answers based on that time. The length of the recall periods needs to be long enough to capture disease fluctuations, but also short enough to promote accurate responses. Mattke et al. (2007) found recall periods vary in length from one week to twelve months, with two weeks being the most commonly used.; however, concensus concerning the optimal length for a recall period is still not clearly defined.

The methods used to estimate the cost of presenteeism were discussed extensively in systematic reviews (Mattke et al. 2007; Schultz, Chen, and Edington 2009; Brooks et al. 2010; Johns 2010; Zhang, Bansback, and Anis 2011; Her and Kavanaugh 2012; Krol, Brouwer, and Rutten 2013). The two most commonly used methods included the human capital approach (HCA) and the friction cost approach (FCA). The HCA adopts a patient perspective and estimates the cost of presenteeism by multiplying unproductive time by the wage. The advantage of this method is that it is simple to apply; however, it has also been criticised for potentially over-estimating the cost of presenteeism. The FCA was developed to calculate the cost of absenteeism by totalling the hours not

worked until a replacement employee was hired, afterwhich production levels prior to absence are fully restored (Koopmanschap et al. 1995). The FCA is difficult to justify as an appropriate method for the valuation of presenteeism because there is no need to replace the sick worker; the sick worker is still present at work.

Contextual factors are another aspect that differs across presenteeism instruments. Contextual factors describe those aspects of work that may influence the impact of lost productivity. Pauly et al. (2002) suggest that the lost productivity caused by one employee may cause further negative productivity impacts elsewhere in a firm. Pauly et al. (2002) proposed three criteria that need to be taken into account when attempting to capture the interdependent effects of lost productivity: 1) the degree to which the employee is replaceable; 2) the extent to which the employee works as part of a team; and 3) the time sensitivity of the output produced by the employee. Zhang, Bansback, and Anis (2011) also discussed the criteria proposed by Pauly and colleagues (2002) and suggested that this research could be used to capture the differences in characteristics between jobs and alternative workplaces. Zhang et al., also highlighted the use of compensation mechanisms. Compensation mechanism are those that reduce the amount of lost productivity (Jacob-Tacken et al. 2005; Nicholson et al. 2006), for example a team member may be able to pick up extra work to help the ill team member at no extra cost in terms of having to work overtime in order to complete the work.

Perspectives

The study perspective affects the types of resource use and costs that can be included in an economic evaluation. However, whilst economic theory advocates the adoption of a societal perspective, other health systems in some countries, including the United Kingdom and Belgium, adopt a healthcare sector perspective. Adopting a healthcare perspective means only those costs and benefits that fall on the healthcare sector should be considered in an economic evaluation; therefore, the cost of lost productivity should not be included.

Theory

Economic theory that underpins the concept of absenteeism is well developed; however, there lack an economic theory for presenteeism (Johns 2010). The review by Zhang et al. (2011) reports the limitations of the economic theories used to underpin the methods to value the cost of lost productivity. The HCA and FCA assume that productivity is equal to the market wage which represents the marginal revenue product of labour of a

perfectly competitive firm. Zhang et al. (2011) ignore the complexities of the imperfect market relating to discriminative wages and instead focus on the theoretical limitations highlighted by Pauly et al. (2002). Pauly et al. (2002) suggest that the allowances for sick days and protection against fluctuating wages will mean that employees will accept a wage rate that is lower than the value of the marginal productivity of the worker. Therefore, the value of lost productivity will exceed the value of the wage because the wage is lower than marginal productivity. Zhang et al. (2011) also conclude that the value of lost productivity will exceed the value of the wage rate if a job type involved team based work, unavailability of substitutes, and produces highly time sensitive output.

Future Research Areas

There is consensus in published systematic reviews that there is currently no 'gold standard' method for identifying, measuring and valuing the cost of presenteeism (Mattke et al. 2007; Schultz, Chen, and Edington 2009; Brooks et al. 2010; Johns 2010; Zhang, Bansback, and Anis 2011; Her and Kavanaugh 2012; Krol, Brouwer, and Rutten 2013). The reviews largely conclude that future research needs to focus on the development and validation of the methodologies used to measure and value presenteeism as a cost. Contextual factors ranging from organisational features, personality differences, financial incentives (sick pay), to team dependence and time sensitivity of output were reported by the majority of reviews as an area that requires further consideration. Two reviews also concluded that more research is needed to take into account the effects of lost productivity in the context of unpaid work (Krol, Brouwer, and Rutten 2013; Zhang, Bansback, and Anis 2011).

However, there are other areas of further research that have not been highlighted by the reviews presented here. Firstly, there is a clear lack of economic theory used to underpin the concept of presenteeism. Zhang et al. (2011) described the relevant labour economic theories used to underpin productivity and then discussed the economic model proposed by Pauly et al. (2002) that considers theories of absenteeism, which are then extended and applied to presenteeism in Pauly et al. (2008). The economic theories reported by Zhang et al. (2011) apply to the cost conversion of productive time in terms of the relevant wage that could be used in when converting lost productive time into a cost unit, relevant for both absenteeism and presenteeism estimates. However, there is no comment made about the economic theory underpinning the concept of presenteeism in terms of lost productive time whilst at work.

The second area that is not mentioned by the reviews is the alternative ways lost productivity can be valued and entered into an economic evaluation. The reviews report lost productivity as a cost that would affect the numerator in an incremental cost-effectiveness ratio (ICER); however it is also possible to treat lost productivity as a non-monetary outcome that can be included in the denominator of the ICER (Gold et al. 1996).

The aim of this systematic review was to examine the extent to which methods of presenteeism are underpinned by economic theory.

Research Aims and Objectives

The primary aim of this systematic review is to assess the extent to which methods used to quantify presenteeism are based on economic theory. As a secondary aim the review will also report key gaps in empirical research relating to the: 1) quantification of the volume of presenteeism; 2) identification of the relevant unit costs; and 3) if, and how, methods are used to value presenteeism as a non-monetary outcome. Information from the identified studies will be used to address four objectives:

1. To identify economic theories described in the development of existing measures of presenteeism
2. To identify and describe how presenteeism identified
3. To identify the relevant unit costs used to value presenteeism
4. To identify and describe whether, and if so how, presenteeism has been valued as a non-monteary benefit

Methods

The systematic review will involve three stages in order to identify all relevant studies of the development of self-reported presenteeism instruments.

Search strategy: Stage one

A search strategy consisting of relevant index and free-text terms will be run the electronic database PubMed to identify all existing systematic reviews regarding the methods used to quantify presenteeism.

Search strategy: Stage two

The most appropriate systematic review will be selected using pre-defined criteria. The search strategy from the selected systematic review will be updated and ran in appropriate electronic databases.

Search strategy: Stage three

References of the included studies will also be screened in order to identify further relevant studies.

Inclusion Criteria

There are two sets of different inclusion criteria that will need to be applied in stage one and stage two of the search. The inclusion criteria that will be applied to each stage of the search are described in the following section.

Stage one

The inclusion criteria is designed to identify those systematic reviews that focus on the methods used to quantify lost productivity associated with any mental or physical condition. For the purposes of this review, a systematic review is defined as a review that designed and applied a systematic search strategy to identify relevant studies. The systematic reviews must specify that they are examining the methods used to measure or quantify lost productivityassocaited with presenteeism. The search for relevant systematic reviews will be restricted those published inclusive of and after January 2000; this was decided because of the lack of presenteeism studies published before this time (Aronsson, Gustafsson, and Dallner 2000). Systematic reviews that are published in English or have a readily available translation in English will be included in the review. Table one summarises the inclusion criteria that will be applied in stage one.

Table 1 Inclusion Criteria for stage one

Inclusion Criteria
Systematic review
Health-related Presenteeism
Valuation of productivity loss at work
Published inclusive of and after January 2000
English or English translation

Stage Two

The studies that will be considered relevant for inclusion are those that focus on developing methods that measure lost productivity associated with presenteeism. The studies that propose methods to quantify presenteeism must do so in the context of mental or physical disease areas. The studies must be published in English or have a readily available translation into English. Any non-English studies that are relevant but do not have an English translation available will be recorded as such to minimise bias.

Table two summarises the inclusion criteria that will be applied in stage two. Empirical studies that measure presenteeism will not be included in this review.

Table 2 Inclusion Criteria

Inclusion Criteria
Methods that quantify presenteeism study
Health-Related Presenteeism
English or English Translation

Study Selection

The studies identified by the search strategies will be managed and de-duplicated using Reference Manager 12. The study selection process will be conducted in two-phases. First, the titles and abstracts of each identified systematic review will be double screened by 2 reviewers (CJ and SV). The reviewers should ensure that all of the inclusion criteria are met before selecting the systematic review or study. The second stage will involve double-screening the full-text studies (CJ and SV). Any discrepancies regarding the inclusion of reviews or studies identified at stage one or two will be resolved through discussion with a third reviewer (KP).

Data extraction

The data will be extracted by one reviewer (CJ) using tailored data collection forms. The extracted data will be verified by a second reviewer (KP). Information extracted from the studies will include the following information;

- Author, year of publication
- Aims and objectives
- Population type, for example rheumatoid arthritis patients
- Perspective and time horizon adopted
- Economic theories used to underpin measures of presenteeism
- Instruments and tools used to identify and value presenteeism
- Cost information including methodologies used to convert presenteeism into a cost, the methodological issues regarding cost conversion, and the cost units used
- Valuation of the dis-benefits of presenteeism

Data Synthesis

The results from the data extraction process, from stages one and two, will be summarised as part of a narrative synthesis. The results found from the sysemtatic review will be used to highlight key gaps of the methods used to quantify presenteeism. The results collected through stage two of the review will discuss the extent to which economic theories are used to underpin methods used to quantify presenteeism.

Dissemination

The final systematic review will be put for consideration of publication in a peer-reviewed health economics based journal.

References

Aronsson, G., Gustafsson and M. Dallner. 2000. 'Sick but yet at Work. An Empirical Study of Sickness Presenteeism'. *Journal of Epidemiology and Community Health* 54 (7):502–9. https://doi.org/10.1136/jech.54.7.502.

Brooks, A., S. E. Hagen, S. Sathyanarayanan, A. B. Schultz, and D. W. Edington 2010. 'Presenteeism: Critical Issues'. *Journal of Occupational and Environmental Medicine* 52 (11):1055–67. https://doi.org/10.1097/JOM.0b013e3181f475cc.

Burton, W. N., D. J. Conti, C.Y. Chen, A. B. Schultz, and D.W. Edington 2001. 'The Impact of Allergies and Allergy Treatment on Worker Productivity'. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine* 43 (1):64–71.

Collins, J. J., C. M. Baase, C. E. Sharda, R. J. Ozminkowski, S. Nicholson, G. M. Billotti, R. S. Turpin, M. Olson, and M. L. Berger. 2005. 'The Assessment of Chronic Health Conditions on Work Performance, Absence, and Total Economic Impact for Employers': *Journal of Occupational and Environmental Medicine* 47 (6):547–57. https://doi.org/10.1097/01.jom.0000166864.58664.29.

Goetzel, R. Z., S. R. Long, R. J. Ozminkowski, K. Hawkins, S. Wang, and W. Lynch. 2004. 'Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers'. *Journal of Occupational* 46 (4):398–412.

Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein. 1996. *Cost-Effectiveness in Health and Medicine*. New York: OUP USA.

Her, M., and A. Kavanaugh. 2012. 'Critical Analysis of Economic Tools and Economic Measurement Applied to Rheumatoid Arthritis'. *Clinical and Experimental Rheumatology* 30 (4):S107–11.

Heuvel, S. G., G. A. van den Geuskens. W. E. Hooftman, L. L. J. Koppes, and S. N. J. van den Bossche. 2009. 'Productivity Loss at Work; Health-Related and Work-Related Factors'. *Journal of Occupational Rehabilitation* 20 (3):331–39. https://doi.org/10.1007/s10926-009-9219-7.

Huver, B., S. Richard, and N. Vaneecloo. 2012. 'Sick but at Work. An Econometric Approach to Presenteeism.' *15th IZA European Summer School in Labor Economics*.

Jacob-Tacken, K., H. M., M. A. Koopmanschap, W. J., Meerding, and J. L. Severens, 2005. 'Correcting for Compensating Mechanisms Related to Productivity Costs in Economic Evaluations of Health Care Programmes'. *Health Economics* 14 (5):435–43. https://doi.org/10.1002/hec.948.

Johns, G. 2010. 'Presenteeism in the Workplace: A Review and Research Agenda'. *Journal of Organizational Behavior* 31 (4):519–42. https://doi.org/10.1002/job.630.

Koopmanschap, M. A., F. F. Rutten, B. M. van Ineveld, and L. van Roijen. 1995. 'The Friction Cost Method for Measuring Indirect Costs of Disease'. *Journal of Health Economics* 14 (2):171–89.

Krol, M., and W. Brouwer. 2014. 'How to Estimate Productivity Costs in Economic Evaluations'. *PharmacoEconomics* 32 (4):335–44. https://doi.org/10.1007/s40273-014-0132-3.

Krol, M., W. Brouwer, and F. Rutten. 2013. 'Productivity Costs in Economic Evaluations: Past, Present, Future'. *PharmacoEconomics* 31 (7):537–49. https://doi.org/10.1007/s40273-013-0056-3.

Mattke, S., A. Balakrishnan, G.Bergamo, and S. J. Newberry. 2007. 'A Review of Methods to Measure Health-Related Productivy Loss'. *American Journal of Managed Care* 13 (4):211–17.

Nicholson, S., M. V. Pauly, D. Polsky, C. Sharda, H. Szrek, and M. L. Berger. 2006. 'Measuring the Effects of Work Loss on Productivity with Team Production'. *Health Economics* 15 (2):111–23. https://doi.org/10.1002/hec.1052.

Pauly, M. V., S. Nicholson, D. Polsky, M. L. Berger, and C.Sharda. 2008. 'Valuing Reductions in on-the-Job Illness: "presenteeism" from Managerial and Economic Perspectives'. *Health Economics* 17 (4):469–85. https://doi.org/10.1002/hec.1266.

Prasad, M., P. Wahlqvist, R. Shikiar, and Y. C. T. Shih. 2004. 'A Review of Self-Report Instruments Measuring Health-Related Work Productivity - A Patient-Reported Outcomes Perspective'. *Pharmacoeconomics* 22 (4):225–44. https://doi.org/10.2165/00019053-200422040-00002.

Schultz, A. B., C-Y. Chen, and D. W. Edington. 2009. 'The Cost and Impact of Health Conditions on Presenteeism to Employers: A Review of the Literature'. *PharmacoEconomics* 27 (5):365–78.

Schultz, A. B., and D. W Edington. 2007. 'Employee Health and Presenteeism: A Systematic Review'. *Journal of Occupational Rehabilitation* 17 (3):547–79. https://doi.org/10.1007/s10926-007-9096-x.

Tillett, W., C. de-Vries, and N. J. McHugh. 2012. 'Work Disability in Psoriatic Arthritis: A Systematic Review'. *Rheumatology* 51 (2):275–83. https://doi.org/10.1093/rheumatology/ker216.

Zhang, W., N. Bansback, and A. H. Anis. 2011. 'Measuring and Valuing Productivity Loss due to Poor Health: A Critical Review'. *Social Science & Medicine* 72 (2):185–92. https://doi.org/10.1016/j.socscimed.2010.10.026.

Appendix 3.2

Search Strategy for Systematic Reviews of Methods used to Quantify Presenteeism

1. (Systematic adj review).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
2. systematic.mp.
3. "Peer Review"/ or review.mp.
4. ((change or loss) adj3 product$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
5. (health-related adj5 producti$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
6. (self-report$ adj instrument$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
7. (health-related adj 5 work-functio$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
8. presenteeism.mp.
9. (method$ adj5 presenteeism).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
10. (method$ adj5 producti$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
11. (measur$ adj5 presenteeism).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
12. (measur$ adj5 producti$).mp. [mp=hw, ab, ti, ct, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, ui]
13. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12

Appendix 3.3

PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	N/A
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	N/A
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	

FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

Appendix 3.4

Electronic Search Strategies

OVID Databases:
Medline (1946 to September week 3 2015),
Embase (1980 to week 40 2015),
CENTRAL (August 2015)
PsycINFO (1806 to September week 4 2015)

1. American Productivity Audit.tw.
2. "angina-related limitations at work question?aire".tw.
3. endicott work productivity scale.tw.
4. "health and labo?r question?aire".tw.
5. "health and performance question?aire".tw.
6. "health and productivity question?aire".tw.
7. "health and work question?aire".tw.
8. health-related productivity question?aire*.tw.
9. "lam employment absence and productivity scale".tw.
10. (migraine disability assessment adj2 (question?aire or survey or scale or score*)).tw.
11. MIDAS.tw. and migraine*.mp.
12. 10 or 11
13. (productiv* or presenteeism or absenteeism or work* or employ*).mp.
14. 12 and 13
15. "migraine work and productivity loss question?aire".tw.
16. (osterhaus and (work* or productivity or presenteeism)).tw.
17. (osterhaus adj3 technique).tw.
18. "productivity and disease question?aire".tw.
19. PRODISQ.tw.
20. (quantity adj2 quality adj (method or instrument)).tw.
21. (Stanford* adj5 Presenteeism Scale).tw.
22. "work and health interview".tw.
23. work performance scale.tw.
24. "work productivity and activity impairment*".tw.
25. WPAI*.tw.
26. (US National Health and Wellness Survey).tw. and (productiv* or presenteeism or absenteeism or work*).mp.
27. "health and work performance question?aire".tw.
28. work productivity short inventory.tw.
29. wellness inventory.tw.
30. work role functioning question?aire.tw.
31. (or/1-9) or (or/14-30)
32. limit 31 to english language
33. (work productivity survey or WPS-RA).tw.
34. valuation of lost productivity.tw.
35. work limitations question?aire.tw.
36. (33 or 34) not 32
37. limit 36 to english language
38. 35 not 32
39. limit 38 to english language
40. 32 or 37

41. 32 or 37 or 39

ISI platform databases:
Web of Science (1900 - Nov 6, 2015)

1 TS="angina-related limitations at work questionnaire"
2 ((TS="endicott work productivity scale")) AND Language=(English)
3 (TS=("health and labor questionnaire" OR "health and labour questionnaire")) AND Language=(English)
4 ((TS="health and productivity questionnaire")) AND Language=(English)
5 ((TS="health and work questionnaire")) AND Language=(English)
6 ((TS="health-related productivity questionnaire")) AND Language=(English)
7 ((TS="lam employment absence and productivity scale")) AND Language=(English)
8 ((TS="migraine work and productivity loss questionnaire")) AND Language=(English)
9 ((TS="Stanford Presenteeism Scale")) AND Language=(English)
10 ((TS="health and work performance questionnaire")) AND Language=(English)
11 ((TS="work and health interview")) AND Language=(English)
12 ((TS="work performance scale")) AND Language=(English)
13 ((TS="work productivity short inventory")) AND Language=(English)
14 ((TS="American Productivity Audit")) AND Language=(English)
15 ((TS=(osterhaus and (work or productivity)))) AND Language=(English)
16 TS=("work productivity and activity impairment*")
17 TS=WPAI*
18 TS="wellness inventory"
19 TS=("work productivity survey")
20 TS=("Work role functioning*" SAME limitations)
21 TS=(osterhaus SAME technique)
22 TS=("productivity and disease questionnaire")
23 (TS=(migraine disability assessment SAME (score* OR scale OR questionnaire OR survey))) AND Language=(English)
24 (TS=(MIDAS AND migraine)) AND Language=(English)
25 (TS=Osterhaus) AND Language=(English)
26 (#23 OR #24 OR #25) AND Language=(English)
27 (TS = (work* or productivity or performance or presenteeism or absenteeism or employ*)) AND Language=(English)
28 (#26 AND #27) AND Language=(English)
29 (TS=("work role functioning questionnaire")) AND Language=(English)
30 (TS="health and performance questionnaire") AND Language=(English)
31 TS=("valuation of lost productivity")
32 TS=("work limitations questionnaire")
33 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #28 OR #29 OR #30 OR #31 OR #32) AND Language=(English)

EBSCO Databases:

CINAHL (1937 – Nov 6, 2015)
Business Source Complete (1886 – Nov 6, 2015)

Search ID#	Search Terms
S28	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 OR S24 OR S25
S25	"work limitations questionnaire"
S24	"valuation of lost productivity"
S23	"work productivity survey"
S22	"health and performance questionnaire"
S21	"productivity and disease questionnaire"
S20	osterhaus AND (technique or presenteeism or absenteeism or productivity or work* or employ*)
S19	"Work role functioning questionnaire"
S18	"wellness inventory"
S17	WPAI*
S16	"work productivity and activity impairment*"
S15	"American Productivity Audit"
S14	"work productivity short inventory"
S13	"work performance scale"
S12	"work and health interview"
S11	"health and work performance questionnaire"
S10	"Stanford Presenteeism Scale"
S9	"migraine work and productivity loss questionnaire"
S8	"lam employment absence and productivity scale"
S7	(("migraine disability assessment" n2 (score* or scale or survey or questionnaire)) OR (MIDAS AND migraine)) AND (work* OR productivity OR presenteeism OR absenteeism OR employ*)
S6	"health-related productivity questionnaire"
S5	"health and work questionnaire"
S4	"health and productivity questionnaire"
S3	("health and labor questionnaire") OR ("health and labour questionnaire")
S2	"endicott work productivity scale"
S1	"angina-related limitations at work questionnaire"

Proquest databases:
ABI Inform (1970 - Nov 6, 2015)
ALL(("American Productivity Audit") OR ("angina-related limitations at work questionnaire") OR ("endicott work productivity scale") OR ("health and labor questionnaire") OR ("health and labour questionnaire") OR ("health and productivity questionnaire") OR ("health and work questionnaire") OR ("health and performance questionnaire") OR ("health-related productivity questionnaire") OR ("lam employment absence and productivity scale") OR ("migraine work and productivity loss questionnaire") OR ("Stanford Presenteeism Scale") OR ("work performance scale") OR ("work and health interview") OR ("health and work performance questionnaire") OR ("productivity and disease questionnaire") OR WPAI* OR ("work productivity and activity impairment*") OR ("work productivity short inventory") OR ("wellness inventory") OR ("work role functioning questionnaire") OR ("valuation of lost productivity") OR ("work productivity survey") OR ("work limitations questionnaire") OR ((("migraine disability assessment questionnaire") OR (MIDAS AND migraine) OR Osterhaus) AND (productivity or presenteeism or work* or employ*))

Appendix 3.5

PRISMA checklist item results for the selection of presenteeism measures for the prediction model

Study	**PRISMA Checklist Item Results**																	**Total**
	1	**2**	**3**	**4**	**6**	**7**	**8**	**9**	**10**	**11**	**14**	**17**	**18**	**24**	**25**	**26**	**27**	/17
Abma et al 2012		X	X	X	X	/	X	X	X	/	X	X	X	/	/	X		13
Brooks et al 2010		X	X			/							X	X	X	X		6.5
Brown et al 2014		X	X		X	X		/	X	/		X	X	X	/	X		10.5
Despiegel et al 2012		X	X		X	/		X	X	X		/	X	X	X	X		11
Filipovic et al 2011		X	X	X		X		X				/	X	X	X	X		9
Loeppke et al 2003		/	X	X		X			X	X			X	X		/		7.5
Lofland et al 2004		X	X		/	X	/			X			X	X	X	X	X	10
Mattke et al 2007		X	X		/	/							X	X	X	X	X	8
Nieuwenhuijsen et al 2010		X	X	X	X	/	X	X		X	X	/	X	X	X	X		13
Noben et al 2014	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		15
Ospina et al 2015	X	X	X		X	X		X	X	X	X	X	X	X	X	X		14
Prasad et al 2004		X	X		X	X	/			X			X	X	X	X	X	10.5
Roy et al 2011	X	X	X			X	/		X		X	X	X	X	X	X		11.5
Tang et al 2015		X	X	X	X	X	/						X	X	X	X		9.5
Uegaki et al 2011	X	X	X		X	X	X	X	X		/	X	X	X	X	X		13.5
Verstappen et al 2012		X	X		X	X	/	X	X	X			X	X	X	X		11.5
Williams et al 2007	X	X	X	X	X	X		X	X		X	/	X	X	X	X		13.5

Appendix 4.1

To what extent is productivity included in Economic Evaluations of Workplace interventions that manage musculoskeletal disorders, and how? A systematic review protocol

Background

The impact of poor health has been directly linked to productivity loss through absenteeism and presenteeism (Johns 2010). Absenteeism and presenteeism are distinct but related concepts. Absenteeism, refers to the amount of time absent from work directly attributed to ill health (Soeren Mattke et al. 2007; Zhang, Bansback, and Anis 2011; Krol, Brouwer, and Rutten 2013). Presenteeism, as defined by many in the literature, describes the choice made by an employee to attend work whilst being ill enough to justify sick leave (Huver, Richard, and Vaneecloo 2012) and therefore leading to an overall reduction in work performance (productivity) (Schultz and Edington 2007; Krol, Brouwer, and Rutten 2013).

In 2005, 38% of all occupational diseases were reported as musculoskeletal disorders (MSDs) in twelve European Union member states (Eumusc, 2013). MSDs describe a set of inflammatory and degenerative conditions that cause pain in the muscles, tendons, nerves, and joints, affecting the body's ability to function (Punnett and Wegman 2004). MSDs may present in the neck, back, upper and lower limbs. MSDs can affect people across all ages of the population from young children through to older adults (Burton and Kendall 2014). Disability caused by MSDs can affect an individual's ability to work and maintain employment (Burton and Kendall 2014). In the UK, 48% of all new cases of work-related illness were attributable to MSDs in 2013/14 making MSDs the most common cause of work-related illness (HSE, 2014). The total cost of MSDs to society, including direct medical costs of treatment and indirect costs associated with lost productivity, was estimated at over £7 billion a year (2007 prices) (HSE, 1999; Bevan, Passmore, and Mahdon 2007). One of the greatest contributing factors to the total cost of MSDs was the cost incurred through losses in productivity (Bevan, Passmore, and Mahdon 2007).

The potential impact health problems may have on productivity can be complex. Lost productivity caused by a health disorder may negatively affect the quality and quantity of output produced by the individual (Zhang, Bansback, and Anis 2011), and overall

team production levels (Nicholson et al. 2006). Workplace interventions (WPIs) are strategies or programmes that are designed to help people improve symptoms and functional ability (Public Health England 2014). WPIs may also help those with MSDs maintain employment (Hoving et al. 2014). Many WPIs are largely directed towards the individual patient, the patient's workplace, and healthcare or other services for which the patient has access (Palmer et al. 2012). Table1summarises the different types of WPIs.

Table 1 Types of non-pharmacological workplace interventions (WPIs)

Intervention level	Non-pharmacological Interventions	Description
Individual	Exercise or physical therapy	To increase flexibility and strength of the muscles
	Psychological therapy	To change the behaviour and attitude of the patient. To help patients cope.
	Vocationally focussed with education	Help patients overcome social barriers including attitudes and perceptions of work
	Rehearsal of safer working techniques	To avoid unnecessary strain on the body
Workplace	Ergonomic or psychosocial risk assessments aimed at the individual	To identify and control risks within the workplace
	Ergonomic changes to the physical environment	Equipment changes, such as computer displays. Environment changes, such as temperature and lighting.
	Job modifications	Flexible working hours and lighter duties to help prevent further damage
Service	Assessment and coordinated action plan	Developed by a multidisciplinary team or case manager
	Consultation with occupational physician	-
	Education of primary-care doctors and/or occupational physicians and/or agreements between them	To improve communication between work and primary care and/or occupational physicians
	Access to external support and referral services	Further support available outside of the workplace

Source: Palmer et al. (2012), HSE. (2013)

The wide variety of WPIs that have been developed has led to the need for an assessment of their cost-effectiveness. A cost-effectiveness analysis can be done through the use of economic evaluation. Economic evaluation is used to compare the costs and outcomes of two or more interventions (Drummond et al. 2015). The inclusion

of productivity as a cost in economic evaluations is determined by the perspective adopted. For example, a societal perspective includes all costs, including those associated with productivity, to be included in a cost-effectiveness analysis. Similarly, an employer's perspective would also allow for the inclusion of productivity as a cost. In contrast, as practiced in the United Kingdon, a healthcare perspective includes only those costs that are applicable to the healthcare system. Therefore, a healthcare perspective does not allow for the inclusion of productivity as a cost in its analyses.

In 2007, the UK's Department of Health commissioned the National Institute for Health and Care Excellence (NICE) (Hillage et al. 2008) to systematically review the literature that examined the cost-effectiveness of WPIs that either prevented or reduced long-term sickness absence, or helped people to return to work after a period of long-term sickness absence. The review found eleven cost-effectiveness studies ofWPIs, 10 of which focussed on interventions for MSDs and one for mental health. NICE's 2008 review reported the cost-effectiveness of WPIs of each included study and found that most interventions were designed to aid return to work.

In a later review the aim was to investigate the methods used to estimate the indirect-costs of absenteeism and presenteeism in economic evaluations from the perspective of the employer (Uegaki et al. 2010). The authors included economic evaluations that performed a cost-effectiveness analysis (CEA), cost-utility analysis (CUA), cost-benefit analysis (CBA), and cost-analysis as defined by Drummond et al (2015). The authors also included studies that compared the costs and consequences of two or more interventions in which the monetary outcomes were limited to changes in healthcare use and/or productivity. These studies were labelled as "financial appraisals" in order to make a distinction between cost-analysis studies where health-related productivity was considered as a cost. The authors identified 34 studies of which 17 focussed on MSDs, seven on influenza, two on mental health, two on migraines, and one on general health. Estimates of lost productivity associated with absenteeism were found in all 34 studies; however, presenteeism estimates were found in only 6 studies. Uegaki et al. (2011) concluded that there was wide variation in the methods used to identify, measure, and value lost productivity, a result that is consistent with many systematic reviews of methods used to quantify the impact of lost productivity.

The analysis of the literature made by NICE (2008) and Uegaki et al's reviews focus on determining the cost-effectiveness of WPIs and the methods used to estimate health-

related productivity loss respectively. At present, there is little information that has been consolidated to examine how lost productivity estimates were included in economic evaluations of WPIs. The proposed systematic review will examine the extent to which lost productivity estimates are included in economic evaluations of WPIs for MSDs and how these estimations of lost productivity have been incorporated and presented in economic analyses. This systematic review will evaluate all economic evaluations of WPIs for MSDs that have been conducted from the perspective of the patient, the employer, the healthcare system, and society.

Review Question

The aim of this systematic review is to assess the extent to which, and how, the impacts on productivity have been identified and quantified in economic evaluations of WPIs aimed at managing MSDs. The review will also aim to describe how the estimates of lost productivity have been incorporated into economic analyses. The evidence identified will be used to address the following research objectives;

1. How many economic evaluations of WPIs designed to manage MSDs have included estimates of lost productivity in their analyses?
2. In the studies that have included estimates of lost productivity, what types of health-related lost productivity, absenteeism or presenteeism, have been included in the analysis?
3. What methods have been used to measure and value the costs and/or dis-benefits of productivity loss in workplace intervention studies?
4. How have the estimates of lost productivity, as a cost or benefit, been incorporated into the analysis of workplace interventions for MSDs?

Search Strategy

The search for economic evaluations of WPIs for MSDs will comprise three stages. Stage one will involve updating the electronic search ran by Hillage et al. (2008). Stage two will involve collecting both included and excluded economic evaluations identified by Hillage et al. (2008) review. The third stage will involve sifting through references of relevant economic evaluations, identified in stages one and two to identify any further potential studies.

Electronic Search

To identify relevant studies for this review, an update of the search conducted by NICE (Hillage et al. 2008) will be undertaken. The search strategy used by Hillage et al., (2008) will be adapted to identify economic evaluations of WPIs for MSDs. The search strategy will include search terms used by Hillage et al., combined with the relevant economic evaluation search filters published by NHS EED. Free text and index terms related to presenteeism and reduced work performance will also be included in the final search strategy. An example of the search strategy used in Medline can be found in appendix seven.

The databases that will be used to search for economic evaluations of workplace interventions for MSDs will include;

1. Medline
2. Embase
3. Econlit

Hillage et al., (2008) references

The economic evaluations that were included and excluded in the Hillage et al. (2008) systematic review will be assembled for screening.

Alternative Sources

The references of each relevant economic evaluation identified by the electronic search and the search conducted by Hillage et al (2008) will be screened for possible extra relevant studies.

Inclusion and Exclusion Criteria

The inclusion criteria will be applied in order to identify relevant economic evaluations of workplace interventions for those with MSDs is described in this section using PICOS.

Musculoskeletal Disorders

MSD defines a wide range of musculoskeletal conditions from back pain to rheumatoid arthritis. For the purposes of this review, MSDs that will be included are rheumatic and musculoskeletal disorders (RMDs) as defined by (Al Maini et al. 2015). Table two lists the RMDs that will be included. Musculoskeletal injuries as a result of work will not be included in this study.

Table 2 Rheumatic Musculoskeletal Disorders (RMDs) included in this systematic review

Rheumatic Musculoskeletal Disorders	**Other Rheumatic Musculoskeletal Disorders**
Low back pain	Ankylosing spondylitis
Neck pain	Arthralgia
Osteoporosis	Behcet's disease
Osteoarthritis (Hip or knee)	Carpal tunnel tyndrome
Rheumatoid Arthritis	Epicondylitis
Gout	Fibromyalgia
Other RMDs	Myofascial pain
	Psoriatic Arthritis
	Scoliosis
	Seronegative rheumatism
	Soft tissue rheumatism
	Systemic lupus erythematosus
	Tendinitis
	+ many other unclassified RMD conditions

Source: Al Maini et al. 2015

Participants

The study participants that will be included in this review are adults (16 years plus) with MSDs who are in full-time or part-time employment.

Intervention

The interventions that are relevant for inclusion into this review are WPIs that aim to help individuals with MSDs maintain employment, prevent sickness absence, and reduce presenteeism. WPIs that are relevant to this review are non-pharmacological programmes and strategies that are delivered by health care professionals, employers, or both.

Outcomes

The outcomes refer to the benefits associated with implementing aWPI. These outcomes may include:

- Reduced presenteeism
- Length of sickness absence
- Recurring sickness absence
- Number of people who have returned to work
- Number of sick days
- Prevention of permanent work disability

Study Design

Economic evaluations, as defined by Drummond et al (2015), that compare the costs and outcomes of two or more interventions will be included in this review. CBA, CUA, and CEA will be accepted as appropriate methods used for economic evaluation. Cost-analyses and cost-minimisation analyses will not be included in this systematic review.

Language and Date restrictions

The systematic review by Hillage et al. (2008) identified studies published between 1990 and 2007; therefore, the search strategy that will be used for the purposes of this systematic review will be run from the 1st January 2007 to the present day. Studies that are published in English or have a readily available translation in English will be included in the review. Table three presents a concise list of the inclusion criteria that will be used for this review.

Table 3 Inclusion Criteria

Inclusion Criteria
Musculoskeletal Conditions (MSDs)
Non-pharmacological workplace interventions
Economic Evaluations
Adults aged 16 years plus in full/part-time work
English or English translation
Published inclusive of and after 2007 to present date

Study Selection

The study selection process will involve double screening of all the studies identified by the electronic search, the Hillage et al. (2008) review and references of relevant studies. The selection process will be conducted in two-stages. First, each of the identified studies will have their titles and abstracts screened by two reviewers (CJ and SV). Second, a full-text version of the selected studies will be obtained and screened by two reviewers (CJ amd KP). Any discrepancies concerning the inclusion or exclusion of studies will be resolved through discussion with a third reviewer (SV).

Data extraction

Data extraction will be carried out by one reviewer (CJ) using two tailored data collection forms. The following information to summarise the economic evaluations of WPIs for MSD will be recorded;

- Author, Year
- Population/ MSD

- Objective
- Study Type
- Intervention
- Perspective and Time Horizon
- Outcomes
- Costs
- Results

For a detailed assessment of the treatment of productivity in the included economic evaluations, the following information will be recorded:

- Is productivity loss treated as an outcome or a cost
- Productivity valuation methods
- Productivity estimates
- How the estimates are incorporated into the final analysis

Quality Assessment

The quality of each identified relevant economic evaluation will be assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) (Husereau, Drummond, Petrou, Carswell, Moher, Greenberg, Augustovski, Briggs, Mauskopf, Loder, et al. 2013)

Data Synthesis

The extracted data will be tabulated and summarised as part of a narrative synthesis.

References

Anis, A., W. Zhang, P. Emery, H. Sun, A. Singh, B. Freundlich, and R. Sato. 2009. 'The Effect of Etanercept on Work Productivity in Patients with Early Active Rheumatoid Arthritis: Results from the COMET Study'. *Rheumatology (Oxford, England)* 48 (10):1283–89. https://doi.org/10.1093/rheumatology/kep239.

Bevan, S., E. Passmore, and M. Mahdon. 2007. 'Fit for Work? Musculoskeletal Disorders and Labour Market Participation'.

Burton, K., and N. Kendall. 2014. 'Musculoskeletal Disorders'. *BMJ* 348 (February):bmj.g1076. https://doi.org/10.1136/bmj.g1076.

Drummond, M. F., M. J. Sculpher, G. W. Torrance, B. J. O'Brien, and G. L. Stoddart 2015. 'Methods for the Economic Evaluation of Health Care Programmes'. 4th Edition. Oxford University Press Catalogue. Oxford University Press. https://ideas.repec.org/b/oxp/obooks/9780198529453.html.

Hillage, J., J. Rick, H. Pilgrim, C. Carroll, and A. Booth. 2008. 'Evidence Review 1:Review of the Effectiveness and Cost Effectiveness of Interventions, Strategies, Programmes and Policies to Reduce the Number of Employees Who Move from Short-Term to Long-Term Sickness Absence and to Help Employees on Long-Term Sickness Absence Return to Work'. http://www.nice.org.uk/nicemedia/pdf/PH19evidenceReview1.pdf.

Hoving, J. L., D. Lacaille, D. M. Urquhart, T. J. Hannu, J. K. Sluiter, and M. H. W. Frings-Dresen. 2014. 'Non-Pharmacological Interventions for Preventing Job Loss in Workers with Inflammatory Arthritis'. *The Cochrane Database of Systematic Reviews* 11:CD010208. https://doi.org/10.1002/14651858.CD010208.pub2.

Husereau, D., M. Drummond, S. Petrou, C. Carswell, D. Moher, D. Greenberg, F. Augustovski et al. 2013. 'Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force'. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 16 (2):231–50. https://doi.org/10.1016/j.jval.2013.02.002.

Huver, B., S. Richard, and N. Vaneecloo2012. 'Sick but at Work. An Econometric Approach to Presenteeism.' *15th IZA European Summer School in Labor Economics*.

Johns, G. 2010. 'Presenteeism in the Workplace: A Review and Research Agenda'. *Journal of Organizational Behavior* 31 (4):519–42. https://doi.org/10.1002/job.630.

Kavanaugh, A., J. S. Smolen, P. Emery, O. Purcaru, E. Keystone, L. Richard, V. Strand, and R. F. van Vollenhoven. 2009. 'Effect of Certolizumab Pegol with Methotrexate on Home and Work Place Productivity and Social Activities in Patients with Active Rheumatoid Arthritis'. *Arthritis and Rheumatism* 61 (11):1592–1600. https://doi.org/10.1002/art.24828.

Krol, M., W. Brouwer and F. Rutten, 2013. 'Productivity Costs in Economic Evaluations: Past, Present, Future'. *PharmacoEconomics* 31 (7):537–49. https://doi.org/10.1007/s40273-013-0056-3.

Maini, M. Al., F. Adelowo, J. Al. Saleh, Y. Al. Weshahi, G. R. Burmester, M. J. Cutolo, J. Flood, et al. 2015. 'The Global Challenges and Opportunities in the Practice of Rheumatology: White Paper by the World Forum on Rheumatic and Musculoskeletal Diseases'. *Clinical Rheumatology* 34 (5):819–29. https://doi.org/10.1007/s10067-014-2841-6.

Mattke, S., A. Balakrishnan, G. Bergamo, and S. J. Newberry. 2007. ‘A Review of Methods to Measure Health-Related Productivy Loss’. *American Journal of Managed Care* 13 (4):211–17.

Nicholson, S., M. V. Pauly, D. Polsky, C. Sharda, H. Szrek, and M. L. Berger. 2006. ‘Measuring the Effects of Work Loss on Productivity with Team Production’. *Health Economics* 15 (2):111–23. https://doi.org/10.1002/hec.1052.

Palmer, K. T., E. C. Harris, C. Linaker, M. Barker, W. Lawrence, C. Cooper, and D. Coggon. 2012. ‘Effectiveness of Community- and Workplace-Based Interventions to Manage Musculoskeletal-Related Sickness Absence and Job Loss: A Systematic Review’. *Rheumatology* 51 (2):230–42. https://doi.org/10.1093/rheumatology/ker086.

Public Health England. 2014. ‘Local Action on Health Inequalities: Workplace Interventions to Improve Health and Well-Being’. PHE publications. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/355773/Briefing5a_Workplace_interventions_health_inequalities.pdf.

Punnett, L., and D. H. Wegman. 2004. ‘Work-Related Musculoskeletal Disorders: The Epidemiologic Evidence and the Debate’. *Journal of Electromyography and Kinesiology*, State of the art research perspectives on muscoskeletal disorder causation and control, 14 (1):13–23. https://doi.org/10.1016/j.jelekin.2003.09.015.

Schultz, A. B., and D. W. Edington. 2007. ‘Employee Health and Presenteeism: A Systematic Review’. *Journal of Occupational Rehabilitation* 17 (3):547–79. https://doi.org/10.1007/s10926-007-9096-x.

Uegaki, K., M.C. de Bruijne, A. J. van der Beek, W. van Mechelen, and M. W. van Tulder. 2011. ‘Economic Evaluations of Occupational Health Interventions from a Company’s Perspective: A Systematic Review of Methods to Estimate the Cost of Health-Related Productivity Loss’. *Journal of Occupational Rehabilitation* 21 (1):90–99. https://doi.org/10.1007/s10926-010-9258-0.

Uegaki, K., M. C. de Bruijne, L. Lambeek, J. R. Anema, A. J. van der Beek, van W. Mechelen, and M. W. van Tulder. 2010. ‘Economic Evaluations of Occupational Health Interventions from a Corporate Perspective - a Systematic Review of Methodological Quality’. *Scandinavian Journal of Work, Environment & Health* 36 (4):273–88.

van Vollenhoven, R F., M. A. Cifaldi, S. Ray, N. Chen, and M. H. Weisman. 2010. ‘Improvement in Work Place and Household Productivity for Patients with Early Rheumatoid Arthritis Treated with Adalimumab plus Methotrexate: Work Outcomes and Their Correlations with Clinical and Radiographic Measures from a Randomized Controlled Trial Companion Study’. *Arthritis Care & Research* 62 (2):226–34. https://doi.org/10.1002/acr.20072.

Zhang, W., N. Bansback, and A. H. Anis. 2011. ‘Measuring and Valuing Productivity Loss due to Poor Health: A Critical Review’. *Social Science & Medicine* 72 (2):185–92. https://doi.org/10.1016/j.socscimed.2010.10.026.

Appendix 4.2

Search Strategy for Economic Evaluations of Workplace Interventions for Musculoskeletal Conditions

Medline, Embase and Econlit via OVID SP

(1st January 2007 to 17th September 2017)

1. Long term absen$.ti,ab.
2. Long term sick$.ti,ab.
3. exp sick leave/
4. (sick$ adj3 leave).ti,ab.
5. (sick adj3 absen$).ti,ab.
6. (work adj3 absen$).ti,ab.
7. (return$ adj3 work$).ti,ab.
8. work readiness.ti,ab.
9. Sick$ benefit$.ti,ab.
10. Disability leave.ti,ab.
11. (injur$ adj3 claim$).ti,ab.
12. (stay$ adj3 work$).ti,ab.
13. (participat$ adj3 employ$).ti,ab.
14. (attend$ adj3 work$).ti,ab.
15. (attend$ adj3 employ$).ti,ab.
16. Absenteeism/
17. (reduc$ adj3 work$ adj3 perform$).ti,ab.
18. ((sick$ or illness$ or employee$) adj3 absenteeism).ti,ab.
19. (welfare adj3 work$).ti,ab.
20. (sicklist$ or sick list$).ti,ab.
21. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21
22. Economics/
23. exp "costs and cost analysis"/
24. Economics, Dental/
25. exp economics, hospital/

26. Economics, Medical/

27. Economics, Nursing/

28. Economics, Pharmaceutical/

29. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

30. (expenditure$ not energy).ti,ab.

31. value for money.ti,ab.

32. budget$.ti,ab.

33. or/22 - 32

34. ((energy or oxygen) adj expenditure).ti,ab.

35. (metabolic adj cost).ti,ab.

36. ((energy or oxygen) adj expenditure).ti,ab.

37. 34 or 35 or 36

38. 33 not 37

39. letter.pt.

40. editorial.pt.

41. historical article.pt.

42. 39 or 40 or 41

43. 38 not 42

44. exp animals/ not humans/

45. 43 not 44

46. bmj.jn.

47. "cochrane database of systematic reviews".jn.

48. health technology assessment winchester england.jn.

49. journal of medical economics.jn.

50. 46 or 47 or 48 or 49

51. 45 not 50

52. 21 and 51

53. limit 52 to yr=”2007 –Current”

Appendix 4.3

Definitions of Rheumatic Musculoskeletal Conditions

RMDs defined by Al Maini et al. (2015)
Low back pain
Neck pain
Osteoporosis
Osteoarthritis
Rheumatoid Arthritis
Gout
Ankylosing Spondylitis
Arthralgia
Behcet's disease
Carpel tunnel syndrome
Epicondylitis
Fibromyalgia
Myofascial pain syndrome
Psoriatic arthritis
Scoliosis
Seronegative spondyloarthropathies
Soft tissue rheumatism
Systemic lupus erythematosus
Tendinitis
(Other unclassified RMD conditions)

Appendix 4.4

Chec-list to assess the quality of economic evaluations

	CHEC-list	YES	NO
1.	Is the study population clearly described?		
2.	Are competing alternatives clearly described?		
3.	Is a well-defined research question posed in answerable form?		
4.	Is the economic study design appropriate to the stated objective?		
5.	Is the chosen time horizon appropriate in order to include relevant costs and consequences?		
6.	Is the actual perspective chosen appropriate?		
7.	Are all important and relevant costs for each alternative identified?		
8.	Are all costs measured appropriately in physical units?		
9.	Are costs valued appropriately?		
10.	Are all important and relevant outcomes for each alternative identified?		
11.	Are all outcomes measured appropriately?		
12.	Are outcomes valued appropriately?		
13.	Is an incremental analysis of costs and outcomes of alternatives performed?		
14.	Are all future costs and outcomes discounted appropriately?		
15.	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?		
16.	Do the conclusions follow from the data reported?		
17.	Does the study discuss the generalizability of the results to other settings and patient/client groups?		
18.	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?		
19.	Are ethical and distributional issues discussed appropriately?		

Appendix 4.5

Consolidated Health Economic Evaluation Reporting Standards (CHEERS) quality reporting criteria for economic evaluations

The table presents a list of the main items included in the CHEERs checklist to assess the quality of the reporting standards of economic evaluations.

Section/Item	Item Number	Reported on page no./line no.
Title and abstract		
Title	1	
Abstract	2	
Introduction		
Background and objectives	3	
Methods		
Target populations and subgroups	4	
Setting and location	5	
Study Perspective	6	
Comparators	7	
Time horizon	8	
Discount rate	9	
Choice of health outcome	10	
Measurement of effectiveness	11a 11b	
Measurement and valuation of preference-based outcomes	12	
Estimating resources and costs	13a 13b	
Currency, price date and conversion	14	
Choice of model	15	
Assumptions	16	
Analytic methods	17	
Results		
Study parameters	18	
Incremental costs and outcomes	19	
Characterising uncertainty	20a 20b	
Characterising heteroscedasticity	21	
Discussion		
Study findings, limitations, generalisability, and current knowledge	22	
Other		
Source of funding	23	
Conflicts of interest	24	

Appendix 5.1

Participant Information Sheet for Qualitative Study

Exploring the impact of health status and well-being in people with Inflammatory Arthritis and how this may affect an individual's ability to work
Working People with Rheumatoid Arthritis, Ankylosing Spondylitis or Psoriatic Arthritis

Participants Information Sheet

You are being invited to take part in a research study designed to identify those aspects of rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis that may affect the ability to work effectively. The research will contribute towards the PhD thesis of Miss Cheryl Jones, at The University of Manchester. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take your time to read through the following information carefully and discuss with others if you wish. If you have any further questions, please do not hesitate to contact us for further information or clarification. Please take some time to decide whether or not you wish to take part in this study. Thank you for reading.

Who will conduct the research?

The research will be conducted by Cheryl Jones, PhD student, who is based in the Manchester Centre for Health Economics at The University of Manchester.

What is the purpose of this research?

We know that patients with arthritis may have difficulties at work and be less efficient than people who do not have arthritis. The aim of this research is to identify any aspects of rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis that may have an impact of how you work.

Why have I been chosen?

You have been chosen because you have been diagnosed with rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis and are currently working for pay.

What would I be asked to do if I took part?

By agreeing to take part, you will participate in a telephone interview designed to talk about your experience of working with a chronic disease such as rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis. The telephone interview will be audio-recorded using a special recording device. The interview should last around 30 minutes and no longer than 45 minutes.

What happens to the data collected?

Your interview responses, along with the responses of others who have also participated in the study, will be transcribed within 15 days of your interview by Cheryl Jones. Your transcribed responses will be made anonymous, after which it will not be possible to remove your responses or your consent for participation in this study. Your responses will then be examined to identify the most common and important aspects that affect employees with rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis and their ability to work effectively.

How is confidentiality maintained?

Your name and any other personal information from which you could be identified will be kept separately from your other data. No-one outside the research team will have access to any identifying information and all identifiable information will be stored securely and handled according to the 1998 Data Protection Act. All data collected as part of the interview will be stored in another database using unique numbers. Quotations will be included in the report however these will be made anonymous.

What happens if I do not want to take part or if I change my mind?

If you decide to take part in this study you are free to withdraw yourself from the study at any time and you do not have to provide reasons for your withdrawal. However, it will not be possible to withdraw your data after it has been made anonymous.

How do I provide my consent to take part in this study?

If you would like to take part in the study you will be asked to provide your consent to participate in the study over the telephone before the interview takes place.

Will I be paid for participating in the research?

You will receive no payment for taking part in this research.

What is the duration of the research?

The interview will last no longer than 45 minutes.

Where will the research be conducted?

The time and date of the telephone interview can be chosen by you, at a time that is most convenient for you. The researcher will be based at The University of Manchester.

Will the outcomes of the research be published?

The outcomes of the research will contribute towards the PhD thesis of Cheryl Jones. Findings will be submitted for publication in peer-reviewed journals, and for presentation at national and international conferences. All responses from your participation will remain strictly anonymous. Once the study has been published, a copy of the results will be available on charity/support group websites including NRAS, NASS and PsAZZ.

What happens if the study has recruited its maximum number of participants?

This study has specified a maximum number of participants that need to be recruited. Once this maximum has been reached it will not be necessary to continue collecting further data. If you express your interest in participating in the study, and the maximum participant levels have already been reached, an email will be sent to you to let you know that we are no longer collecting data and that your participation in this study is not required at this time.

What if something goes wrong?

If there are any issues regarding this research that you would prefer not to discuss with members of the research team, please contact the Research Governance and Integrity Team by either writing to 'The Research and Governance and Integrity Manager, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL', by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

If there are any issues regarding this research that you would prefer to discuss with a support group or charity please contact the relevant organisation listed on the next page.

Who is funding this research?

This research is funded by Arthritis Research UK (ARUK) and the Medical Research Council (MRC) grant reference: 20665.

Who has reviewed the research project?

Ethical approval for this study has been granted by The University of Manchester Ethics Committee *reference number: 16144.*

Contact for Further Information

Lead Researcher

Miss Cheryl Jones

Manchester Centre for Health Economics

Room 4.306, Jean McFarlane Building

The University of Manchester

Oxford Road

M13 9PL

Email: cheryl.jones@manchester.ac.uk

Supervisor

Dr. Suzanne Verstappen

Manchester Academic Health Science Centre

Room 2.600

Stopford Building

Oxford Road

M13 9PT

Email: suzanne.verstappen@manchester.ac.uk

Support Groups and Charities Contact Details

Support Group / Charity	Contact Details
National Rheumatoid Arthritis Society (NRAS)	Tel: 0800 298 7650 Email: helpline@nras.org.uk Website: www.nras.org.uk
National Ankylosing Spondylitis Society (NASS)	Tel: 020 8741 1515 Email: asknass@nass.co.uk Website: www.nass.co.uk
The Psoriasis Association	Tel: 08456 760 076 Email: mail@psoriasis-association.org.uk Website: www.psoriasis-association.org.uk
PsAZZ Support Group	Website: http://psazzgroup.wix.com/psazz
Arthritis Research UK (ARUK)	Website: www.arthritisresearchuk.org

Appendix 5.2

Interview Schedule for employees with RA, AS or PsA

Exploring the impact of health status and well-being in people with Inflammatory Arthritis on presenteeism in the workplace

Interview Schedule

Employees with RA, AS or PsA

Introduction

Introduction:

Hello, my name is Cheryl Jones. I am a PhD student in the Manchester Centre for Health Economics, at The University of Manchester. Thank you for agreeing to take part in my research. I plan to talk to you for around 30 to 45 minutes, to understand what aspects or symptoms of your condition (RA, AS, PsA) make work potentially more challenging or difficult for you.

Before we begin, I must remind you that our conversation will be recorded, but any information that you give will remain fully anonymous. Please do not mention any individual's names or any sensitive information regarding the company you work for. If names are mentioned please don't worry, I will remind you to avoid using names and the interview will continue. All names and other sensitive information that might come up in the interview will be removed from the transcripts. Your participation is voluntary, and you are free to end the session at any time without giving any reason. Please also understand that there are no right or wrong answers, I am interested in your experiences of working with RA, AS or PsA. Feel free to ask any questions at any time, during or after the interview.

Are you still happy to take part?

YES NO

Semi- structured Interview Schedule

- Press record on the recording device now

Section A: Health Details

1. Please tell me what rheumatic/arthritic condition you have been diagnosed with?
    a. When were you diagnosed?
    b. How long did it take to be diagnosed?
    c. Are you taking any medication for your arthritis at the moment?
    d. What are you taking?
    e. Do you think it is effective?

2. In your opinion, how severely does your condition affect you day-to-day?

3. In terms of your condition, do you experience 'good' days and 'bad' days?
    - *If yes;*
    a. How does a 'bad' day differ in terms of their effect on your day-to-day activities compared to a 'good' day?
    b. How often do you experience 'bad' days?
    c. How would you describe the progression of your disease? *(increasing over time, stable, in remission)*
    - *If No;*
    d. How would you describe the progression of your disease? *(increasing over time, stable, in remission)*

Section B: Job Details

The following questions are about your current job

4. What is your job title?

5. How would you describe the type of industry that you work in? (for example construction, retail)

6. During a typical working day, what tasks do you perform?

7. Do you work as part of a team?
    - *No;*

a. You work independently throughout the working week?

- *Yes;*

a. How many people work in your team?

b. Do you manage anyone in this team?

c. How many people are you managed by?

8. On those days you are absent from work, can your colleagues fill in for you and complete some of your work?

- *No;*

a. What happens to your work?

- *Yes;*

b. Who is able to do your work whilst you are away?

c. How much of your work can be done by your colleagues?

Section C: Health and Work Performance

9. Bearing in mind your condition, do you feel you are able to achieve what you want to achieve at your job? *(for example, to gain a promotion or to do your job to the best of your ability)*

- *Yes;*

a. Are there any aspects of your job or workplace that help you to achieve your goals?

- *No;*

b. What do you feel limits your ability to achieve your goals at work?

10. You mentioned earlier that you experience 'good' days and 'bad' days, in terms of your condition. How does your condition affect your ability to work on a 'bad' day compared to a 'good' day?

- Does your condition affect you at work when you are experiencing a 'good' day? If so, how?

11. Does your condition reduce the amount of work you can complete in a given amount of time?

- *No;*
- *Yes;*

a. What happens to the work that is not completed? For example;
   i. Work overtime
   ii. Colleagues take work to complete
   iii. Nothing – output is reduced

12. Does your condition reduce the quality (as opposed to quantity) of work you are able to produce?
    - *No;*
    - *Yes;*
        a. In what ways do you think your quality of work is reduced?

13. Has your condition ever compromised the safety of yourself or others at your workplace?
    - *No;*
    - *Yes;*
        a. Can you describe what happened?

14. In the past, have you had to change your job or alter your job, for example the hours you work, the types of tasks you complete, because of your condition?
    - *No;*
    - *Yes;*
    a. Can you please describe what happened?
    b. The changes that were made to your job because of your condition, how did they help you work more effectively?

15. Imagine you are able to make changes or further changes to your current workplace or job that would help you to manage your condition and allow you to work more effectively, what would you change? *(for example, starting work at 10am, going to the gym at lunch, raising my desk and standing to do work)*
16. Have you told any person at work (colleagues or managers) about your rheumatic condition?
    - *Yes;*
    a. Who did you tell? *(no names, please describe as colleague or manager)*

b. What made you decide to tell this person?*(trust, supportive workplace)*

c. What was their reaction?

- *No;*

b. What has stopped you from telling someone at work about your condition?

   i. What would help you feel confident to tell your colleagues or manager about your condition?

Work Performance and Other Issues

17. Other than your rheumatic condition, are there any other issues you believe affect your ability to work effectively?

- *No; go to next section*
- *Yes;*

a. Please describe what they are

- **Press stop on the recording device now**

Closing of the interview:

Thank you very much. That's the end of the interview. Thank you for your participation, your responses have been very helpful.
If you have any further questions, feel free to contact either myself or my supervisor – our contact details can be found on the participant information sheet. Have a good day. Bye.

Appendix 5.3

Absenteeism Themes

The inductive themes identified and incorporated into the main study were those that have a direct effect on presenteeism. However, two further themes were identified that impact productivity loss in terms of absenteeism.

Theme 1A: Appearance

The first of these two themes is appearance. A person's appearance is important promoting self-esteem and confidence. Of course, when something negative happens to a person's physical appearance it can make them behave differently. One participant explained that she suffered from a psoriasis rash across her face and it was enough to prevent her from going to work.

'I had psoriasis on my face, psoriasis on my hands urm…. I didn't go off [work] from the psoriasis on my hands but it when it was on my face my skin was… I looked like I had been really badly sunburnt and no amount of make-up can cover that…I looked horrendous and I thought there is no way I am going to go to work looking like this' (P101, AS, female).

Theme 2A Hospital Appointments

The second identified theme was hospital appointments. Individuals' diagnosed with RA, AS, or PsA are monitored closely, especially if they are taking a new medication, such as a biologic, and are expected to attend frequent hospital appointments. Participants agreed that they do their best to arrange appoints at the beginning or end of a working day so that it doesn't interfere too much with their work.

'I… just always try and I get early morning appointments and late afternoon appointments, so it doesn't really, because my job I can do anytime, it's not like I need to be at my desk by a certain time and finish at a certain time, my job is you..you can do my job late' (P110, PsA, female)

The quote by P110 also emphasises the importance of having a degree of flexibility in a job where employees are able to compensate for the hours they have been out of work. However, for some, the amount of hospital appointments has meant reducing working hours and prevented progression in their career.

'I would have preferred to keep me longer hours and to have worked at a higher level… but because I have a lot of hospital appointments, I can arrange them on the days I am not working, and obviously rest and deal with the side effects on the days that I am not working' (P102, PsA, female).

Appendix 6.1

Summary of questions on absenteeism and presenteeism included in self-report measures suitable for rheumatic musculoskeletal disorders

Instrument	**Focus:** (Health, RA, etc)	**Absenteeism/Off Sick**	**Response Format Absenteeism**	**Presenteeism Questions**	**Response format Presenteeism**	**Recall Periods: Absenteeism and Presenteeism**	**Length of Survey (no. of questions)**
Functional Status Questionnaire (Jette et al 1986)	Function: Physical Mental Social Role	No absenteeism questions are included (During the past month, how many days did your illness or injury keeps you in bed all or most of the time?)	(Answer: yes or no)	If you were employed during the past month, how was your work performance? 1. Done as much work as others? 2. Worked for short periods of time/taken frequent rests? 3. Worked regular number of hours? 4. Done job accurately as others in similar job? 5. Worked at usual job, but with some changes because of your health? 6. Feared losing your job because of your health?	Answer using points system: 1= All of the time 2 = most of the time 3= some of the time 4= none of the time	Absenteeism: N/A Presenteeism: Month	33
Osterhaus Technique (OT) (Osterhaus et al 1992)	Health Problem	1. During the past 30 days, how many days of work have you missed because of your health problem?	1. Number of days	1. Compared with your normal work performance, please estimate how effective you are at your job when you have symptoms at work?	1. 100% = fully effective. State %	Absenteeism: 30 days Presenteeism: Not clear	12

WPAI (Reilly et al 1993)	Health Problem	1. During the past 7 days, how many hours did you miss from work because of problems associated with your PROBLEM? *Include hours missed on sick days, time you went in late, left early, etc., because of your problem Do not include time you missed to participate in this study.* Answer: no. hours	1. Number of hours	1. During the past 7 days, how many hours did you actually work? 2. During the past 7 days, how much did your PROBLEM affect your productivity while you were working? *Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If PROBLEM affected your work only a little, choose a low number. Choose a high number if PROBLEM affected your work a great deal.*	1. Number of hours 2. Scale from 0 = problem has no effect to 10 = problem completely prevented me from doing my activities	Absenteeism: 7 days Presenteeism: 7 days	6
HLQ (Van Roijen et al 1996)	Health Problems	1. Respondent is asked to indicate on what days, during the past 2 weeks, they performed paid work, were unable to perform paid work due to holiday or other reason, were unable to perform paid work due to health problems.	Code days worked, did not work due to health problem, and no paid work performed	1. Were you hindered by health problems at you paid work over the past 2 weeks? 2. Please indicate for each statement that is mentioned how often if applied to you in the past two weeks. I did go to work but as a result of health problems; 1. I have a problem concentrating 2. I had to go to work at a slower pace 3. I had to seclude myself 4. I found decision-making more difficult	1. Binary: Yes/No 2. Answer with one of the following; Never, sometimes, often, almost always 3. Number of hours	Absenteeism: 2 weeks Presenteeism: 2 weeks	23

				5. I had to put off some of my work 6. Others has to take over some of my work 7. I had other problems, (please state) 3. How many extra hours would you have to work to catch up on tasks you were unable to complete in normal working hours due to health problems over the past two weeks? Answer: no. hours (*Do not count the days on which you were reported sick*)			
Endicott Work Productivity Scale (EWPS) Endicott et al 1997	Attitudes and behaviour that affect work performance and efficiency	1. How many hours did you work last week? 2. How many hours do you usually work or would usually be expected to work? 3. If you missed time off work last week, please note all the reasons why	1. Number of hours per week 2. Number of hours per week 3. Day off (holiday), physically ill, upset/depressed, other	1. During the past week, how frequently did you… a. Arrive at work late or leave work early? b. Take longer lunch hours or coffee breaks c. Just do no work at times when you would be expected to be working? d. Find yourself daydreaming, worrying or staring into space when you should be working? e. Have to do a job over because you made a mistake or your supervisor told you to do a job over? f. Waste time looking for misplaced supplies, materials, papers, phone number, etc.? g. Find you have forgotten to call someone? h. Find you have forgotten to	1. Answer: 0 = never to 4 = almost always	Absenteeism: 1 Week Presenteeism: 1 Week	33

	respond to a request? i. Become annoyed with or irritated by co-workers, boss/supervisors, clients/customers/vendors or others? j. Become impatient with others at work? k. Avoid attending meetings? l. Avoid interaction with co-workers, clients, vendors or supervisors? m. Have a co-worker redo something you had completed? n. Find it difficult to concentrate on the task at hand? o. Fall asleep unexpectedly or become very sleepy while at work? p. Become restless while at work? q. Notice that your productivity for the time spent is lower than expected r. Notice that your efficiency forth time spent is lower than expected? s. Lose interest or become bored with your work? t. Work more slowly or take longer to complete tasks than expected? u. Have your boss/co-workers remind you to do things? v. Not want to return phone calls or put off returning calls? w. Have trouble organising work or setting priorities? x. Have trouble organising work or setting priorities? y. Fail to finish assigned tasks? z. Feel too exhausted to do your	

				work?			
Occupational Role Performance (ORQ) Kopec and Esdaile 1998	Back Pain	Presenteeism only	N/A	1. Would you say that, because of back pain, a. You cut down on the amount of extra work or overtime? b. You work more slowly? c. You take more frequent or longer breaks? d. You are less able to concentrate on your work? e. You have fewer opportunities to upgrade your skills? f. You have less satisfaction with your job? g. You need more help from your co-workers?	1. Respondent chooses from; A lot = 3 Somewhat = 2 A little = 1 Not at all = 0	Not specified	8
The Quality and Quantity Method (QQ) Brouwer, Koopmanschap and Rutten 1999	Consequences of illness	Presenteeism only	N/A	1. Could you indicate how much work you actually performed today during regular hours as compared to normal on the scale below? 2. Could you indicate the quality of the work you performed today as compared to normal on the scale below?	1. Practically nothing 1 2 3 4 5 6 7 8 9 10 Normal 2. Very poor quality 1 2 3 4 5 6 7 8 9 10 Normal	Presenteeism: One day One week	3
Work Productivity Index (WPI) Burton et al	Health Problems	1. Total Illness hours + total short-term disability*hours/ Weeks Employed	1. Lost hours per week due to Absenteeism	1. [100% - (overall score)/(0.5)] x Average staffing hours	1. Lost hours per week due to failure to meet productivity standard	Absenteeism: 1 week Presenteeism: 1 Week	1

1999							
Work Role Functioning Questionnaire (WRFQ) Amick et al 2000	Generic role-specific measure	N/A	N/A	1. Getting to work on time 2 Sticking to a routine or schedule without having to re-arrange or reassign your work tasks 3 Working without taking frequent rests or breaks to avoid discomfort 4. Working the required number of hours 5. Handing very demanding or stressful work situations 6. Doing your work without becoming tense or frustrated 7. Doing your work carefully 8. Satisfying those people who judge your work 9. Feeling a sense of accomplishment 10. Finishing work on time Handling the workload 12. Lifting, carrying, or moving objects at work weighing 10 pounds or less 13. Lifting, carrying, or moving objects at work weighing more than 10 pounds 14. Walking more than one block or climbing up or down on flight of stairs while working 15. Sitting, standing, or staying in one position for longer than 15 minutes while working 16. Bending, twisting, or reaching while working 17. Using hand-operated tools or equipment (for example, a pen,	Answer all with one of the following; None of the time Half of the time All of the time	Presenteeism: 4 weeks	26

			drill, sander, keyboard, or computer mouse) 18. Using your upper body to operate tools or equipment (upper body means arms, head neck, shoulders or upper back) 19. Using your lower body to operate tools or equipment (lower body means legs, knees, feel, lower back) 20. Keeping your mind on your work 21. Keeping track of more than one task or project at the same time 22. Concentrating on your work 23. Remembering things having to do with your work 24.Talking to people in person, in meetings, or on the phone 25. Controlling irritability or anger towards people when working 26. Helping other people to get work done			
Presenteeism Scale (SAHAPS) Pelletier and Koopman 2001 USA (see reference; Lynch, Mercer, and	N/A	N/A	42 items grouped in 2 modules; 3) Demographics 4) Presenteeism	Rate their performance using 5 point Likert scales and 0-100% scales	Presenteeism: 1 month	42

Riedel 2001)							
The Work Limitations Questionnaire (WLQ) Lerner et al 2001	Chronic health problem	N/A	N/A	1. In the past 2 weeks, how much of the time did your physical health or emotional problems make it difficult for you to do the following? a. Do you work without stopping/taking breaks b. Stick to a routine or schedule c. Keep your mind on your work d. Speak with people in person, in meetings or on the phone e. Handle the workload 2. In the past 2 weeks, how much of the time were you able to do the following without difficulty caused by physical health or emotional problems? a. walk or move around different work locations (for example, go to meetings) b. use hand-held tools or equipment (for example, a phone, pen, keyboard, computer, mouse, drill, hairdryer or sander)	Answer all with one of the following; 1 = All of the time (100%) 2 = Most of the time 3 = Some of the time (about 50%) 4 = A slight bit of the time 5 = None of the time (0%) 0 = Does not apply to my job	Presenteeism: 2 weeks	7
Stanford Presenteeism Scale – 6 (SPS-6) Koopman et al 2002	Health Problems	N/A	N/A	1. Because of my (health problem)*, the stresses of my job were much harder to handle 2. Despite having my (health problem)*, I was able to finish hard tasks in my work 3. My (health problem)* distracted me from taking pleasure in my work	Respondent chooses one of the following; 1 = if you strongly disagree with the statement 2= if you somewhat	Presenteeism: Month	6

				4. I felt hopeless about finishing certain work tasks due to my (health problem)* 5. At work, I was able to focus on achieving my goals despite my (health problem)* 6. Despite having my (health problem),* I felt energetic enough to complete all my work. *Note that other words such as "back pain" can be substituted in for "health problems"	disagree 3 = if you are uncertain about your agreement with the statement 4. If you somewhat agree with the statement 5. If you strongly agree with the statement		
Work Instability Scale (WIS) Gilworth et al 2003 UK	Health Condition	N/A	N/A	1. I'm getting up earlier because of my condition 2. I get very stiff at work 3. I'm finding my job is about all I can manage 4. The stress of my job makes my condition flare 5. I'm finding any pressure on my hands is a problem 6. I get good days and bad days at work 7. I can get my job done, I'm just a lot slower 8. If I don't reduce my hours I may have to give up work 9. I am very worried about my ability to keep working 10. I have pain or stiffness all the time at work 11. I don't have the stamina to work, like I used to 12. I have used my holidays so	Answer; True = 1 or false = 0	Presenteeism Not Reported	23

				that I don't have to go sick 13. I push myself to go to work because I don't want to give in to my condition 14. Sometimes I can't face being at work all day 15. I have to no to certain things at work 16. I've got to watch how much I do certain things at work 17. I have great difficulty opening some of the doors at work 18. I have to allow myself extra time to do some jobs 19. It's very frustrating because I can't always do things at work 20. I feel I may have to give up work 21. I get on with work but afterwards I have a lot of pain 22. When I'm feeling tired all the time work's a grind 23. I'd like another job but I am restricted to what I can do			
Work Productivity Short Inventory (WPSI) Goetzel et al	15 medical problems; allergies, respiratory infections, arthritis, asthma,	1. During the past year, estimate the total DAYS you MISSED FROM WORL because of ARTHRITIS/RHEUMATISM and	1. Number of days	1. During a typical 8 hour workday, estimate the total HOURS you were UNPRODUCTIVE because of ARTHRITIS/RHEUMATISM and related symptoms.	1. Number of hours	Absenteeism: 1 year Presenteeism: 8 hour work-day	22

2003	anxiety disorders, depression and bipolar, stress, diabetes, hypertension, migraine and other major headaches, coronary heart disease/ high cholesterol	related symptoms..					
Work Performance Questionnaire (HPQ) Kessler et al 2003	Health Problems	1. In the past 28 days, how many days did you…. a. miss an entire work day because of problems with your physical or mental health? b. miss an entire work day for any other reason (including vacation)? c. miss part of a work day because of problems with your physical or mental health?	1. Number of days	1. The next questions are about the time you spent during your hours at work in the past 4 weeks. Circle the one number from each question that comes closest to your experience. a. How often was your performance higher than most workers on your job? b. How often was your performance lower than most workers on your job? c. How often did you do no work at times when you were supposed to be working? d. How often did you find yourself not working as carefully as you should? e. How often was the quality of	Responses: 1= All of the time 2 = Most of the time 3 = Some of the time 4 = A little of the time 5 = None of the time 2. Circle number between 0 (worst) and 10 (best) 3. Circle number	Absenteeism: 28 days Presenteeism: 4 weeks (28 days)	20

	d. miss part of a work day for any other reason (including vacations)? e. come in early, go home late, or work on your day off?	your work lower than it should have been? f. How often did you not concentrate enough on your work? g. How often did health problems limit the kind or amount of work you could do? 2. On a scale from 0 to 10 where 0 is the worse job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate the usual performance of most workers in a job similar to yours? 3. Using the same 0 to10 scale, how would you rate your usual job performance over the past year or two? 4. Using the same 0 to 10 scale, how would you rate your overall performance on the days you worked during the past 4 weeks? 5. How would you compare your overall job performance on the days you worked during the past 4 weeks with the performance of most other workers who have a similar type of job? 6. This week, how frequently did you: a. Become annoyed with or irritated by co-workers, boss/supervisor,	between 0 and 10. 4. Circle number between 0 and 10. 5. Circle one of the following; 1 – you were a lot better than other workers 2 – you were somewhat better than other workers 3 – you were a little better than other workers 4- you were about average 5 – you were a little worse than other workers 6 – you were somewhat worse than other workers 7 – you were a lot worse than other workers 6. Choose from 1 – 10. 1 = Never 10 = Almost always

				clients/customers/vendors or others? b. Become impatient with others at work? c. How often did you get into conflicts with others at work? d. Become restless while at work? e. Lost interest or become bored with your work? f. Had difficulty concentrating at work? g. Fail to finish assigned tasks? 24. Feel too exhausted to do you work?			
Health-related Productivity Questionnaire – Diary (HRPQ-D) Kumar et al 2003	Health conditions	1. Record number of hours work missed today due to health problem.	1. Number of hours	1. For the hours you did work outside the home, how did your health symptoms impact your effectiveness?	1. Choose one of the following; 90-100%, 75-89%, 50-74%, 49% or less of my usual	Absenteeism: 1 day Presenteeism: 1 day	9 per day
Health and Work Questionnaire (HWQ) Shikar et al 2004	General – not health specific	N/A	N/A	1. How would you and the following people (self, supervisor, co-workers) describe your efficiency this week? 2. How would you and the following people describe the overall quality of your work this week? 3. How would you and the following people describe the overall amount of work you did this	1. Choose number on scale 1 = worst ever to 10 Best possible. 2. Choose number on scale 1 = worst ever to 10 Best possible. 3. Choose	Presenteeism: 1 week	24

				week? 4. Rate your highest level of efficiency this week 5. Rate your lowest level of efficiency this week	number on scale 1 = worst ever to 10 Best possible. 4. Choose number on scale 1 = worst ever to 10 Best possible. 5. Choose number on scale 1 = worst ever to 10 Best possible.		
Work Health Interview (WHI) Stewart et al 2004							
Stanford Presenteeism Scale – 13 (SPS-13) Turpin et al 2004	Health Conditions	1. Because of your primary condition as you identified in question 1, how many work hours did you miss in the past 4 weeks?	1. Answer: Place X on continuous scale 0 to 40+ (number of hours)	1. In thinking about how your primary condition affected your ability to do your job, how often in the past 4 weeks? a. were you able to finish hard tasks? b. Did you find your attention wandering? c. Were you able to focus on achieving work goals? d. Did you feel energetic enough to work? e. Were the stresses of your job hard to handle? f. Did you feel hopeless about	1. Answer with one of the following; Always Frequently About half the time Occasionally Never No Answer 2. Answer: Place X on continuous 1 – 100 scale.	Absenteeism: 4 weeks Presenteeism: 4 weeks	13

				finishing your work? g. Were you able to focus on finding a solution when unexpected problems arose in your work? h. Did you need to take breaks from your work? i. Were you able to work with other people on shared tasks? j. Were you tired because you lost sleep? 2. Given your primary health condition, what percentage of your usual productivity level were you able to achieve while working over the last 4 weeks?			
Work Ability Index (WAI) Ilmarinen et al 2007	Health	1. During the last 12 months: how many whole days have you been off work?	1. Chose one from following: none max. 9 days 10-24 days 25-99 days 100 – 354 days 3. Estimated work impairment due to diseases. Answer: 1 – 6	1. Current work ability compared to highest work ability. Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability? 2. Work ability in relation to the demands. a. How do you rate your current work ability with respect to the physical demands of your work? b. How do you rate your current work ability with respect to the mental demands of your work? 3. Is your illness or injury a hindrance to your current job?	1. Scale 0 to 10. 0 = Completely unable to work 10 = work ability at its best 2. Very good, rather good, moderate, rather poor, very poor 3. Circle more than one alternative if needed: a. There is no hindrance/ I have no diseases b. I am able to do my job, but it	Absenteeism: 12 months Presenteeism: No recall period specified	25

					causes symptoms c. I must sometimes slow down my work pace or change my work methods d. I must often slow down my work pace or change my work methods e. Because of my condition, I feel I am able to do only part time work f. In my opinion, I am entirely unable to work		
Work Productivity Survey for Rheumatoid Arthritis (WPS-RA) Osterhaus, Purcaru and Richard 2009 USA	Rheumatoid Arthritis	1. How many days did the patient miss work because of arthritis? (if none please write 0)	1. Number of days	1. How many days in the last month was the patient's productivity at work reduced by half or more because of arthritis? 2. In the last month, how much has arthritis interfered with the patient's work productivity (work outside of home) on a scale of 0 – 10, where 0 = 'no interference' and 10 = 'complete interference'	1. Number of days 2. Scale; 0= no interference 10 = complete interference	Absenteeism: 1 Month Presenteeism: 1 Month	10
Short Form HLQ (Van Roijen et al	Health Problems	1. Did health problems oblige you to be off work	1. Binary Response: Yes/ No	1. Was your performance adversely affected by health problems during the past month?	1. Answer: No, not at all, Yes slightly, Yes,	Absenteeism: Month	11

2010)	at any time in the past month? 2. Were you off work for a period longer than the past month because of health problems?	2. Binary response: Yes/No and report number of days	2. On how many days during the past month did you perform paid work, although you were bothered by health problems? (*Please do not count the days on which you did not work at all because you called in sick*) 3. Please rate how well you performed on the days you went to work even though you were bothered by health problems. *(1 indicates a much worse performance than usual and 10 that your work was not affected.)* 4. Please indicate how often each statement was applicable to you during the past month. I went to work, but as a result of health problems…. • • • • • • • 1. I had concentration problems 2. I had to work at a slower pace 3. I had to work in seclusion 4. I had more difficulty making decisions 5. I had to postpone work 6. Others had to take over my work 7. I had different problems, namely:…. 5. If you had to catch up on all the work you were unable to perform over the past month because of	very much 2. Number of days 3. Scale; 0= Worse than usual 10 = Performed as usual 4. Answer with one of the following; Never, sometimes, often, almost always 5. Number of hours	Presenteeism: Month

				health problems, how many hours of work would you be forced to make up? *(The days on which you failed to work at all because you reported ill do* ***not count****)*			
Well-being Assessment for Productivity (WBA-P) Prochaska et al 2011							
The Valuation of Lost Productivity (VOLP) Zhang et al 2012	Health	1. In the past 3 months, how many work days in total have you been absent from work because of your health (any physical, mental, emotional problems or symptoms)? *Please include work days you missed due to your health, and/or partial work days where you went in late or left early due to your health (e.g. doctor appointments). Do*	1. Number of work days	1. In the past 7 days, have you gone to work? 2. Think of all the work you have completed during the past 7 days. Would you complete the same work in LESS TIME if you did NOT experience any health problems (*i.e., any physical, mental, or emotional problems or symptoms)*? 3. If yes (to 2 above), please indicate the time you took to complete all your work in the past 7 days and the time you would take to complete the same work if you did NOT experience any health problems: Answers: a) Time taken to complete all of my work during the past 7 days b) Time I would take to complete the same work if I did NOT	1. Binary response: Yes/No 2. Binary response: Yes/No 3a. Number of hours 3b. Number of hours 4. Scale: 0 to 10/. 0 = health had no effect on my work to 10 = I could not do any work at all due to my health.	Absenteeism: 3 months Presenteeism: 7 days	26

		not include any work days you missed to participate in this study.		experience any health problems (***should be less than a)*** 4. In the past 7 days, to what extent was your performance at work affected by YOUR HEALTH which you were working? (*Please think of any physical, mental or emotional problems or symptoms)*			
Composite Work Functioning Approach Boezeman et al 2015	Health	Questions taken from the WLQ		Questions taken from the WLQ		Not stated	
iMTA Productivity Cost Questionnaire (iPCQ) Bouwmans et al 2015	Health Problem	1. Have you missed work in the last 4 weeks? How many days?	1. Yes or No, number of days	1. During the last 4 weeks have there been days in which you worked but during this time were bothered by physical or psychological problems? 2. How many days at work were you bothered by physical or psychological problems? (Only count the days at work in the last 4 weeks) 3. On the days that you were bothered by these problems, was it perhaps difficult to get as much work finished as you normally do? On these days how much work could you do on average?	1. Yes or No 2. Number of work days 3. Scale 0 – 10. 0 = On these days I could not do anything to 10 5 = I was able to do half as much as I do normally 10 = I was able to do just as much as I normally do	Absenteeism: 4 weeks Presenteeism: 4 weeks	12 questions (4 pages)

Appendix 6.2

The ResearchNow Recruitment Process

ResearchNow specialises in the collection of health related data. ResearchNow recruit individuals to their database from which they can be invited to participate in research. Potential participants who join ResearchNow's panel do not receive an incentive when they initially opt-in to the internet panel. ResearchNow profile their participants at the recruitment stage and regularly throughout their time as a panellist. Up to 300 data points are collected to describe each individual included in their panel. The data is used to target potentially relevant participants and invite them to participate in studies that are applicable to them. ResearchNow contact their internet panel through email invitation to complete a questionnaire. The email invitation is a standard format and states the length of the survey, the reward they can earn if they qualify and complete the survey, and the general topic area of the survey. Panellists are told the financial reward for filling out a survey only when they choose to open the email invitation to the study sent by ResearchNow; the financial reward is not mentioned in the subject line of the email. ResearchNow has implemented this 'double opt-in' process to reduce the financial pressure their panellists may feel to participate. ResearchNow pays the participants directly once they have registered their bank account details. The research participants are paid a sum by ResearchNow. The reward varies but even if participants are not eligible for the study, they receive a small fee for considering the study.

Appendix 6.3

Online survey for individuals working with rheumatoid arthritis or ankylosing spondylitis

MANCHESTER 1824
The University of Manchester

Q1

Working with Rheumatoid Arthritis

A survey about how rheumatoid arthritis may affect your ability to work

Please click the arrow below to continue to the survey

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

AboutStudy

About this study

For this survey, commissioned by an external consulting agency, we would like to have a better understanding about health and wellness and for this reason we may ask questions related to your health conditions which will be analysed in aggregate format only.

Be assured that all information shared in this survey is for research purposes only and will be kept confidential.

All data will be processed in adherence to Market Research Society's Code of Conduct and Data Protection Act 1998.

Some of the questions may seem repetitive; however please answer all of the questions included in this survey, they are there for an important reason.

There are no right or wrong answers; we are interested in your experiences.

Would you like to take part in this study?

- Yes, I would like to take part
- No, I do not want to take part

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

QA

Should I complete this survey?

Has your hospital doctor diagnosed you with any of the following conditions?

- Multiple sclerosis
- Constipation
- Rheumatoid Arthritis
- Migraines
- Irritable Bowel Syndrome (IBS)
- Diabetes
- Epilepsy
- Heart Disease
- Kidney Disease
- Other, please specify; _____

QB

Do you work in a job that pays you? *(This can be a full-time or part-time job)*

- Yes
- No

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

PIS

Participant Information Sheet

You are being invited to take part in a research study designed to investigate the effects of working with rheumatoid arthritis.

The research will contribute towards the PhD thesis of Miss Cheryl Jones, at The University of Manchester.

Before you decide, it is important for you to understand why the research is being done and what it will involve.

Please take your time to read through the following information carefully and discuss with others if you wish.

What would I be asked to do if I took part?
You will be asked to fill-out a 10 to 15 minute online questionnaire. The questionnaire will ask you about your rheumatoid arthritis, your occupation, and your ability to work in your job with your condition.

Who will conduct the research?
The research will be conducted by Cheryl Jones, PhD student, who is based in the Manchester Centre for Health Economics at The University of Manchester.

What is the purpose of this research?
The purpose of this research is to develop a method that helps employers to understand and quantify the productive benefits that can be gained from investing in workplace interventions. Investment into effective workplace interventions may help employees with health conditions to continue working.

Why have I been chosen?
You have been chosen because you have been diagnosed with rheumatoid arthritis and are currently working in paid employment.

What happens to the data collected?
Once you have filled in the online questionnaire, your responses will be sent through and stored on a secure-password protected database provided by The University of Manchester.

How is confidentiality maintained?
The survey will not ask you to reveal sensitive information, such as your name, address or postcode. Also, your responses will be assigned with a unique number that ensures your answers remain anonymous.

How do I provide my consent to take part in this study?
You will provide your consent to take part in the study at the beginning of survey.

What happens if I do not want to take part or if I change my mind?
If you decide to take part in the study you are free to withdraw until you finish answering the survey. Once you have reached the end of the survey it will be impossible to withdraw your responses from the study because it will immediately be made anonymous.

Will I be paid for participating in the research?
You will be paid for participating in this research by ResearchNow.

What is the duration of the research?
The questionnaire should take you 10 to 15 minutes to complete.

Where will the research be conducted?
You may complete the online questionnaire wherever you have access to a computer. You may also complete the questionnaire at a time that is most convenient for you.

Will the outcomes of the research be published?
The outcomes of the research will contribute towards the PhD thesis of Cheryl Jones and her supervisory team. Findings will also be submitted for publication in peer-reviewed journals, and for presentation at national and international conferences.

Where will the collected data be stored?
Your responses will be stored on a secure password protected database provided by The University of Manchester.

How and who else may use this data?
The outcomes of this research may be used for future projects, for example applying for a fellowship

grant award. The data may also be shared with internal and external researchers for other projects, however all access to the data will have to be approved by Cheryl Jones.

If there are any issues regarding this research that you would prefer to discuss with a support group or charity please contact the relevant organisation listed at the bottom of this page.

Who is funding this research?
This research is funded by Arthritis Research UK (ARUK) and the Medical Research Council (MRC) grant reference: 20665.

Who has reviewed the research project?
Ethical approval for this study has been granted by The University of Manchester Ethics Committee reference number: 16461

Please click on the arrow to continue to the next step;

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Consent

Consent

Please note that your consent to take part in this study will be recorded once you have started the questionnaire.

Your consent is taken for two reasons; 1) to take part in the study, and 2) to allow us to store and use the data in future research projects.

You are free to withdraw from the survey any time before you finish the survey.

Once you have submitted your answers it will not be possible to withdraw them from the study.

Your answers are anonymous and will be kept strictly confidential. Unlike other surveys you may have answered, the results will only be used by The University of Manchester. No results or data will ever be passed to a private company or organisation.

Please choose an option;

- I would like take part and start the survey
- I do NOT wish to take part

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

gender

Section 1: About You

What is your gender?

- Male
- Female
- Prefer not to say

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Education

Section 1: About You

What is the highest qualification you have?

- University Education and Professional/Vocational equivalents
- A levels, Vocational level 3 and equivalents
- GCSE/O level grade A* - C, Vocational level 2 and equivalents
- Qualifications at level 1 and below
- Other qualifications: level unknown (including foreign qualifications)
- No qualifications
- Prefer not to say

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Age

Section 1: About You

What is your age today?

- 18 to 24 years
- 25 to 29 years
- 30 to 34 years
- 35 to 39 years
- 40 to 44 years
- 45 to 49 years
- 50 to 54 years
- 55 to 59 years
- 60 to 64 years
- 65 to 69 years
- 70 or above
- Prefer not to say

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Ethnicity

Section 1: About You

What is your ethnicity?

- White British/English/Scottish/Welsh/Northern Irish/Irish
- Any other White background
- Mixed: White and Black Caribbean/ White and Black African/ White and Asian
- Any other mixed background
- Asian: Indian/Pakistani/Bangladeshi
- Asian: Chinese
- Any other Asian background
- Black: African/ Caribbean
- Any other Black background
- Arab
- Any other ethnic group
- Other, please specify; _____

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Jobtitle

Section 2: Your Work

What is the title of your main job?
(e.g. primary school teacher, chartered accountant, cashier)

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Sector

Section 2: Your Work

What kind of business, industry or service is your working organisation?
(e.g. construction, education, healthcare, agriculture, manufacturing, public services, government)

- Aerospace, Defense, Security
- Agriculture
- Asset Management
- Automotive
- Banking and Capital Markets
- Business Services
- Capital Projects and Infrastructure
- Charities
- Chemicals
- Communications
- Education
- Emerging Markets
- Engineering and Construction
- Entertainment and Media
- Financial Services
- Forest, Paper & Packaging
- Government and Public Sector
- Healthcare
- Hospitality and Leisure
- Insurance
- Manufacturing
- Metals
- Mining
- Oil and Gas
- Pharmaceuticals and Life Sciences
- Power and Utilities
- Private Equity
- Real Estate
- Retail and Consumer
- Technology
- Transport and Logistics
- Other, please specify; _____

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

JobStatus

Section 2: Your Work

Which of the following best describes your current employment status? *(Tick only one)*

- ◯ Full-time as an employee
- ◯ Part-time as an employee
- ◯ Self-employed
- ◯ Other, please specify; _____

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Dayswork

Section 2: Your Work

On average, how many days do you work per week at this job?

- ◯ 1
- ◯ 1.5
- ◯ 2
- ◯ 2.5
- ◯ 3
- ◯ 3.5
- ◯ 4
- ◯ 4.5
- ◯ 5
- ◯ 5.5
- ◯ 6
- ◯ 6.5
- ◯ 7

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Hourswork

Section 2: Your Work

On average, how many hours do you work per week at this job?

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Income

Section 2: Your Work

What is your average annual income before tax from your main job?
(If you have more than one job, please report only your income earned from your main job)

- Less than £10,000
- £10,000 - £19,999
- £20,000 to £29,999
- £30,000 to £39,999
- £40,000 to £49,000
- £50,000 to £59,999
- £60,000 to £69,999
- £70,000 to £79,999
- £80,0000 and more
- I do not know
- I prefer not to say

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

YearofDiagnosis

Section 3: Your Health

In what year were you diagnosed with rheumatoid arthritis by your hospital doctor?

Please state the year you were diagnosed: _____

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Medication

Section 3: Your Health

What medication are you currently taking for your rheumatoid arthritis?

- ☐ Biologic response modifiers also known as biologics e.g. etanercept (Enbrel®), rituximab (Rituxan®), adalimumab (Humira®)
- ☐ Disease-modifying anti-rheumatic drugs (DMARDs); e.g. methotrexate (Rheumatrex® or Trexall®), sulfasalazine (SSZ®), hydroxychloroquine (Plaquenil®)
- ☐ Oral corticosteroids; e.g. betamethasone (Celestone®), prednisone (Ravos®)
- ☐ Nonsteroidal anti-inflammatory medications (NSAIDs); e.g. naproxen, ibuprofen
- ☐ Other analgesics; e.g. tramadol (Ultram®), oxycodone (OxyContin®)
- ☐ None
- ☐ I don't know

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

MedsEffective

Section 3: Your Health

In your opinion, is your current medication effective?

- Yes, my medication is effective
- My medication is fairly effective
- No, my medication is not effective
- I don't know
- I have no opinion

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

RAPID3

Section 3: Your Health

The next set of questions ask you about your rheumatoid arthritis using a survey called the Routine Assessment of Patient Index Data (RAPID3)

Please tick the best answer for your abilities at this time.

Over the last week, were you able to:

	without **any** difficulty	with **some** difficulty	with **much** difficulty	**unable** to do
Dress yourself, including tying shoelaces and doing buttons?	○	○	○	○
Get in and out of bed?	○	○	○	○
Lift a full cup or glass to your mouth?	○	○	○	○
Walk outdoors on flat ground?	○	○	○	○
Wash and dry your entire body?	○	○	○	○
Bend down to pick up clothing from the floor?	○	○	○	○
Turn regular taps on and off?	○	○	○	○
Get in and out of a car, bus, train, or airplane?	○	○	○	○
Walk two miles or three kilometers, if you wish?	○	○	○	○
Participate in recreational activities and sports as you would like, if you wish?	○	○	○	○
Get a good night's sleep?	○	○	○	○
Deal with feelings of anxiety or being nervous?	○	○	○	○
Deal with feelings of depression or feeling blue?	○	○	○	○

RapidP

How much pain have you had because of your condition and how do you think your health is affecting how you've been **over the past week**?

Please indicate below how severe your pain has been and how you think you are doing:

Rapid2

Considering all the ways in which illness and health conditions may affect you at this time, please

indicate below how you are doing:

0% 100%

Arthritis Research UK

MRC Medical Research Council

centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

EQ5DMob

Section 3: Your Health

The next set of questions ask you about your current health using a survey called the EuroQol Five Dimensions (EQ5D)

Please indicate your health today, taking everything into account.

Under each heading, please tick ONE option that best describes your health TODAY.

Mobility;

- I have **no** problems walking about
- I have **slight** problems in walking about
- I have **moderate** problems in walking about
- I have **severe** problems in walking about
- I am **unable** to walk about

EQ5Dselfc

Self-care;

- I have **no** problems with washing or dressing myself
- I have **slight** problems with washing or dressing myself
- I have **moderate** problems with washing or dressing myself
- I have **severe** problems with washing or dressing myself
- I am **unable** to wash or dress myself

EQ5Dus

Usual Activities (e.g. work, study, housework, family or leisure activities);

- I have **no** problems doing my usual activities
- I have **slight** problems doing my usual activities
- I have **moderate** problems doing my usual activities
- I have **severe** problems doing my usual activities
- I am **unable** to do my usual activities

EQ5Dpain

Pain and/or Discomfort;

- I have **no** pain or discomfort
- I have **slight** pain or discomfort
- I have **moderate** pain or discomfort
- I have **severe** pain or discomfort
- I have **extreme** pain or discomfort

EQ5DAnx

Anxiety and/or Depression;

- I am **not** anxious or depressed
- I am **slightly** anxious or depressed
- I am **moderately** anxious or depressed
- I am **severely** anxious or depressed
- I am **extremely** anxious or depressed

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

EQ5DVAS

Section 3: Your Health

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

State a number from 0 to 100 that indicates how your health is TODAY in the box below

Your Health Today:

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

SF6Dphy

Section 3: Your Health

The next set of questions also asks you about your health using a survey called the Short Form Six Dimensions (SF6D)

There are six groups of statements, each covering a different aspect of health.

Please tick one statement in each group to show the statement which best describes your health over the past 4 weeks.

Physical Functioning;

- ◯ Your health does **not** limit you in **vigorous** activities
- ◯ Your health limits you a **little** in **vigorous** activities
- ◯ Your health limits you a **little** in **moderate** activities
- ◯ Your health limits you **a lot** in **moderate** activities
- ◯ Your health limits you **a little** in bathing and dressing
- ◯ Your health limits you **a lot** in bathing and dressing

SF6Drole

Role Limitations;

- ◯ You have **no** problems with your work or other regular daily activities as a result of your **physical health** or any **emotional problems**
- ◯ You are **limited** in the kind of work or other activities as a result of your **physical health**
- ◯ You **accomplish less** than you would like to as a result of your **emotional problems**
- ◯ You are **limited** in the kind of work or other activities as a result of your **physical health** and **accomplish less** than you would like as a result of your **emotional problems**

SF6Dsocial

Social Functioning;

- ◯ Your health limits your social activities **none** of the time
- ◯ Your health limits your social activities a **little** of the time
- ◯ Your health limits your social activities **some** of the time
- ◯ Your health limits your social activities **most** of the time
- ◯ Your health limits your social activities **all** of the time

SF6Dpain

Pain;

- ◯ You have **no** pain
- ◯ You have pain, but it does **not** interfere with your normal work (both outside the home and housework)
- ◯ You have pain that interferes with your normal work (both outside the home and housework) a **little** bit
- ◯ You have pain that interferes with your normal work (both outside the home and housework) **moderately**
- ◯ You have pain that interferes with your normal work (both outside the home and housework) **quite a bit**
- ◯ You have pain that interferes with your normal work (both outside the home and housework) **extremely**

SF6Dmental

Mental Health;

- You feel tense or downhearted and low **none** of the time
- You feel tense or downhearted and low a **little** of the time
- You feel tense or downhearted and low **some** of the time
- You feel tense or downhearted and low **most** of the time
- You feel tense or downhearted and low **all** of the time

SF6Dvita

Vitality;

- You have a lot of energy **all** of the time
- You have a lot of energy **most** of the time
- You have a lot of energy **some** of the time
- You have a lot of energy a **little** of the time
- You have a lot of energy **none** of the time

0% 100%

Arthritis Research UK

MRC Medical Research Council

centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

ICECAPFeeling

Section 3: Your Health

The next set of questions ask you about your quality of life using a survey called the ICEpop CAPability measure for Adults (ICECAP-A)

About your overall quality of life

Please indicate which statements best describe your overall quality of life at the moment by choosing ONE option for each of the five groups below.

Feeling settled and secure;

- I am able to feel settled and secure in **all** areas of my life
- I am able to feel settled and secure in **many** areas of my life
- I am to feel settled and secure in a **few** areas of my life
- I an **unable** to feel settled and secure in **any** areas of my life

ICECAPALove

Love, friendship and support;

- I can have **a lot** of love, friendship and support
- I can have **quite a lot** of love, friendship and support
- I can have **a little** love, friendship, and support
- I **cannot** have **any** love, friendship and support

ICECAPAIndep

Being Independent;

- I am able to be **completely** independent
- I am able to be independent in **many** things
- I am able to be independent in **a few** things
- I am **unable** to be at all independent

ICECAPAAchieve

Achievement and progress;

- I can achieve and progress in **all** aspects of my life
- I can achieve and progress in **many** aspects of my life
- I can achieve and progress in **a few** aspects of my life
- I **cannot** achieve and progress in **any** aspects of my life

ICECAPEnjoy

Enjoyment and pleasure;

- I can have **a lot** of enjoyment and pleasure
- I can have **quite a lot** of enjoyment and pleasure
- I can have **a little** enjoyment and pleasure
- I **cannot** have **any** enjoyment and pleasure

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

RAaffectwork

Section 4: Your Rheumatoid Arthritis and Work

From your point of view, does your rheumatoid arthritis affect your ability to work?

- Yes
- No

0% 100%

WPAI1

Section 4: Your Rheumatoid Arthritis and Work

The next set of questions ask you about how rheuamtoid arthritis affects your work using a survey called the Work Productivity Activity Impairment(WPAI)

Are you currently working for pay?

- Yes
- No

0% 100%

MANCHESTER 1824
The University of Manchester

Sickleave

Section 4: Your Rheumatoid Arthritis and Work

Are you currently on sick leave?

- Yes
- No

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Datesoffsick

Section 4: Your Rheumatoid Arthritis and Work

On what date did your current sick leave period being?

For example, if your sick leave began on the 5th of June 2016 please type 05/06/2016

0% 100%

Arthritis Research UK | MRC Medical Research Council | centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Timesickleave

Section 4: Your Rheumatoid Arthritis and Work

How many days have you been on sick leave during the past month?

If you have been off sick for half a day please write 0.5

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

WPAI2

Section 4: Your Rheumatoid Arthritis and Work

During the past 7 days, how many hours did you miss from work because of problems associated with your rheumatoid arthritis?

Include hours you missed on sick days, times you went in late, left early, etc., because of your rheumatoid arthritis. Do not include time you missed to participate in this study.

Number of hours:

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

WPAI3

Section 4: Your Rheumatoid Arthritis and Work

During the past 7 days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?

Number of hours:

0% 100%

WPAI4

Section 4: Your Rheumatoid Arthritis and Work

During the past 7 days, how many hours did you actually work?

Number of hours:

0% 100%

MANCHESTER 1824
The University of Manchester

WPAIscale1

Section 4: Your Rheumatoid Arthritis and Work

During the past 7 days, how much did your **rheumatoid arthritis** affect your **productivity** while you were working?

Think about days you were limited in the amount and kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If your rheumatoid arthritis affected your work only a little, choose a low number. Choose a high number if your rheumatoid arthritis affected your work a great deal.

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

WAPIscale2

Section 4: Your Rheumatoid Arthritis and Work

During the past 7 days, how much did your **rheumatoid arthritis** affect your ability to do you regular **daily activities**, other than work at a job?

By regular activities, we mean the usual activities you do, such as work around the house, shopping, childcare, exercising, studying, etc. Think about times you were limited in the amount of kind of activities you could do and times you accomplished less than you would like. If your rheumatoid arthritis affected your activities only a little, choose a low number. Choose a high number if your rheumatoid arthritis affected your activities a great deal.

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Toldemployer

Section 4: Your Rheumatoid Arthritis and Work

Have you told your employer that you have rheumatoid arthritis?

- Yes, and he/she is supportive
- Yes, and he/she is not supportive
- No
- I do not have an employer
- Prefer not to answer

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Disabled

Section 4: Your Rheumatoid Arthritis and Work

For the purposes of your job, do you consider yourself to be disabled?

A person is defined as having a disability for the purposes of discrimination law if he or she has a physical or mental impairment which has a substantial and long-term adverse effect on his or her ability to carry out normal day to day activities.

- Yes, because of my rheumatoid arthritis
- Yes, but because of another condition I have, please specify;
- Yes, because of my rheumatoid arthritis and another condition I have, please specify;
- No, I am not disabled

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

QQ01

Section 4: Your Rheumatoid Arthritis and Work

The next 3 questions are very similar to those asked previously. The questions are about how your rheumatoid arthritis affects your work using a survey called the Quantity and Quality Method (QQ Method)
Please answer all of these questions.

Could you indicate ***how much*** work you performed ***today*** (or the last working day) as compared to normal on the scale below?

0% 100%

QQ01

Section 4: Your Rheumatoid Arthritis and Work

The next 3 questions are very similar to those asked previously. The questions are about how your rheumatoid arthritis affects your work using a survey called the Quantity and Quality Method (QQ Method)
Please answer all of these questions.

Could you indicate ***how much*** work you performed ***today*** (or the last working day) as compared to normal on the scale below?

0% 100%

MANCHESTER 1824
The University of Manchester

QQ02

Section 4: Your Rheumatoid Arthritis and Work

Could you indicate the ***quality*** of the work you performed ***today*** (or the last working day) as compared to normal on the scale below?

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

QQ03

Section 4: Your Rheumatoid Arthritis and Work

Please choose on the scale below the degree of ***efficiency*** you consider yourself to have worked with ***on the days*** you did go to work while suffering from rheumatoid arthritis.

On this scale 10 means your work was not affected and 1 means that you were hardly capable of performing work.

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Feedback01

Section 5: Your Feedback

The next 3 questions allows you to give us your thoughts about working with rheumatoid arthritis and the design of the survey.

Do you have any further comments to make about your rheuamtoid arthritis and work?

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Feedback02

Section 5: Your Feedback

How easy was it to answer the questions in this survey?

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

MANCHESTER 1824
The University of Manchester

Feedback03

Section 5: Your Feedback

Do you have any general comments you would like to make about this survey?

0% 100%

Arthritis Research UK
MRC Medical Research Council
centre for musculoskeletal health & work

Appendix 6.4

Summary of the results of the pilot study: Assessing the potential performance of online survey used to quantify presenteeism using HRQoL measures

A pilot study is essentially a small scale trial run that generates useful preliminary information about the potential performance of the study. The results of a pilot study may indicate potential issues with the current design of the study thereby informing the researcher of the need to make necessary adjustments before launching the full study (van Teijlingen and Hundley 2002). Pilot studies are also used to identify any issues regarding the practicalities of carrying out the research which may include: reducing the likelihood of human error and assessing the technical runnings of the study (van Teijlingen and Hundley 2002).

Aim

The purpose of this pilot study was to assess the potential performance of the survey used to quantify presenteeism using HRQoL. The objectives to meet this aim were to:

1. Test the technical performance of the online survey. This involved assessing the survey's ability to:
    a. excluded potential participants who did not meet the inclusion criteria set for this study;
    b. populate the data collection spreadsheet correctly
2. Assess the responses provided by the sample; did the responses make logical sense?
3. Generate descriptive statistics about the sample's background, job, disease severity, health status, and levels of presenteeism
4. Explore the relationship between three health related quality of life (HRQoL) measures and two presenteeism measures

Methods

The pilot study was conducted on the 20th April 2017. The survey was uploaded online using Sawtooth Lighthouse Studio 9.3.1 software [Sawtooth Software 2017]. Links to the survey were sent to ResearchNow and email invitations to potential participants who may have been interested in taking part in the study were sent.

Sample

The aim of the pilot study was to collect a sample of 35 employed individuals (full-time or part-time) working with rheumatoid arthritis (RA). Initially, the full study aimed to recruit a sample size of 350 employed individuals working with RA; therefore, as recommended by ResearchNow, 10% of the full study sample size was recruited. (Since this analysis was undertaken, the sample size for the full study increased to 550).

Analysis

The data collected by the survey was cleaned and analysed using STATA (v.12). The survey was assessed for its ability to reject individuals who did not meet the inclusion criteria which was done using two screening questions: 1) Has your hospital doctor diagnosed you with any of the following conditions?; 2) Do you work in a job that pays you? (This can be a full-time or part-time job). Only those individuals who confirmed they have RA and work in a paid job were accepted into the survey.

The software's function that allowed questions in the survey to be skipped where necessary to maintain a logical flow was also checked. Importantly, the data collection spreadsheet housed within Sawtooth's software was also checked to ensure that data recorded by the survey was populated and uploaded correctly. The process for exporting data from Sawtooth to a Microsoft Excel spreadsheet was also tested; this was an important step to ensure the data could be uploaded into STATA.

Descriptive statistics were generated after the data was cleaned using STATA. The relationship between the HRQoL and presenteeism was explored using scatterplots and Spearman's Rank correlation coefficients.

Results

The following section reports the performance of the survey, descriptive statistics of the sample and the relationship between HRQoL and presenteeism. The results were used to inform the analysis of the main study.

Technical performance

The online survey was assessed for its ability to exclude potential participants who did not meet the inclusion criteria. ResearchNow invited 123 individuals to take part in the study; of the 123 individuals 36 met the inclusion criteria assessed by the screening questions. Two individuals, whilst they met the inclusion criteria, were excluded because they did not provide their consent to take part in the study. The results confirmed the survey was able to successfully exclude potential participants who were

not eligible to take part. For individuals who did not meet the inclusion criteria the survey re-directed them to a webpage that thanked those for their interest in the study but that they were not eligible to complete this study.

The data collected from the online survey was correctly and accurately populated and uploaded into Sawtooth's spreadsheet. The data was also successfully transferred from Sawtooth into a Microsoft excel sheet. The time it took for the 32 individuals to complete the survey was approximately 11 minutes.

Data Cleaning

The data collected from the survey was transferred from Sawtooth to Microsoft Excel and then uploaded into STATA to being data cleaning. Two further observsations were dropped from the sample because they reported inconsistent answers. First, the Work Productivity Activity Impairment Questionnaire (WPAI) asks the following question: 'Are you currently working for pay?' to which the individual answered 'no; however, the same observation also answered 'yes' to the screening question 'Do you work in a job that pays you?'. The second observation was dropped because they reported ten for the WPAI and ten for the QQ method; a logically impossible result.

Descriptive Statistics

In total, the pilot sample consisted of 32 individuals working with RA. Table one summarises the characteristics of the sample including their gender, age, ethnicity and level of education. Females represented 52% of the sample which was lower than would be expected because females are more likely to develop RA at a ratio of 3:1 (D. Symmons et al. 2002). The greatest proportion of individuals was aged between forty and sixty; an expected pattern given the inclusion criteria into this study was individuals who have a paid job and therefore not retired. Ethnic minority groups were poorly represented with 96.8% of the sample being White British. Education levels including GCSEs, A levels, and degrees were, relatively, equally represented.

Table 1 Summary Characteristics of Pilot Sample

Characteristics	**n= 32**	**%**
Gender (female)	16	51.6
Age		
25 - 29	1	3.2
30 -34	2	6.5
35 - 39	4	12.9
40-44	6	19.4
45-49	4	12.9
50 -54	1	3.2
55-59	8	25.8
60-64	1	3.2
65-69	3	9.7
70+	1	3.2
Ethnicity		
White British	30	96.8
Asian: Indian	1	3.2
Education		
Degree	9	29.03
A levels	10	32.3
GCSE	11	35.5
Level 1	1	3.2

Table two presents job characteristics of the individuals included in the sample. The sample consisted of 6.5% full-time employees, 37.5% part-time and zero self-employed. The majority of the sample worked in a non-manual job. Eighty-five percent of the sample reported their income levels up to £40,000.

Table 2 Job Characteristics

Job Characteristics	Total Sample	
	N	**%**
Full-time employee	29	61.3
Hours per week	40.8	-
Non-manual	13	68.4
Manual	6	31.6
Part-time employee	12	38.7
Hours per week	19.7	-
Non-Manual	9	75.0
Manual	3	25.0
Self-employed	0	
Hours per week		
Non-manual		
Manual		
Income		
Less than £10,000	5	16.1
£10,000 - £19,999	8	25.8
£20,000 - £29,999	7	22.6
£30,000 - £39,000	6	19.4
£40,000 - £49,000	1	3.2
£50,000 - £59,999	1	3.2
£60,000 - £69,999	0	0.0
£70,000 to £79,999	2	6.5
Prefer not to say	1	3.2

Table three presents information about the sample's health. Seventy-five percent of the sample was classified by the Routine Assessment of Patient Index Data 3 (RAPID3) as having a high level of disease severity. Over 50% of the sample had RA for nine years or less; an expected trend for two reasons: 1) the peak age of incidence of RA for women is at the age of 45 years (Humphreys et al. 2013); and 2) only individuals who currently work were included in the sample. The number of individuals using biologics or disease-modifying anti-rheumatic drugs (csDMARDs) was far lower than would have been expected given NICE guidelines to treat RA early using aggressive medication (NICE 2009).

Table 3 Health Characteristics

Health	**Total Sample**	
	N	**%**
Disease Severity (RAPID)		
High	23	74.2
Medium	8	25.8
Low/Remission	0	0.0
Years with RA or AS (years)		
> 5	11	35.5
5 – 9	8	25.8
10 – 14	1	3.2
15 – 19	2	6.5
20+	9	29.0
Medication		
Biologics		
Biologics only	3	9.7
csDMARDs only	2	6.5
Biologics and csDMARDS	6	1.9
	1	3.2
csDMARDs		
Oral Corticosteroids	7	22.6
NSAIDs	6	19.4
Other Analgesics	15	48.4
None	10	32.3
	2	6.5
Number of Medications		
0	2	6.3
1	21	65.6
2	6	18.8
3	3	9.4

RAPID3: Routine Assessment of Patient Index Data Three; csDMARDS: Disease Modifying Anti-Rheumatic Drugs; RA: Rheumatoid Arthritis; AS: Ankylosing Spondylitis

Health Status, Capability, and Presenteeism

Table four shows the mean index utility scores for the EuroQol Five Dimensions (EQ5D-5L) (Gudex 2005), Short-Form Six Dimensions (SF6D) (Brazier, Roberts, and Deverill 2002)and ICEpop CAPability measure for Adults (ICECAP-A) (Al-Janabi et al. 2013). The SF6D reported the highest level of HRQoL. The EQ5D-5L and SF6D did not report full health (1) levels.

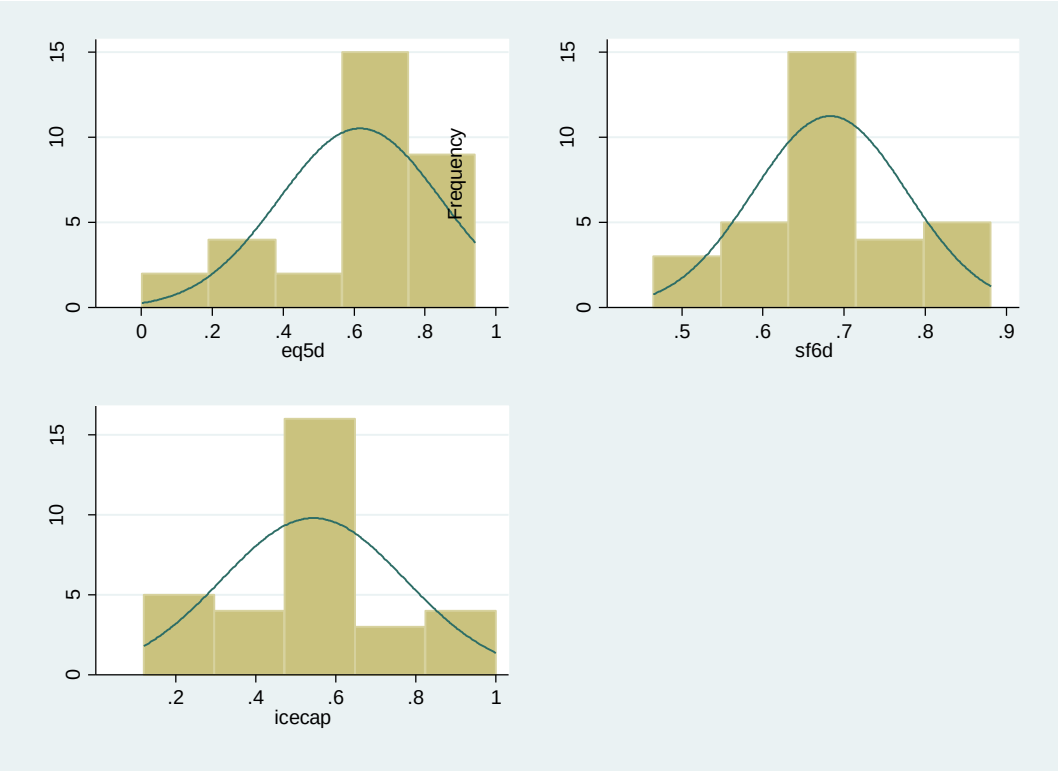
Table 4 Health Status and Capability mean utility scores

Measure	Mean	Min.	Max.
EQ5D-5L	0.627	0.002	0.942
SF6D	0.676	0.465	0.810
ICECAP-A	0.544	0.12	1.000

*Categories of the strength of the correlations: very weak (0 to 0.19); weak (0.2 to 0.39); moderate (0.40 to 0.59); strong (0.6 to 0.79); and very strong (0.8 to 1). EQ5D-5L: EuroQol Five Dimensions Five Level; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop CAPability measure for Adults

Figure one shows the distribution of the EQ5D-5L, SF6D and ICECAP-A. The EQ5D-5L and was skewed to the right and the SF6D and ICECAP-A was approximately normal.

Figure 1 Distribution of health related quality of life measures



EQ5D-5L: EuroQol Five Dimensions Five Level; SF6D: Short Form Six Dimensions; ICECAP-A: ICEpop CAPability measure for Adults

Presenteeism

Table five shows the mean presenteeism scores measured by the WPAI and the QQ method. The mean scores for presenteeism reported by the two measures were approximately equivalent.

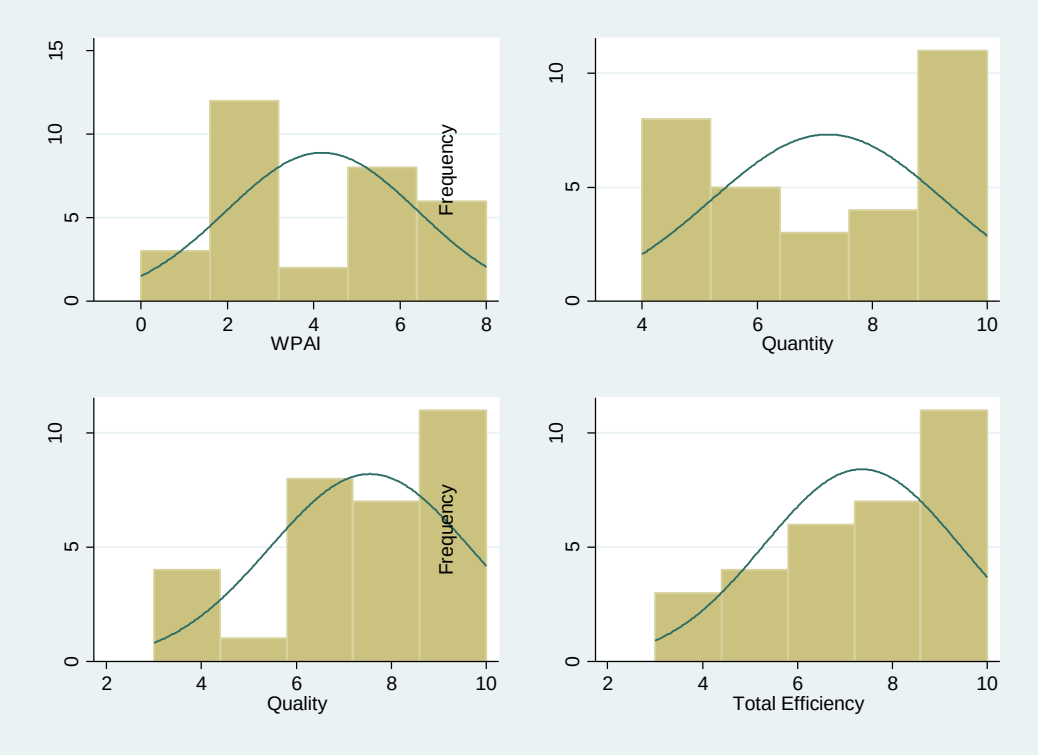
Table 5 Presenteeism

Presenteeism	Mean	Min	Max
WPAI	4.4	0	8
Quantity (QQ Method)	7.3	4	10
Quality (QQ Method)	7.6	3	10
Total Efficiency (QQ Method)	7.4	3	10

WPAI: Work Productivity Activity Impairment Index; QQ Method; Quantity and Quality Method

Figure two presents histograms of the distrbutions of the WPAI and three QQ method meausres of presenteeism.

Figure 2 The distributions of presenteeism



WPAI: Work Productivity Activity Impairment Index; Quantity: measure of quantity of work completed by the Quantity and Quality Method (QQ method); Quality: measure of quality of work completed by the QQ method: Total Efficiency: measure of efficiency of work completed by the QQ method

The Relationship between HRQoL and Presenteeism

The following scatterplots in figures three, four, five and six show the relationship between HRQoL and presenteeism. The relationship appears to be stronger for the WPAI and HRQoL measures compared with the QQ method. However, the direction of the relationship for each HRQoL and presenteeism measure was as expected.

Figure 3 The relationship between health related quality of life and presenteeism measured by the Work Productivity Activity Impairment questionnaire

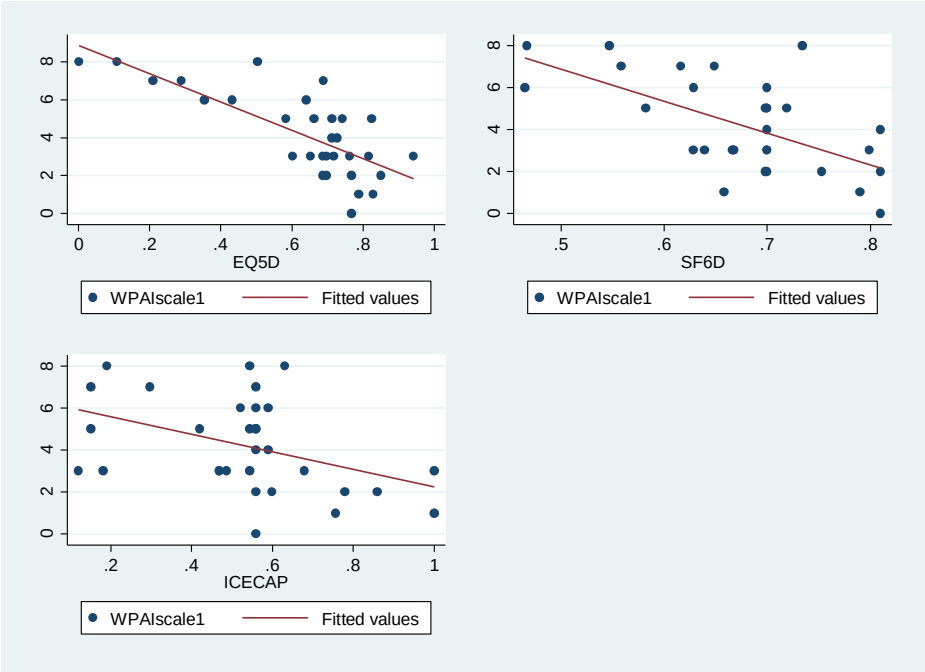


Figure 4 The relationship between health related quality of life and the quantity of work completed measured by the Quantity and Quality method

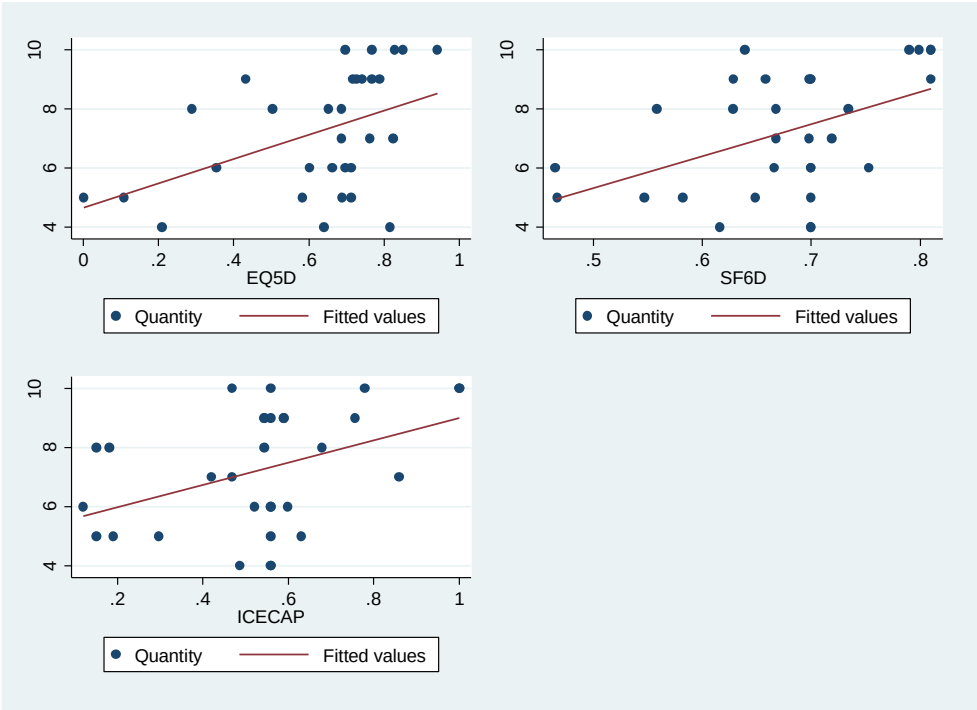


Figure 5 The relationship between health related quality of life and the quality of work completed measured by the Quantity and Quality method

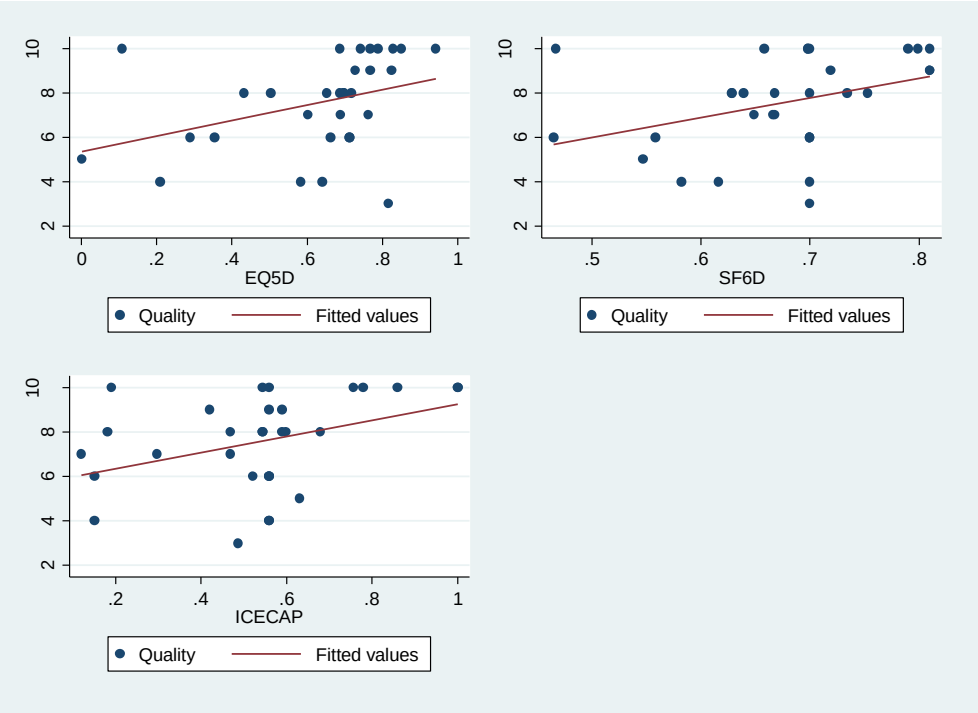


Figure 6 The relationship between health related quality of life and efficiency of work completed measured by the Quantity and Quality method

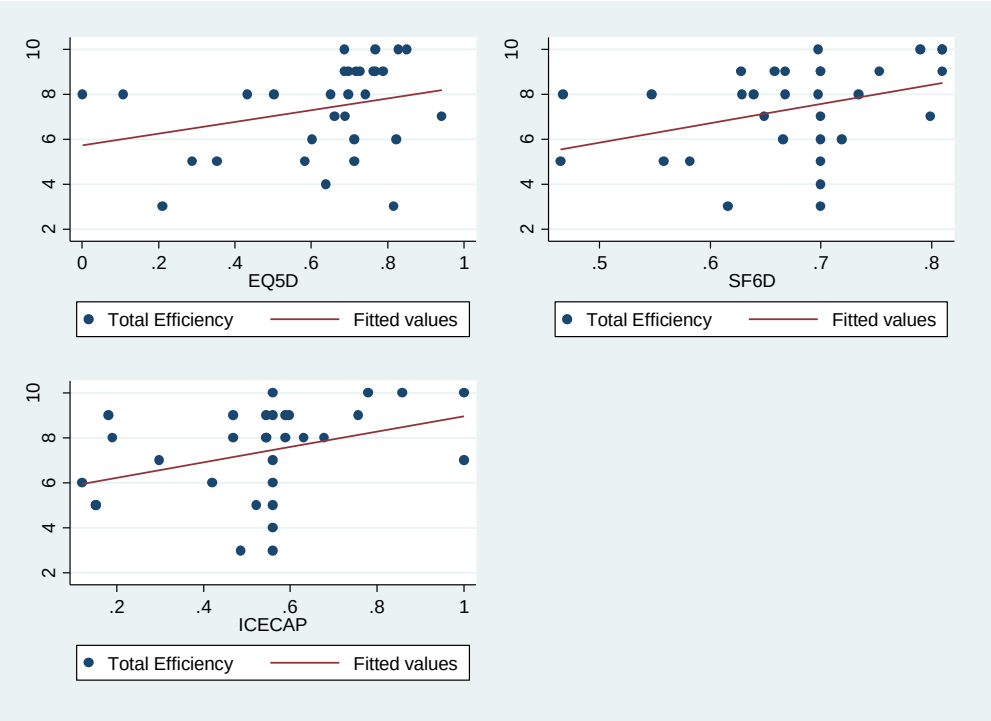


Table six shows Spearman's Rank correlation scores for HRQoL and presenteeism. The correlations for the WPAI ranged from moderate to strong compared with the QQ method which ranged from weak to moderate.

Table 6 HRQoL and Presenteeism Spearman's Rank Correlations

Health Status	Presenteeism			
	WPAI	Quantity	Quality	Total Efficiency
EQ5D	-0.7176	0.5323	0.5080	0.3995
SF6D	-0.5107	0.4218	0.3833	0.3720
ICECAP-A	-0.4041	0.3579	0.3924	0.4431

*Categories of the strength of the correlations: very weak (0 to 0.19); weak (0.2 to 0.39); moderate (0.40 to 0.59); strong (0.6 to 0.79); and very strong (0.8 to 1). WPAI: Work Productivity Activity Impairment Index; Quantity: measure of quantity of work completed by the Quantity and Quality Method (QQ method); Quality: measure of quality of work completed by the QQ method: Total Efficiency: measure of efficiency of work completed by the QQ method

Discussion

The results of the pilot analysis confirmed that the survey flowed in a logical manner, excluded individuals who did not meet the inclusion criteria and skipped questions where necessary. The average time taken to complete the survey was within the 15 minutes initially set. The survey also populated and uploaded data into Sawtooth's spreadsheet correctly and it was possible to simply transfer the data into Microsoft excel.

The pilot study highlighted some potential issues that needed to be addressed before launching the full study. First, the responses to the two questions about working for pay generated potentially confusing answers. One participant confirmed they worked for a paid job and also that they do not *currently* work for pay. It was unclear what the participant was trying to convey; was the participant on sick-leave but still employed in a paid job therefore fitting the criteria of being in paid employment but not *currently* working or did they not currently work in a paid job. As a result of this confusion, the questionnaire was altered to include the following three questions if the response to 'Are you currently working for pay?' (WPAI question) was answered 'no';

1. Are you on sick leave? (Yes/No)
2. When did sick leave begin? (Date)
3. How many days on sick leave? (Number of days)

The descriptive statistics of the sample's characteristics, job and health were largely as expected. A slight cause for concern was the low number of individuals who reported being on biologics or csDMARDs. NICE, as mentioned earlier, recommend treating RA as early as possible using csDMARDs, or, if they fail, biologics. This may suggest that the individuals answering the survey do not have RA but may in fact have Osteoarthritis

(OA), a disease that is not auto-immune and does not require treatment using csDMARDs or biologics. As a result of this finding, ResearchNow was contacted to confirm how they identify and categorise individuals on their database. ResearchNow reported that the respondents are put through an extensive set of screening questions to help identify and categorise them so surveys can be targeted to them. Specifically, individuals with health conditions are presented with a list of diseases they may have; RA and OA are listed separately to help avoid the chance of individuals self-reporting the wrong health condition.

The distributions of the measures of presenteeism are difficult to interpret and as a result it is difficult to state the types of models that could be used to predict presenteeism using HRQoL data. The distribution of the WPAI suggest that individuals have reported five for their level of presenteeism when potentially some may have meant four; this can be seen by the large number of individuals reporting five compared to the few that reported four. This phenomenon may arise because of a psychological tendency to go for the mid-point when using Likert scales (Dillman 2006). The quality and total efficiency distributions of the QQ method show a positive distribution; however, the quality measure also seems to be suffering from over reporting of five.

The relationship between HRQoL and presenteeism measured by the WPAI was stronger than the relationships between HRQOL and presenteeism measured using the QQ method. The correlations, especially, those for the WPAI and HRQoL, are strong enough to enable the development of a prediction model.

Conclusion

The pilot study confirmed the online survey and the technology used to record and transfer the data worked correctly. The pilot study highlighted the potential for confusing responses from two questions regarding whether the individual currently works for pay. Three questions were added to the survey to accommodate and rectify the confusion. The descriptive statistics were largely in line with expectations with the exception of the number of individuals taking biologics or csDMARDs. ResearchNow was contacted to explain more about how they identify and categorise individuals to complete health-related surveys. The correlations between HRQoL and presenteeism were strong indicating the potential to quantify the link between HRQOL and presenteeism.

References

Brazier, J, J. Roberts, and M. Deverill. 2002. 'The Estimation of a Preference-Based Measure of Health from the SF-36'. *Journal of Health Economics* 21 (2):271–92.

Dillman, D. A. 2006. *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Wiley.

Gudex, C.. 2005. 'The Descriptive System of the EuroQol Instrument'. In *EQ-5D Concepts and Methods: A Developmental History*, edited by Paul Kind, Richard Brooks, and Rosalind Rabin, 19–27. Springer Netherlands. http://link.springer.com/chapter/10.1007/1-4020-3712-0_2.

Humphreys, J. H., S. M. M. Verstappen, Kimme L. Hyrich, Jacqueline R. Chipping, Tarnya Marshall, and Deborah P. M. Symmons. 2013. 'The Incidence of Rheumatoid Arthritis in the UK: Comparisons Using the 2010 ACR/EULAR Classification Criteria and the 1987 ACR Classification Criteria. Results from the Norfolk Arthritis Register'. *Annals of the Rheumatic Diseases* 72 (8):1315–20. https://doi.org/10.1136/annrheumdis-2012-201960.

Al-Janabi, H., T. J. Peters, J. Brazier, S. Bryan, T. N. Flynn, S. Clemens, A. Moody, and J. Coast. 2013. 'An Investigation of the Construct Validity of the ICECAP-A Capability Measure'. *Quality of Life Research* 22 (7):1831–40. https://doi.org/10.1007/s11136-012-0293-5.

NICE. 2009. 'Rheumatoid Arthritis'. 2009. https://www.nice.org.uk/guidance/cg79.

Symmons, D., G. Turner, R. Webb, P. Asten, E. Barrett, M. Lunt, D. Scott, and A. Silman. 2002. 'The Prevalence of Rheumatoid Arthritis in the United Kingdom: New Estimates for a New Century'. *Rheumatology* 41 (7):793–800. https://doi.org/10.1093/rheumatology/41.7.793.

Teijlingen, E.van, and V. Hundley. 2002. 'The Importance of Pilot Studies'. *Nursing Standard (Royal College of Nursing (Great Britain): 1987)* 16 (40):33–36. https://doi.org/10.7748/ns2002.06.16.40.33.c3214.

Appendix 7.1

Publication strategy

Chapter	Title of study or focus of study	Target Journals	Progress
Chaper 3	Title: Economic Theory and Self-Reported Measures of Presenteeism in Musculoskeletal Disease	Currentt Rheumatology Reports (see appendix nineteen)	Published
Chapter 4	Title: To examine the extent to which the impact of productivity is explicity and consistently reported in economic evaluations of workplace interventions for musculoskeletal disorders: A systematic review	Applied Health Economics and Health Policy, Value in Health	Pending submission
Chapter 5	Focus: Exploring the extent to which HRQoL measures are able to capture presenteeism caused by rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis: A qualitative study	Social Science and Medicine, Value in Health, Quality of Life Research	Submission planned for 2018
Chapter 5	Focus: Conceptual overlap between three HRQoL measures	Social Science and Medicine, Value in Health, Quality of Life Research	Submission planned for 2018
Chapter 6	Focus: Development of a prediction model for presenteeism caused by rheumatoid arthritis, ankylosing spondylitis or psoriatic arthritis	Medical Decision Making, Value in Health, European Journal of Health Economics	Submission planned for 2018

Appendix 7.2

Plans for dissemination of research

The table below lists plans of dissemination of the results of the thesis presented.

Chapter	Aim	Target Audiences	Forum
Chapter six	Poster presentation to discuss the methods and results of this chapter	Health Economists	Health Economics Study Group conference (HESG), City University London 10th January
Thesis	Oral Presentation of the results from the whole thesis	Rheumatologists and epidemiologists working within the area of musculoskeletal conditions and their impact on an individual's ability to work	Scientific meeting Centre for Musculoskeletal Health and Work annual scientific meeting held on the 13th February 2018 in Southampton
Chapter five	Abstract	Rheumatologists and epidemiologists working in the field of rheumatology; not necessarily working in the area of rheumatology and work	British Society for Rheumatology annual conference held in Liverpool on the 1st May 2018