THE ECONOMICS OF BACK PAIN: ALTERNATIVE APPROACHES TO PRODUCTIVITY COST ESTIMATION IN ECONOMIC EVALUATION OF HEALTHCARE

By

JESSE B.L KIGOZI

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Health Economics Unit

School of Health and Population Sciences

University of Birmingham

February 2014

ABSTRACT

The thesis investigates approaches to estimating productivity costs in economic evaluation, focusing on application of the friction cost approach (FCA) using low back pain as a case study. Individual studies of validating a Single-Item Presenteeism Question (SIPQ), estimation of the friction period by occupation in the United Kingdom (UK), and comparison of sickness certification records with self-reported data are reported. Further, the thesis explores the impact of the approaches on cost-effectiveness estimates using the FCA.

Results suggest SIPQ is a valid and responsive measure, and that self-reported data provides more complete data than sickness certification records. Stratified friction periods estimates were used in the FCA approach to generate absenteeism costs. This is the first time the FCA is used in a cost effectiveness study to report productivity costs, presenteeism, compensation mechanisms and multiplier effects in UK. In this case study, consideration of full productivity costs and applying detailed friction periods did not alter interpretation of the cost-effectiveness estimates.

Further testing of these approaches in the UK is required, considering growing evidence and merit for accurate estimates of productivity costs. Overall, the research contributes to methods for assessing productivity in economic evaluation, and further illustrates the feasibility of using them in the UK.

DEDICATION

To God be the glory

ACKNOWLEDGEMENTS

This thesis has benefited significantly from the support and guidance of my supervisors, Pelham Barton, Joanna Coast, Sue Jowett, and Martyn Lewis. I deeply appreciate their commitment, leadership, insightful guidance and technical expertise throughout this doctoral research process. I am particularly grateful for the time they dedicated to reading, advising, challenging and discussing this work throughout the period of my research.

The work presented in the thesis was supported by the Arthritis Research UK Primary Care Centre Doctoral Research Award as part of an NIHR programme of work funded through Keele University. I would like to acknowledge the funders and colleagues associated with the Keele University primary care and health sciences who provided important support during the PhD: Particular thanks go to Martyn Lewis and David Whitehurst.

I would also like to thank my colleagues in the Health Economics Unit, for their support. To my colleagues and friends, Lazaros, Raymond, Kennedy, Anita, Lillian and others with whom I discussed aspects of my PhD and who commented on chapters, thank you.

Many thanks to the brethren and special friends at the Deeper Life Bible Church, and in the Deeper Life Campus Fellowship, thank you all for your constant encouragement, support and prayers through this challenging PhD experience. A special thank you to Mum, Dad, and all my siblings, I love you all and I know you all can’t wait for this to end so you can have more of me!

Finally, all I am, and all I have, I give the glory to the most High God. Thank You for Your kindness and abundant love for me.

DECLARATION

I declare that the work presented in this thesis was carried out in accordance with the requirements of the University of Birmingham. The work is original, except where indicated in the write-up. No part of this thesis has been submitted for any other academic award.

SIGNED: Jesse Kigozi DATE:

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFP	Adjusted friction periods
ALWQ	Angina-related limitations at work questionnaire
ASHE	Annual Survey of Hours and Earnings
CBA	Cost-benefit analysis
CBT	Cognitive behavioural therapy
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CIPD	Chartered Institute of personnel and development
COI	Cost of illness
COPD	Chronic obstructive pulmonary disorder
CPI	Consumer price index
CUA	Cost-utility analysis
DLA	DLA piper LLP
EHCA	Employer Health and Coalition of Tampa assessment instrument
EQ-5D	EuroQol
ES	Effect size
EWPS	Endicott Work Productivity scale
FCA	Friction cost approach
GRI	Guyatt responsiveness index
HAD	Hospital Anxiety and Depression scale
HAQ	Health and Assessment Questionnaire
HCM	Human capital method
HLQ	Health and Labour Questionnaire
HPQ	Health and Work Performance Questionnaire
HRPQ-D	Health Related Productivity Questionnaire Diary
IBS	Impact Back Study
ICER	Incremental cost-effectiveness ratio
IMPaCT Back	Implementation study to improve Patient Care through Targeted treatment for back pain
IQWiG	Institute for Quality and Efficiency in Healthcare
LBP	Low back pain
MIDAS	Migraine Disability Assessment Questionnaire
MWPLQ	Migraine work and productivity loss questionnaire
NHS	National Health Service
NHS EED	National Health Service economic evaluation database
NICE	National Institute for Health and Care Excellence (NICE)
NSAIDS	Non-steroidal anti-inflammatory drug

NSLBP	Non-specific low back pain
NS-SEC	National statistics socio-economic classification
ONS	Office of national statistics
OST	Osterhaus technique
PCS	Pain and Catastrophising scale
QALYs	Quality-Adjusted Life years
QQ	Quantity and Quality questionnaire
RMDQ	Roland Morris Disability questionnaire
SBT	STarT Back Trial
SD	Standard Deviation
SF-12 MCS	SF-12 mental health component summary measure
SF-12 PCS	SF-12 physical component summary measure
SIC	Standard industrial classification
SIPQ	Single-item Presenteeism Question
SOC	Standard occupational classification
SPS	Stanford Presenteeism Scale
SRM	Standardised response mean
STarT Back	Subgrouping for Targeted Treatment
TSK	Tampa scale for Kinesiophobia
UFP	Unadjusted friction periods
UoB	University of Birmingham
WALS	Work activity limitations scale
WHI	Work and health interview questionnaire
WIS	Work Instability Scale
WLQ	Work limitations questionnaire
WL-WRF	Work Role and Functioning scale
WPA	US Washington panel approach
WPAI	Work Productivity and Activity Impairment Questionnaire
WPAI-GH	Work Productivity and Activity Impairment Questionnaire General Health
WPSI	Work productivity short inventory
WTP	Willingness-to-pay

CHAPTER ONE INTRODUCTION

1.1 Background

The increasing demands on limited healthcare resources have raised questions about how to make resource allocation decisions. Economic evaluations are often used as a means of guiding decision makers in the efficient allocation of scarce healthcare resources. In economic evaluation research, the costs and benefits of an intervention are compared. They are used to determine the relative efficiency of new interventions and to determine whether to fund a new intervention or not (Drummond et al., 2005). Application of these analyses, it is argued, could potentially result in cost containment, healthcare demand management and maximisation of benefits from healthcare spending (Morris et al., 2007).

The role of economic evaluation has become increasingly important in recent years. While the impact of these evaluations on final decision making may be uncertain (Buxton, 2006; Williams et al., 2008), a recent review of national pharmaco-economic guidelines shows that most countries now use economic evaluation in funding decisions of pharmaceutical interventions (Knies et al., 2010). For instance, in the UK, since 1999 the National Institute for Clinical excellence, now known as the National Institute for Health and Care Excellence (NICE), has produced national level guidance on effectiveness and cost-effectiveness of devices and interventions for the National Health Service (NICE, 2008; NICE, 2013). One of the main issues of controversy in most economic guidelines is the normative question of which perspective should be considered in such analyses (Drummond and Rutten, 2008). The study perspective is the viewpoint from which an economic evaluation is conducted (Drummond et al., 2005). The choice of perspective is important as it determines the costs and benefits to be considered in an evaluation. Most national guidelines recommend performing economic evaluation from either (i) a narrow perspective, often that of a health service (payer), incorporating only costs falling within the healthcare budget, or (ii) the societal perspective, which includes all relevant costs and benefits (Knies et al., 2010; Pritchard and Sculpher, 2000).

In the UK context, economic evaluation research is often conducted using an extra-welfarist approach with an associated NHS (payer) perspective for assessing resource use (Brouwer et al., 2008; NICE, 2013). Elsewhere, however, others have advocated for a societal perspective (Gold et al., 1996; Brouwer et al., 1997a). Arguments in favour of this perspective are related to the basic principles of economic evaluation founded in welfare economics (Byford and Raftery, 1998; Johannesson, 2009). In addition, some argue that adopting a narrow perspective would be to deny the reality of costs falling outside the healthcare budget which could lead to biased health policies for society (Drummond et al., 2005). Adopting a societal perspective implies that all relevant costs and effects should be included, irrespective of who bears the costs and receives the benefits (Gold et al., 1996).

If the principle of a societal perspective is adopted, one of the main areas of controversy is the estimation of productivity costs of illness. Productivity costs are defined as 'Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid' (Brouwer et al., 1997a) (p 254). Broadly, in the context of paid labour, productivity costs can emerge as a result of absenteeism and presenteeism. Absenteeism is when an individual is away from work due to illness, disability (morbidity) or premature death (mortality). Productivity losses can also emerge as a result of reduced productivity at work (presenteeism) due to disease which often precedes or follows absenteeism (Brouwer et al., 2002).

There have been various methodological debates on the estimation of productivity costs, including whether, and how, to incorporate such costs in economic evaluation studies (Brouwer et al., 1997a; Koopmanschap et al., 1997; Koopmanschap et al., 1995; Liljas, 1998; Brouwer and Koopmanschap, 2005; Johannesson and Karlsson, 1997). In addition, there are several on-going debates around additional work related situations such as compensation mechanisms and multiplier affects that can either reduce or increase productivity losses related to disease. Not all absenteeism or presenteeism results in reduced production as internal compensation within an organisation can potentially reduce productivity (Jacob-Tacken et al., 2005; Severens et al., 1998; Krol et al., 2012). Additionally, there can be productivity losses resulting from negative effects of employee sickness on team productivity, output and substitutability (Nicholson et al., 2006; Zhang et al., 2012; Krol et al., 2012; Pauly et al., 2008). Such costs are often termed as 'multiplier effects'. The multiplier is defined as, "*the cost of an absence as a proportion (often greater than one) of the absent worker's daily wage*", (Nicholson et al., 2006) (p 112). Little is known about how to incorporate these costs in economic evaluations and the overall impact of including compensation costs and multiplier effects in economic evaluation studies.

The significant costs associated with productivity loss increase the need to ensure rigorous methodological developments in estimation of these costs. Furthermore, the lack of consensus regarding the most appropriate method for valuing both absenteeism and presenteeism costs necessitates further developmental work in this area (Pritchard and Sculpher, 2000; Knies et al., 2010). In the literature, three distinct approaches to valuation of productivity costs have been considered. These are the human capital approach, the friction cost approach, and the US panel Washington approach. The human capital approach is the traditionally used method, with productivity benefits based on the present value of lost time due to illness, disability or death (Weisbrod, 1961; Sculpher, 2001). This method has been criticised for generating unrealistic estimates of productivity loss (Koopmanschap and van Ineveld, 1992). The friction cost approach, was developed by health economists from the Netherlands as a result of perceived limitations and unrealistic theoretical assumptions of the human capital approach. The approach limits productivity loss to the friction period – time required to restore production to original levels (Koopmanschap et al., 1995; Koopmanschap and van Ineveld, 1992). A proxy to this period is the time it takes to replace and train a worker either from the labour market or from the unemployed ranks (Koopmanschap et al., 1997). Friction costs then broadly encompass lost production during the friction period and the costs of hiring and training new individuals (Koopmanschap and Rutten, 1996). In an alternative consideration of productivity loss, the US panel approach recommended valuation of productivity costs through a generic preference measure such as the Quality Adjusted Life Years (QALY) (Gold et al., 1996).

Much of the empirical research in relation to these three approaches has focussed on the US panel approach. The theory in relation to the friction cost approach does not appear to have been translated into routine practice in economic evaluation studies. Little is currently known about whether and how this method is applied in different country contexts. In addition, prospective research is needed to better understand components of productivity loss when using the friction cost approach in chronic episodic diseases such as low back pain. The application of the FCA in retrospective studies of chronic illnesses may provide a foundation for alternative ways of assessing the frequency, length of friction periods and friction costs in other disease areas and in future investigations. There is therefore merit for further exploration and developments on the practical application of the FCA.

This thesis considers methodological issues relating to alternative approaches to valuation and inclusion of productivity costs in economic evaluation in the United Kingdom (UK), with a particular focus on the friction cost approach and applications in the clinical area of low back pain. Low back pain is used as a case study in this methodological research as productivity costs are particularly relevant for this area. Economic evaluation of stratified management for care of patients with low back pain is used to explore issues relevant to productivity costs and economic evaluation research.

1.2 The case of low back pain

As a vehicle for exploring the feasibility and implications of the alternative methodological approaches developed in this thesis, the stratified management care intervention for low back pain sufferers is used as a case-study. This section briefly summarises the burden and importance of the clinical area of low back pain with a comprehensive overview provided in Chapter Five of this thesis.

Low back pain is a health condition associated with significant burden on primary and secondary healthcare services in the UK. It is a very common disorder and nearly five million people develop back pain each year in the UK (Hay et al., 2008). Additionally, low back pain will affect 80-85% of people over their life-time (Hoy et al., 2010a). The condition mainly affects individuals of working age leading to work absence and disability which are associated with significant economic costs to society (Ostelo et al., 2008). Low back pain has been found to be very costly in the UK, with estimated direct healthcare costs of £1.6 billion and indirect costs totalling £10.7 billion in 1998 (Maniadakis and Gray, 2000). It has been estimated that treating the different types of back pain costs the NHS over £1 billion per year (NICE, 2009b). Although indirect costs comprise the majority of the economic burden of low back pain, there is still uncertainty about the methodological approaches used to estimate these costs.

1.3 Productivity costs issues addressed in this thesis

The main questions addressed in this thesis include controversies and unresolved issues regarding the measurement and inclusion of productivity costs of absenteeism and presenteeism within a UK setting. The overall aim of the thesis is to contribute to the theoretical and methodological literature involved in the practical valuation of productivity costs using the friction cost approach and related to the inclusion of absence from work (absenteeism), reduced productivity (presenteeism) and wage-related multiplier effects in economic evaluation.

These questions are:

1. How common is the application of the friction cost approach in valuing productivity costs in UK economic evaluation studies?
2. How has presenteeism been measured and applied in economic evaluation research?
3. What is the appropriate length of a friction period for the UK?
4. Can sickness certification records be used as a proxy for self-reported work absence data?
5. Is the Single-Item Presenteeism Question (SIPQ) a valid and responsive measure among low back pain patients?
6. How does the inclusion of productivity costs using the alternative valuation methods explored in this thesis impact on the results of the cost-effectiveness analyses?

Economic analysis of stratified primary care management in patients with low back pain is used to explore the methodological and empirical issues in this thesis. Using data from the Subgrouping for Targeted Treatment (STarT Back) Trial and IMplementation study to improve Patient Care through Targeted treatment for back pain (IMPaCT) Back study, this thesis demonstrates the application of alternative approaches for estimating productivity costs within an economic evaluation framework. It furthers explores the impact of these approaches on overall productivity and cost-effectiveness analysis results

1.4 Structure of the PhD thesis

The literature reviews conducted in this thesis informed the empirical studies on the validity for the Single-Item Presenteeism Question, feasibility of comparing the sickness certification records with self-report absence records and application of the friction approach in the UK, in order to provide further evidence of alternative methods for valuing productivity loss within economic evaluation. These methodological approaches were then used in the available data sets to generate cost-effectiveness estimates of subgrouping in back pain patients from a societal perspective. Overall, the thesis is reported in nine chapters, the content of which is outlined below:

Chapter Two provides a review of the literature on the theoretical and methodological concepts of economic evaluation and applications of productivity costs. It also discusses the theoretical foundations of economic evaluation and the role of productivity costs within the economic evaluation framework. It then highlights important literature regarding past and on-going debates on the estimation of productivity costs and their inclusion in economic evaluation.

Chapter Three systematically reviews literature on economic studies that have involved applications of the friction cost approach. The review focuses on identifying cost-of-illness and economic evaluation studies that have applied the friction cost approach to productivity cost valuation. This chapter therefore reviews important under-explored questions for future research on productivity costs in economic evaluation.

Chapter Four systematically reviews literature on instruments that have been used in measuring presenteeism and in generating presenteeism monetary estimates. The systematic review focuses on identifying economic evaluation studies that have reported on the measurement and valuation approaches for presenteeism.

Chapter Five provides an overview of the clinical area of low back pain covered in this research. It details the incidence and prevalence of low back pain, and the current treatment strategies available for treatment of the condition. Clinical and economic evidence is provided for the various treatment strategies. Next, evidence of the economic burden associated with low back pain is provided with a particular focus on the UK. Finally a detailed description of the data sets used in this thesis is provided.

Chapter Six describes methodological developments in the application of the friction cost approach. An investigation is carried out to obtain a UK specific length of friction period estimate. The investigation goes beyond the traditional friction cost approach to explore the estimation of more detailed friction periods disaggregated by occupation. Moreover, the comparability of sickness certification records with self-reported work absence records is explored. An investigation is carried out to determine whether sickness certification records can be used as an alternative to self-reported sickness absence records in their absence within economic evaluation.

Chapter Seven examines the construct validity and responsiveness of a Single-Item Presenteeism Question (SIPQ) tool among low back pain sufferers. The validity of the tool is tested by means of correlation analysis, known-group validity and divergent validity. Responsiveness is tested by means of correlational analysis and distribution-based analysis.

Chapter Eight reports findings from the societal cost-utility analysis of stratified care in low back patients from the STarT Back Trial and IMPaCT Back studies. The chapter includes empirical applications of methodological work developed in Chapters Six and Seven and provides an insight into the influence of productivity costs in economic evaluations. The inclusion of full productivity costs of absenteeism and presenteeism using the friction cost approach, and the impact on overall cost-effectiveness outcomes is investigated and reported. Further, the effect of using friction period values disaggregated by occupation is examined. The impact of incorporating effects of back-related sickness on team work and substitutability on productivity cost estimates, and cost-effectiveness outcomes, is presented.

Chapter Nine brings together the aims and objectives of the thesis as set out at the beginning of the thesis. It provides a discussion of the findings, strengths and limitations of the research. The implications for policy and research are discussed and the overall conclusions of the thesis provided.

CHAPTER TWO ECONOMICS AND LOST PRODUCTIVITY

2.1 Introduction

This chapter consists of two parts. Sections 2.1 to 2.4 provide an overview of the theoretical and economic foundations of economic evaluation, and the general approaches to carrying out economic evaluation. Sections 2.5 to 2.9 focus on the role of productivity costs in economic evaluation, what they are and how they are estimated and applied. The chapter also includes a discussion of the main debates and developments in the estimation of productivity costs in relation to both absenteeism and presenteeism and their inclusion in economic evaluation. Finally, an overview of unresolved issues is provided in Section 2.10 to identify areas for further research.

2.2 Theoretical foundations of economic evaluation

Economics is a discipline that is concerned with the allocation of scarce resources. The discipline of economics that is used in guiding the efficient allocation of scarce resources within the healthcare sector is known as health economics (Witter, 2000). To contribute to the resource allocation process, health economists rely on both normative and positive economic theories originating from the discipline of economics (Morris et al., 2007; Coast et al., 2008a). Positive economics is objective and uses theories to *predict* social phenomena, while normative economics is based on value judgements and aims to inform how resources *should* be allocated (Boadway and Bruce, 1984). For the most part, health economics has focused on normative analysis of health and healthcare, particularly through economic evaluation (for example, whether decision makers should fund a particular intervention).

Economic evaluation is a tool used in health economics with the aim of guiding decision makers when choosing between different interventions by comparing their costs and benefits (Drummond et al., 2005). It stems from the realisation that healthcare resources are scarce, and yet market mechanisms are unable to optimally allocate these resources, hence decision makers have to choose between interventions (Williams, 1983; Folland and Goodman A, 2001; Gafni et al., 2012). As a result, there has been a rise in methodological and analytical approaches to economic evaluation of healthcare interventions. The majority of these developments have focused on the more recent analytical frameworks of cost-effectiveness analysis (CEA) and cost-utility analysis (CUA) with a lesser focus on the more traditional cost-benefit analysis (CBA) (Drummond et al., 2005; McIntosh et al., 2010).

2.2.1 Welfare economics and extra-welfarism

Economic evaluation has become increasingly popular in recent years because of its ability to guide decision makers in choosing between programmes (Brouwer and Koopmanschap, 2000). However, this increased popularity has also been followed by controversy, evidenced by the substantial literature on the theoretical foundations of economic evaluation methodology with a particular focus on cost-effectiveness analysis and cost-utility analysis (Garber and Phelps, 1997; Weinstein et al., 1997; Brouwer and Koopmanschap, 2000). Much of the theoretical debate in relation to the foundations of economic evaluation is associated with differences in views on key methodologies in the normative analysis of health and healthcare. These divergent views are based on distinctions between the approaches of 'welfarism' and 'extra-welfarism' (Brouwer et al., 2008).

2.2.1.1 Welfare economics

While the underlying theory of economic evaluation can be traced back to traditional welfare economics, there has been debate around embedding the theory and practice of economic evaluation in the alternative extra-welfarist framework (Brouwer and Koopmanschap, 2000; Birch and Donaldson, 2003). Welfare economics is a theoretical framework concerned with systematic analysis of resource allocation, exclusively in terms of individual utilities and is the principal normative tradition in economics (Drummond et al., 2005; Morris et al., 2007). Welfarism assumes that social welfare (utilities) is an aggregation of individual welfare (utilities) and that individuals are the best judges of their own welfare (Gold et al., 1996; Coast, 2004). In principle, welfarism limits the evaluative space to individuals (Brouwer et al., 2008).

Economic evaluation is based on Paretian welfare economics. The Pareto principle suggests that optimality in resource allocation occurs if one person can be made ‘better off’ without another being made ‘worse off’ (Tsuchiya and Williams, 2001). This principle is, however, impractical as inevitably some people will benefit and some will lose out when policy decisions are made. It is because of this limitation in evaluations that Kaldor (1939) and Hicks (1939), proposed the less restrictive compensation principle (Kaldor, 1939; Hicks, 1939; Layard and Glaister, 1994). According to these proponents, a welfare gain to society could be achieved if the potential gainers (in theory) compensate potential losers and remain better off (Kaldor, 1939; Pearce, 1971). There were questions raised on the equity implications of applying this approach, particularly in the process of evaluating gainers and losers, and how to determine the reallocation of resources. The proponents of this principle, however, argued that the compensation is only hypothetical and that the overall aim should be to assess global welfare gains and not the distribution of this welfare (Brazier et al., 2007).

The cost-benefit analysis (CBA) approach, often based on willingness-to-pay assessments is one type of economic evaluation rooted in the theory of welfare economics that attempts to provide guidance on societal preferences (McIntosh et al., 2010). However, equity concerns have been raised about the appropriateness of using CBA in decision-making because its valuation of outcomes is performed in monetary terms (Drummond et al., 2005). In addition, it has been reported that individuals find it difficult to value health in monetary terms (Ryan et al., 2003). For this reason, the 'extra-welfarist' theoretical framework has been suggested, which goes beyond the traditional welfarist approach (Brouwer et al., 2008).

2.2.1.2 Extra-welfarism

The normative framework of analysis that is not restricted to a function of individual utilities is what is commonly known as extra-welfarism (Brouwer et al., 2008). The focus of welfare economics in the past has been considered too narrow, with suggestions to broaden its perspective to include aspects such as individuals' capabilities (Brouwer et al., 2008). This alternative framework, first proposed by Culyer (1989), draws on a number of different theories including, but not limited to, the capability approach developed by Amartya Sen (Culyer, 1989; Cookson, 2005; Coast et al., 2008b).

In healthcare, extra-welfarism typically extends the evaluative space to one of health, with the main objective being to maximise health from a given budget (Culyer, 1989;

Johannesson, 1995; Brouwer et al., 2008). Health gains are often measured in terms of quality-adjusted life years (QALYs) (Drummond et al., 2005). Priority for resource allocation is placed on those interventions and individuals where the highest health gains in relation to additional costs are realised (Hauck et al., 2004). Nevertheless, it is this specific focus on health in current practice that has caused some to question whether extra-welfarism does indeed offer anything extra above traditional welfarism (Birch and Donaldson, 2003). A practical application of the extra-welfarist approach using QALYs in the UK healthcare sector is seen in the economic guidelines for economic analysis provided by the National Institute for Health and Care Excellence (NICE, 2013). In the recent guidelines, there is an explicit requirement to use the QALY when assessing benefits for evaluations to be submitted to the cost-effectiveness panel (NICE, 2013).

Overall, there are two major theoretical frameworks that decision makers can draw on in making decisions about resource allocation within the health sector. These are the traditional welfarist economics framework and the extra-welfarist framework linked to the capability approach. Each of these frameworks has their associated limitations and the choice between approaches remains a matter of debate.

2.3 Economic Evaluation Theory and Practice

Economic evaluation involves identifying, measuring and comparing the costs and benefits of alternative interventions (Drummond et al., 2005). Evaluation of costs and benefits can be conducted in different ways, which primarily differ in whether and how associated health benefits are considered. The techniques, which include the cost-minimisation analysis, cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, and cost-consequence analysis, are described below.

2.3.1 Cost-minimisation analysis (CMA)

A cost-minimisation analysis is a form of economic evaluation in which an assumption is made that the health outcomes of two or more interventions are equal, and hence the objective is on minimising cost by choosing the least costly intervention (Gray et al., 2010). The attraction with this method lies in its ease of analysis and interpretation (Dakin and Wordsworth, 2013). However, it has been criticised for failure to focus on estimating the joint density of costs and benefits and in quantifying uncertainty surrounding the incremental cost-effectiveness ratio (ICER) (Briggs and O'Brien, 2001). As a result, the cost-minimisation analysis technique has been limited because of its applicability only in rare situations.

2.3.2 Cost-effectiveness analysis (CEA)

In a cost-effectiveness analysis, costs are calculated and the value of health benefits assessed in terms of natural units such as life years gained (Drummond et al., 2005). The difference between the costs, divided by the difference between the benefits, expressed in natural outcomes determines what is known as the incremental cost-effectiveness ratio (ICER) (Drummond et al., 2005). Interventions are then compared against each other based on cost per unit of effectiveness. CEA aims to guide resource allocation through maximisation of health benefits produced within a given budget (Morris et al., 2007).

Nevertheless, the CEA methodology is associated with some limitations. CEA is not suited for resource allocation when interventions have different outcomes because it uses natural units related to the particular intervention, and is therefore appropriate for answering questions on technical efficiency and not allocative efficiency (Drummond et al., 2005).

2.3.3 Cost-utility analysis (CUA)

Increased research efforts into generic measures better suited for allocative efficiency have led to the development of a CEA variant known as cost-utility analysis. The CUA is a form of economic evaluation in which health outcomes are presented as a single generic measure. CUA is based on the extra-welfarist theoretical framework, with health benefits expressed in quality adjusted life years (QALY) (Drummond et al., 2005). The QALY is a generic measure of health outcomes, which combines both aspects of quality-of-life (morbidity) and quantity of life (mortality) into a single measure (Weinstein and Stason, 1977; Williams, 1985). In CUA, the quality-adjusted sets of weights (utilities) are applied to the time spent in potential health states resulting from an intervention. This provides the number of quality adjusted life years hence incorporating the effects of an intervention on both the quality and quantity of life. QALY weights are anchored on a 0 (equivalent to death) to 1 (signifying full health) scale (Drummond et al., 2005). Cost-utility is commonly measured by a ratio (ICER) expressed as an additional cost per QALY gained. Cost-utility analysis has been widely used in decision making within the UK (Buxton, 2006; NICE, 2013) and increasingly in other European countries (Schwappach and Boluarte, 2007; Dagenais et al., 2009).

2.3.4 Cost-benefit analysis (CBA)

CBA is the traditional economic evaluation approach rooted in welfare economics, in which costs and benefits to society are valued and expressed in monetary terms and weighed against each other (McIntosh et al., 2010). The aim of CBA studies is to guide decision makers in prioritising programmes that yield the greatest gain in net benefits to society (Boardman et al., 2006).

The valuation of outcomes in monetary terms provides a platform for the comparison of costs and benefits of an intervention to determine its net social benefit. The net social benefit is the difference between the incremental programme benefits and the incremental programme costs (Dolan and Edlin, 2002). A positive net social benefit signifies that the programme is worthwhile and vice versa. The attraction with this approach is that it enables comparisons to be made both within and across different sectors of the economy (Gafni, 2006).

The two broad methods used in monetary valuation of health outcomes are the human capital approach and the Willingness-to-pay (WTP) or contingent valuation approach (CV). In the human capital approach, the value of monetary benefits is estimated through production gains to the economy (Johannesson, 1996; Drummond et al., 2005). These gains accrue from increased production of individuals at their workplace as a result of treatment, and are valued based on wage-rates. However, this approach has been criticised for ignoring aspects such as estimates of quality of life gains and individuals outside of the labour market (Johannesson and Jonsson, 1991).

Instead, the contingent valuation or willingness-to-pay approach, which provides a framework for assessing the sacrifice individuals are willing to make in order to enjoy benefits of an intervention, has been proposed as a more appropriate method for estimating benefits (Frew, 2010). A growing number of direct stated preference exercises, performed through survey methods based on hypothetical market scenarios, have been reported within the health economics literature (Ryan et al., 2001; McIntosh et al., 2010). Nevertheless, apart from recent methodological advancements in assessing health outcomes through WTP studies (McIntosh et al., 1999), practical application of CBA methods within the economic analysis of healthcare remains relatively limited (Cookson, 2005). This is partly because of the ethical concerns about attaching monetary values on the length and quality of life especially in countries traditionally not based on payment systems (Gafni, 2006). These methodological challenges have meant CBA has rarely been used within healthcare; instead economic evaluation remains dominated by cost-effectiveness analysis and cost-utility analysis.

2.3.5 Cost-consequence analysis (CCA)

In a cost-consequence analysis, costs and benefits are calculated but not combined into cost-effectiveness ratios or QALYs (Russell et al., 1996). Rather, CCA results are reported in a tabular form with all relevant costs and consequences disaggregated and presented separately (Brazier et al., 2007; Gray et al., 2010). Decision makers are then allowed to interpret, prioritise, and assign their own value weights to the tabulated costs and consequences (Coast, 2004). Cost-consequence analysis is often used as the first step to other forms of economic evaluations.

Roberts *et al.,* (2012) recently used a cost-consequence analysis to investigate the costs and benefits of partner notification models for sex partners of individuals with chlamydia, gonorrhoea and non-gonococcal urethritis (Roberts et al., 2012). Healthcare costs were categorised into clinical and pharmaceutical costs and resource data, with the outcome comprising the number of partners treated by strategy listed. As the study was based on an exploratory trial, the authors argued that in this context of cost-consequence, '*Resource use data collected, costs estimated and the results of this study will be useful for informing development of future RCTs*' (Roberts et al., 2012) (p20).

Proponents of this approach have argued that it provides useful information that can enable decision makers to determine returns on investment without requiring detailed costs and effectiveness analysis (Wilkinson, 1999). Furthermore, CCA provides decision makers with the flexibility to tailor cost analyses to their own setting and perspective (Mauskopf et al., 1998). The limitation with this approach however is that decision makers are left with the challenge of interpreting the costs and consequences from the results and in attaching appropriate weights for the different outcomes in question (Gray et al., 2010).

2.4 Decision making in economic evaluation: the threshold approach

Decision makers in many health-care systems are faced with challenges concerning optimal approaches of maximising health benefits subject to healthcare budget constraints (Eichler et al., 2004). The most commonly used analytical framework is the cost-utility variant of the cost-effectiveness approach as evidenced by the various economic evaluation guidelines introduced by many countries (Knies et al., 2010). In order to be transparent and explicit in priority setting, decision makers need a set of decision rules to determine *what is* and *what is not* cost-effective, hence the need for some kind of acceptable threshold value (Eichler et al., 2004).

The threshold approach is the preferred method used by the United Kingdom's National Institute for Health and Care Excellence, an organisation that provides guidance to the National Health Service (NHS) (Culyer et al., 2007). The method of decision-making adopted for this thesis and discussed in this section is based on cost-effectiveness within the UK healthcare context.

In this context, the threshold approach involves estimating an incremental cost-effectiveness ratio (ICER). This is the ratio of additional costs and additional effects generated by an intervention, expressed as the incremental cost per Quality Adjusted Life Year (QALY) gained (see equation 2.1 below).

Equation 2.1

$$\frac{Costs(A) - Costs(B)}{QALYsA - QALYsB} = \frac{\Delta Costs}{\Delta QALYs} = ICER$$

The ICER value generated is then compared against a threshold ICER value of a QALY, above which access to an intervention or programme could be restricted or limited (van Hout et al., 1994).

Internationally, countries have mostly been implicit about an acceptable threshold ICER value, although possible threshold values have been associated with some countries (Appleby et al., 2007; Bridges et al., 2010). In the UK, NICE has been associated with a threshold cost per QALY gained value purported to range between £20,000-30,000 (Rawlins and Culyer, 2004; McCabe et al., 2008). The empirical basis for the cost per QALY threshold range used by NICE or how decisions are taken in relation to the threshold, however, remains somewhat unclear (Appleby et al., 2009). The main debate on the NICE threshold has been on the empirical basis for this threshold and whether the current threshold value should be changed over time (Towse, 2009; Raftery, 2009). Previous work on assessing the social value of a QALY in the UK assessed how QALYs from different healthcare recipients should be weighted and how this potentially affects the existing NICE threshold (Donaldson et al., 2011). The majority of methods used for aggregating the data available resulted in values of a QALY of around £18000 to £40,000 (Donaldson et al., 2011). Additional survey work on this research found no convincing evidence for changing the NICE threshold. Some have argued that in a budget-constrained healthcare system such as the United Kingdom's NHS, the threshold should be based on opportunity cost (Claxton et al., 2013). This is because as new interventions are funded, existing services are likely to be displaced within the system. The work on generating a threshold value based on opportunity cost reported that the 'best' estimate of the threshold value using routine data for 2008 to 2010 was estimated to be around £18,000 per QALY.

Similarly, a threshold value of $50,000 per QALY gained is commonly applied in the United States (Bridges et al., 2010). However, questions have been raised about the theoretical or scientific justification for choosing this value (Bridges et al., 2010). The idea of these thresholds remains an important issue among health economists with recommendations for further research on this value (Weinstein et al., 1997; Hirth et al., 2000; Appleby et al., 2007; Appleby et al., 2009).

One area of controversy surrounding the application of CEA results to the threshold approach, relates to the inclusion of costs that go beyond direct healthcare costs (Olsen and Richardson, 1999). In the context of economic evaluations, exclusion or inclusion of productivity costs could affect overall CEA results. A recent study from the Netherlands (Krol et al., 2011) reported that inclusion or exclusion of productivity costs significantly affected incremental costs in a number of studies. This illustrates the importance of clearly reporting all relevant costs and consequences considered in an analysis. Moreover, further scientific research on whether and how to explicitly determine acceptable decision making ICER threshold values when incorporating productivity costs in economic evaluation would be very helpful.

2.5 Defining productivity costs

Economic analysis of healthcare treatments involves a range of costs that are linked with the health sector (direct healthcare costs), patients and their families (indirect healthcare costs), resource use in other sectors, and productivity costs (mainly absence from work, disability, death and reduced productivity costs). Given the focus of the thesis, this section focuses on the indirect costs associated with work-related productivity loss (Drummond et al., 2005). While healthcare costs were the focus of most early research, recent literature has involved various debates and discussions of estimation and valuation of indirect costs of illness, commonly known as productivity costs.

Productivity costs are defined as 'Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid' (Brouwer et al., 1997a) (p254). The United States Washington panel defined productivity costs as 'costs associated with lost or impaired ability to work or engage in leisure activities due to morbidity and lost economic productivity due to death' (Gold et al., 1996) (p181). The US Washington panel definition differs by specifically incorporating the value of lost leisure time. Productivity costs are often classified into paid working time, unpaid working time and leisure time (Sculpher, 2001).

In the context of paid labour, productivity costs often considered in economic evaluation can occur through patients having paid work and being absent from work as a result of sickness, disability or death - together known as absenteeism, and working with limitations due to sickness - known as presenteeism (van den Heuvel et al., 2010; Sculpher, 2001). Although much research on estimating productivity costs has focused on effects of paid working time, there is also debate around the effects of unpaid labour and leisure time in economic evaluation (Pritchard and Sculpher, 2000). While it is important to consider the effects of sickness and care on unpaid labour and leisure time, this thesis focuses solely on aspects of productivity costs that relate to paid labour for which data was collected in the case studies.

2.6 Considerations for including productivity costs in economic evaluations

The question of whether and how productivity costs should be incorporated in economic evaluation forms an important area of debate (Olsen and Richardson, 1999; Pritchard and Sculpher, 2000; Brouwer et al., 2006). In considering the case for whether productivity costs should be included in economic evaluation, three issues are discussed: the normative foundations for performing economic evaluations, the choice of perspective, and the ethical implications of including productivity costs.

2.6.1 Normative foundations

The normative foundation for economic evaluation is either the traditional welfare economic or recent extra-welfarist economic framework (Brouwer and Koopmanschap, 2000). If a welfarist perspective is utilised in an economic evaluation, then all costs and benefits should be included, hence, justifying the inclusion of productivity costs (Brouwer et al., 1997a; Hauck et al., 2004). As economic evaluation has its theoretical foundations in welfare economics, exclusion of relevant productivity costs in an economic analysis would be inconsistent with the theory in this area (Johannesson, 1995).

The extra-welfarist theoretical framework on the other hand focuses on maximising health for a given budget. As a result, productivity costs are normally excluded on grounds that they are not within the decision maker's budget. This, for example, can be seen in the case of UK's NICE which does not recommend inclusion of productivity costs in economic analysis panel submissions in either the reference-or non-reference case analyses (NICE, 2013). As current practice stands, the extra-welfarist approach, which does not explicitly provide for the inclusion of productivity costs, continues to form the normative foundations for economic evaluations in the healthcare arena within the UK. There is growing interest however in applying the capability approach in health economics with the realisation that interventions are associated with various benefits that accrue beyond health (Coast et al., 2008a), and this broader perspective might necessitate a societal assessment of costs.

2.6.2 The choice of perspective

The perspective of the economic analysis generally determines the scope of relevant costs and benefits to be considered. As such, the relevant perspective plays a critical role in whether productivity costs are to be included in the analysis. The two most utilised perspectives are the healthcare perspective (payer) and societal perspective (Gold et al., 1996; Sculpher, 2001; Drummond et al., 2008). Evaluations performed from a healthcare perspective typically include direct healthcare costs and patient related health benefits, and exclude costs such as productivity costs that do not fall within the payer's healthcare budget. However, the limitation with this perspective is that exclusion of these costs when conducting economic analysis might not accurately reflect the true societal costs of an intervention (Krol et al., 2011). Furthermore, adopting such a perspective would be to deny the reality of costs falling outside the healthcare budget which could lead to biased health policies based on inefficient changes in healthcare and loss in total societal welfare (Byford and Raftery, 1998; Sculpher, 2001; Drummond et al., 2005).

In contrast, economic evaluations carried out from a societal perspective should typically include productivity costs (when appropriate) as a requirement for including all relevant costs and benefits. The common practice however is that economic evaluation studies are often restricted to a narrow healthcare perspective (Johannesson, 2009), although there is a strong theoretical preference for considering a societal perspective (Gold et al., 1996; Byford and Raftery, 1998; Drummond et al., 2008), even when conducted from an extra-welfarist perspective.

Indeed, among health economists in the US, Canada, and the Netherlands, the general consensus is that economic evaluation should be carried out from a societal perspective even when conducted from an extra-welfarist perspective (Gold et al., 1996; Torrance et al., 1996; Byford and Raftery, 1998; Johannesson, 2009). Nevertheless, this perspective is not always the most selected option. For example in the UK, recent NICE guidelines, recommend the use of an NHS and Personal Social Services (PSS) (provider) perspective (NICE, 2013).

In Germany, the Institute for Quality and Efficiency in Healthcare (IQWiG) recommends inclusion of costs from the statutory health insurance perspective (Drummond and Rutten, 2008). Some have argued for the use of a two-perspective approach as the standard (Brouwer et al., 2006). In this approach, cost-effectiveness results should be presented with and without broader costs to society. For example, the most recent NICE guidelines from the UK acknowledge the potential relevance of costs outside the NHS and PSS perspective and provide for their inclusion separately from the reference case, but these do not include productivity costs (NICE, 2013). Similar recommendations have been made in US submission guidelines with a societal perspective accepted in a secondary analysis (AMCP, 2012). Only a few guidelines, for example Australian guidelines, explicitly allow submissions for economic evaluations using either a healthcare or societal perspective (PBAC, 2008).

It is therefore argued in this thesis that a societal perspective, which is more in line with economic theory, would be more appropriate for economic evaluation. This is because it allows for a broader assessment of costs irrespective of who incurs them, hence providing the case for the inclusion of productivity costs in economic evaluation.

2.6.3 Ethical implications

One area of contention in economic evaluation is the inclusion of productivity costs and the possible equity implications that are likely to result from this (Pritchard and Sculpher, 2000). Some have argued that incorporating productivity costs in the analysis may lead to the prioritisation of individuals in society who are in employment, at the expense of those who are not in employment (Koopmanschap and Rutten, 1993; Johannesson, 2009). As such, inclusion of these costs could result in resource allocation decisions that discriminate against some population groups, for example, older people who are not in employment. Others have raised concerns that consideration of productivity loss might also lead to ethical issues around valuing productivity loss for women differently compared to men, as a result of probable differences in wage rates (Luce, 1996).

While these ethical concerns are important, there are different options that can be considered in order to minimise their impact. Some have argued, for example, that economic evaluation can be reported with productivity costs included and equity considerations explicitly reported to decision makers (Brouwer et al., 2006). Moreover, cost-effectiveness analysis results can be presented with, and without, productivity costs to enable policy makers to incorporate their own value judgements to weigh up alternative implications of including these costs (Pritchard and Sculpher, 2000; Koopmanschap et al., 2005) .

In dealing with potential differences in wage-rates, health economists have recommended using alternate wages, including a single average, national or sector-specific wage-rate as equity checks (Russell et al., 1996). In relation to older people, some have argued that productivity gains from those who are working can potentially be translated into healthcare benefits for older people (Johannesson, 2009).

To conclude, in spite of important theoretical and ethical implications, there remain strong grounds both theoretically and practically that productivity costs should be considered in economic evaluation. However, the case for incorporating these costs is stronger when considering a welfarist approach than in the current expression of extra-welfarism. In this thesis, the argument used is that it is important to understand the impact of decisions on resources outside the health sector by using both a healthcare and societal perspective.

2.7 Key factors influencing productivity costs

2.7.1 Absence from work (absenteeism)

Absenteeism can be defined as reduced productivity resulting from individuals not being able to report to work due to health problems (Sculpher, 2001). Absenteeism includes various time components including sick days off work, time away from work related to short and/or long-term work disability, or death (Schultz et al., 2009). Traditionally, health economists have concentrated on estimating indirect costs in relation to absenteeism (Beaton et al., 2005).

2.7.2 Reduced productivity at work (presenteeism)

Presenteeism in healthcare is a concept of lost productivity that refers to reduced productivity or performance while at work due to ill health (Schultz and Edington, 2007; Cooper and Dewe, 2008). Sickness will tend to affect workers in various ways with both the quality and quantity of the employee's work being affected (Hemp, 2004). Recent interest in this concept has been driven by reports that it is associated with much more significant costs than absenteeism. For example, a UK Sainsbury Centre for Mental Health study estimates that presenteeism accounts for 1.5 times more lost working time than absenteeism, with presenteeism costs of mental health problems alone estimated at £15 billion per year (SCMH, 2007). Assessing presenteeism may be particularly important in generating theoretical advances in determining how absenteeism starts and how sickness affects and influences workers in the decision to return to the workplace (Johns, 2010).

2.7.3 Compensation mechanisms and multiplier effects

Recent research has shown that actual productivity costs can be limited during working hours as a result of compensation mechanisms (Koopmanschap et al., 1995; Koopmanschap and Rutten, 1996; Zhang et al., 2012; Krol et al., 2012). This is because a sick employee could potentially make up for reduced productivity or absence during normal hours or by putting in extra hours. Additionally, colleagues at the work place may take on some roles in the short-term at no cost in order to compensate for potential productivity losses.

Indeed, researchers who have explored compensation mechanisms found existence of compensations suggesting overall lower productivity costs than is traditionally reported (Severens et al., 1998; Brouwer et al., 2002; Jacob-Tacken et al., 2005; Krol et al., 2012). However, others have argued that productivity losses from a sick employee can also lead to extra costs as a result of the negative impact on productivity of co-workers particularly where team work is involved (Pauly et al., 2002; Pauly et al., 2008). Such costs have been termed as multiplier effects. Multiplier effects have been shown to impact significantly on overall productivity changes in studies investigating smoking cessation among Dutch citizens (Krol et al., 2012), and in treating rheumatoid arthritis among UK citizens (Zhang et al., 2012).

2.8 Measurement of paid lost productivity: theory and practice

2.8.1 Measurement: Absenteeism

Measurement of lost productivity is generally a complex issue with potential practical and administrative challenges (Evanoff et al., 2002; Pole et al., 2006). Absenteeism is measured by identifying the time individuals have missed from work due to sickness. The two most common approaches for measuring lost productivity from work due to sickness involve the use of (i) objective or (ii) subjective measures. Objective measures include administrative records on sickness absence such as employment databases or registries for work absences due to sickness. Subjective measures include self-reported questionnaires or patient diaries.

Administrative records have the advantage of continuous follow up, easy data collection with relatively low-cost means, and the potential to provide large samples of employee sickness absence records without the possibility of recall errors or perception biases (Pole et al., 2006). Nevertheless, administrative datasets have various limitations, particularly when applied in economic evaluation. The obvious challenge is that there are difficulties in accessing these records due to ethical concerns, cost and effort of having to contact different employers of individuals as part of a study (Knies et al., 2010). Further, databases may have incomplete productivity loss data, loss of employee follow-up due to job losses or changes in work roles, and failure in some cases to capture long-term absence from work data (Young et al., 2002).

Consequently, subjective values based on self-reported estimates of sickness absence from questionnaires have often been applied and have become an acceptable approach in economic evaluation research (Ferrie et al., 2005). Productivity loss is usually generated from self-reported questionnaires, interviews or diary recording methods. The diary and interview methods are generally used less frequently, as they are perceived to be time consuming and costly to administer when compared to retrospective self-reported questionnaires (van Poppel et al., 2002). Nevertheless, these approaches tend to provide more valid estimates as they reduce recall bias in comparison to the use of self-reported questionnaires (van den Brink et al., 2005).

A recall period of three months for absenteeism has recently been recommended (Zhang et al., 2011). However, there is no consensus among economists on the appropriate recall period, although it has been argued that this choice should be guided by a trade-off between increased data collection costs and on responder burden (Zhang et al., 2011). The validity of different recall periods however deserves further scientific research.

2.8.2 Measurement: Presenteeism

There has been an increase in the number of self-report instruments developed to measure the impact of presenteeism at the work place. This was partly in response to the inability of administrative datasets to incorporate and capture presenteeism possibly resulting in an underestimation of productivity costs (Evanoff et al., 2002). Researchers have also attempted to derive methods for attaching a monetary value to presenteeism - mostly over short periods of time - based on the output of the presenteeism scores from these instruments, although many questions in relation to transparency and extrapolation of these costs remain (Brooks et al., 2010). Some of these self-report instruments can be used to generate outputs that provide monetary values (Schultz et al., 2009). However, evidence about the most appropriate methods for estimating the economic burden of presenteeism is still lacking (Lofland et al., 2004; Mattke et al., 2007).

The most popular and widely used presenteeism instruments include the work limitations questionnaire (WLQ) (Lerner et al., 2001) and the Stanford presenteeism scale (SPS) (Turpin et al., 2004). For instruments measuring both absenteeism and presenteeism, the Work Productivity and Activity Impairment (WPAI) (Reilly et al., 1993) and Health and Labour Questionnaire (HLQ) have commonly been used (Brooks et al., 2010). However, few studies (Ozminkowski et al., 2004; Tang et al., 2009; Zhang et al., 2010b), have attempted to compare presenteeism results from two or more presenteeism measures, and the results from these have varied depending on the questionnaire chosen. A recall period of one week for presenteeism has been recommended (Zhang et al., 2011), although further empirical studies assessing the validity of different recall periods for presenteeism are needed. A more detailed review of presenteeism measures is provided in the systematic review in Chapter Four of this thesis.

2.8.3 Measurement: Compensation mechanisms and multiplier effects

To comprehensively estimate productivity costs, data are required pertaining to employment status, absenteeism, presenteeism and the job characteristics of sick employees. Such data are not always easy to collect. In previous attempts, respondents have been asked to report whether their absence was compensated for during their absence or whether extra hours had been undertaken to make up for lost work using questionnaires (Jacob-Tacken et al., 2005; Zhang et al., 2012).

Similarly, few studies to date have attempted to generate wage-related multiplier estimates for adjusting for effects of productivity loss on team productivity (Pauly et al., 2002; Pauly et al., 2008; Zhang et al., 2012). In order to estimate wage-related multiplier effects, data would be required on specific work characteristics such as team work dynamics, and availability of substitutes. Such data are difficult to obtain in an economic evaluation setting because of costs and challenges in obtaining ethical approvals. Data for compensation mechanisms and multiplier estimates can be obtained either from employers or employees. However, the evidence shows that the multiplier values generated from employers are always different compared with those of employees (Jacob-Tacken et al., 2005; Zhang et al., 2012).

Therefore, in spite of recent developments in adjusting for compensation mechanisms and multiplier effects of team work and employee substitutability, more comprehensive studies are necessary before consensus about including these costs in economic evaluation can be achieved. Of particular research interest, the most appropriate source for wage-related multiplier values should be investigated and the implications of including these costs on overall productivity costs more explicitly addressed.

2.8.4 Summary of measurement approaches

To conclude, in light of the methodological challenges associated with self-reported measures of absenteeism, for practical purposes economic evaluation research is better suited to use of self-reported questionnaires, as accessibility to administrative records and objective measurement of lost productivity can be very problematic where patients undergoing any single intervention will work in different work places. In the case of presenteeism, the lack of consensus on the most reliable presenteeism instrument among researchers has probably contributed to limitations in application of presenteeism measures within economic evaluation research. Measures adjusting for compensation mechanisms and multiplier effects remain challenging, but there is scope for further research in assessing whether the best approach is to generate reliable estimates from either employers or employees.

2.9 Valuation approaches for paid lost productivity

2.9.1 Valuation of Absenteeism

The methodology for valuing productivity loss has attracted the most attention and debates among health economists in the area of productivity costing. This section therefore highlights the main areas of discussion and research in the theoretical and practical literature on valuing productivity costs in relation to absenteeism from paid employment. Potential areas for further research are also identified. The discussion focuses on the three main approaches to valuing productivity loss in the field of economic evaluation: i) human capital method (HCM) ii) friction cost approach (FCA) iii) US Washington panel approach (WPA). In current practice, productivity costs are often valued in monetary terms and incorporated in the numerator of the cost-effectiveness (C/E) ratio using either the HCM or FCA, compared to the alternative of their inclusion in the denominator using the WPA (Weinstein et al., 1997).

2.9.1.1 Human Capital method (HCM)

Productivity losses in this approach are estimated by multiplying the appropriate wage rate value with the time loss resulting from absence, disability, premature death or productivity gains from the intervention in question (Liljas, 1998; Sculpher, 2001). The traditional method for generating monetary estimates of productivity loss is the human capital method (Weisbrod, 1961). This approach, based on neo-classical human capital theory (Weinstein et al., 1997), uses the gross wage as the unit of value for assessing the present value of lost time due to illness, disability or death (Johannesson, 1996). Thus, productivity costs are quantified in terms of foregone earnings during absence, with time lost being valued until an individual's retirement age (Johannesson and Karlsson, 1997). This approach can lead to the generation of very high productivity cost estimates.

Therefore, although grounded in economic theory, the method has been criticised for estimating potential productivity loss as opposed to actual productivity loss, hence reporting overestimated costs (Koopmanschap et al., 1995; Koopmanschap and Rutten, 1996). The HCM has also been criticised for distributional allocations that favour specific groups of the population. For example, working men tend to earn higher wage rates than women, the elderly and minority groups, leading to different cost estimates that are potentially more favourable for those in employment and in working men (Tranmer et al., 2005). Some have however argued that while most researchers have used a single industry wage rate, a variation of wage-rates such as adopting the minimum wage or group specific wage-rates can be applied to reduce consequent equity issues (Jacobs and Fassbender, 1998; Berger et al., 2001). The search for the most appropriate wage-rate to use, in assessing foregone earnings due to lost labour within economic evaluation, remains an under-explored area.

2.9.1.2 Friction cost approach (FCA)

The high costs associated with the human capital method have attracted many critiques, with the most notable being Koopmanschap and van Ineveld who argued that the human capital method estimates potential costs as opposed to actual costs since it fails to consider compensation mechanisms and replacements dynamics existing within organisations (Koopmanschap and van Ineveld, 1992) . They argued that the existence of involuntary unemployment and compensation mechanisms within organisations would inevitably limit productivity losses related to absence, disability, and death. Koopmanschap and van Ineveld (1992) therefore developed the friction cost approach based on the argument that productivity loss is limited to the period required to restore production levels within an organisation during a period of time known as the friction period (Koopmanschap and van Ineveld, 1992; Koopmanschap et al., 1995). This period depends on the time it takes to replace and train a replacement worker either from the unemployed or employed ranks of the labour market (Koopmanschap et al., 1997). The method assumes compensation occurs as absent employees make up for lost production on return to work or existing internal labour reserves can prevent short term production loss (Liljas, 1998; Pritchard and Sculpher, 2000; Brouwer et al., 2002). In the long-term, replacements could be made to prevent further productivity loss.

The main focus is then on what is known as friction costs which broadly encompass lost production during the short-term period and costs of hiring and training new individuals (Liljas, 1998). In order to generate these costs, information is required on the frequency, the length of friction period, the average duration of a job vacancy as a proxy for the length of a friction period, and the valuation of lost production during the absence period (Koopmanschap et al., 1995). Vacancy durations depend on the efficiency of the labour market to match demand and supply as well as on prevailing unemployment levels. As a result, it is important to have country-specific vacancy durations for the appropriate application of this method.

The friction cost approach also assumes that work absence results in a less than proportional reduction in labour productivity, yielding what is termed as a labour elasticity value (Koopmanschap and Rutten, 1996). This is an adjustment factor normally used to correct for compensation of work absence during normal working hours. A value of 0.8, originally proposed by Koopmanschap et al., (1995), has often been used in practical applications of the friction cost approach, suggesting that productivity losses are limited to 80% of employee output for absences less than the friction period. The value is based on empirical estimates reported by the Dutch Economic Costing Institute (de Koning and Tuyl, 1984). It is unclear whether this is representative of UK firms, and to date, no attempt has been made to generate a labour elasticity value or investigate the existence of internal labour reserves in firms in the UK.

Compared to the human capital method, the friction cost approach often generates lower estimates particularly in the long-term and proponents of the approach argue that it provides more realistic estimates of lost productivity. Nevertheless, the friction cost approach has been criticised both theoretically and practically, primarily from proponents of the human capital method. The main theoretical criticisms are that the method has no foundation in economic theory, that it has impracticable assumptions (Luce, 1996; Johannesson and Karlsson, 1997; Hutubessy et al., 1999) and that it treats lost leisure time as having no value (Liljas, 1998). Moreover, the friction cost approach was criticised for being limited to paid work. In response, proponents of the approach, for example, Koopmanschap et al. (1997) and Brouwer and Koopmanschap (2005) have argued that the method is an attempt to generate more practical and realistic valuations of the productivity adjustments that occur in the labour economy. In addition, they argue that leisure is not treated as having no value, but that rather at the societal level, it is considered to be equal between the ill and replacement individual, hence attracting no associated costs (Koopmanschap et al., 1997; Brouwer and Koopmanschap, 2005).

However, the main challenge to date in operationalising the friction cost approach is that data are required on the average duration of a job vacancy, elasticity value, the frequency of absenteeism episodes, and the duration of each specific period of absence. Such data are not always available nor easy to collect, which might have contributed to limited application of this approach to date (Tranmer et al., 2005). Further research may help to determine alternative approaches that can be used to generate specific absence episodes for individuals and friction period values specific to country contexts.

2.9.1.3 The US Panel Cost-effectiveness Approach (WPA)

The US Panel approach was proposed by the US Panel for Cost-Effectiveness Analysis of Health and Medicine guidelines (Weinstein et al., 1996). In this method, which differs markedly from the friction cost approach and the human capital method, productivity costs components are said to be valued within the denominator (for example through the QALY measure) (Weinstein et al., 1997; Sculpher, 2001; Brouwer et al., 2002). The main argument of the proponents is that individuals consider the effects of income loss during health state valuations (Liljas, 1998; Pritchard and Sculpher, 2000; Brouwer et al., 2002), therefore productivity costs should be included in the denominator of the cost-effectiveness ratio in order to avoid double counting. This is the preferred approach for the US Panel's reference case (Gold et al., 1996).

This approach when first published generated much controversy, sparking off a debate among economists about inclusion of productivity costs in the denominator or numerator (Weinstein et al., 1997; Brouwer et al., 1997b; Brouwer et al., 1997a). Some have argued that the presence of social security systems and private compensation schemes that partially compensate for income losses due to sickness could lead to misinterpretation of productivity costs from a societal point of view when considering the WPA (Brouwer et al., 1997a).

It has also been argued that respondents could find it difficult to determine how health states are linked to productivity and income. Indeed, existing studies investigating whether and how respondents consider income effects when valuing health states have been largely inconclusive (Tilling et al., 2010). Nevertheless, some have found that respondents did not incorporate income effects when valuing health states, except in situations where explicit instructions were given to the respondents to consider them (Krol et al., 2009). Although the WPA potentially offers benefits in preventing double counting of productivity losses, various questions remain (Sculpher, 2001). This is evidenced by a review (Pritchard and Sculpher,

2000) investigating methods that have been used when incorporating productivity costs in economic evaluation studies that found a dearth of studies applying this method within the United States. There were no applications of this method found in any UK studies, possibly due to the lack of consensus on the appropriateness of the method in the literature (Pritchard and Sculpher, 2000).

In recent research contributing to the inclusion of productivity costs in the numerator versus denominator debate, Nyman has renewed calls for the US to move away from the WPA to one of valuing productivity costs in monetary terms and including them in the numerator of the ICER (Nyman, 2012). In order to ensure the exclusion of leisure in the HCM and FCA, Nyman proposes the use of health status measures that explicitly incorporate questions addressing role functions to capture the impact of sickness on leisure time (Nyman, 2012).

This question of whether to incorporate indirect costs within the QALY component therefore remains an important theoretical consideration, although the preferred approach seems to be the incorporation of productivity costs in the CEA monetary component. Further empirical research on whether individuals consider effects of health states on productivity, and consequently income, when valuing health states in the commonly used health state valuation measures would be valuable. The evidence to date has been inconclusive, and further research could provide greater understanding on how the issue of double counting can be addressed in cost-effectiveness analysis.

To conclude, there have been many debates in the economic evaluation literature regarding the most appropriate approach for estimating productivity costs. In spite of the increased literature, however, methodological consensus on the most appropriate method for valuing productivity costs has not been reached, although the human capital method seems to have been used most often in the UK context (Pritchard and Sculpher, 2000). Some have argued that the friction cost approach would be more appropriate in European countries where the level of both hidden and registered unemployment is sometimes substantial (Berger et al., 2001). Applications of the WPA have been limited since not all economic evaluations use QALYs and due to the uncertainty about whether income effects of lost productivity are really captured by current health state measures (Brouwer et al., 1998; Pritchard and Sculpher, 2000).

Further, no consensus exists between the most preferable approach between the HCM and FCA as shown in existing national economic guidelines (NMA, 2005; Zorgverzerkeringen, 2006; CADTH, 2006; Graf von der Schulenberg, 2008). The methodological differences between productivity cost valuation methods, coupled with a lack of consensus on the most appropriate approach for valuing productivity costs, comprise some of reasons that have contributed to the exclusion of productivity costs from economic evaluation studies. There remains a gap therefore, of empirical work testing applications of existing approaches such as the friction cost approach in varying contexts and settings.

2.9.2 Valuation of presenteeism

Few attempts have been made to comprehensively investigate approaches to valuation of presenteeism, with most of the literature limited to development of measurement instruments. Consequently, values for presenteeism have not been presented as regularly and confidently as in the case of absenteeism (Brooks et al., 2010). The methods used in valuing absenteeism, particularly the human capital method, have largely been extended for use in the monetary valuation of presenteeism (Lofland et al., 2004; Brooks et al., 2010). The human capital method has been adjusted to assess monetary loss by multiplying reduced productivity time losses estimated by the appropriate wage rate.

The advantage of using this method is that presenteeism cost estimates can easily be calculated (Mattke et al., 2007). Nevertheless, there are limitations in its application, as when used in valuing absenteeism, since the HCM does not account for additional wage-related teamwork and work compensation costs. For example, a colleague could be assigned to assist an individual working with limitations in the short-term or an additional worker could be hired in the long-run to limit further productivity loss. In such cases, the HCM might over or underestimate presenteeism costs. The FCA on the other hand appears not to have been translated into practical application for valuation of presenteeism.

2.10 Summary and conclusions: unresolved issues in estimation of productivity costs

2.10.1 Progress to date

Various aspects of productivity cost measurement and valuation in relation to both absenteeism and presenteeism have been explored. The debate on the most appropriate method for valuing productivity loss has received the most attention, with preference for the FCA and HCM. However, to date, little progress among health economists has been made towards achieving a consensus on which of these approaches to valuation is most appropriate. Nevertheless, the HCM is more prevalent both generally and specifically within the UK. In relation to the WPA, empirical studies on income effects in health state valuations support the argument of including productivity costs in the numerator (cost side) as opposed to the denominator (effects) side of the equation.

Furthermore, although the presenteeism concept has been well received in scientific literature, and many productivity loss measurement instruments have been developed and validated, the choice of the most appropriate instrument remains a matter of debate. Recent developments have also seen the extension of total productivity loss to include additional components such as effects of compensation mechanisms and aspects of team work on overall productivity costs. In spite of progress in this area, there are still opportunities for further research to facilitate standardisation of methods in economic evaluation.

2.10.2 Discussion of unresolved measurement and valuation issues

The significant costs associated with productivity losses warrant further research to ensure rigorous and standardised approaches are developed. For example, few attempts have been made to compare different self-report methods such as the recall and diary approaches. Another area of research that has received little attention in work absence methodological research relates to the use of sickness certification records, for example, from the general practice sickness certification database in the UK. Only one study (Wynne-Jones, 2008) has investigated the comparability of these records with self-reported questionnaire data over a short-term period of seven days. However, given the importance of accurately estimating sickness absence as a measure of productivity loss, this area provides opportunities for further research, particularly where productivity loss is estimated over longer periods. Economic evaluation studies linked to certification databases provide an opportunity to assess agreement between self-reported sickness absence data and sick leave certification data (where available). This is an area explored further in Chapter Six of the thesis.

Although many productivity cost instruments have been developed, there are various issues that remain unresolved. Most of these instruments currently rely on self-reported data, which are inevitably subject to possible recall errors. However, few studies (Goetzel, 2004) have attempted to measure the effect of recall periods on productivity loss estimates. Few studies (Ozminkowski et al., 2004; Collins et al., 2005) have attempted to compare measurement instruments within the same population, although different instruments have different approaches to eliciting presenteeism estimates. Furthermore, presenteeism measures have mostly comprised multi-item instruments, while few attempts have been made to review and validate single-item measurements of presenteeism. More information is needed on the instruments and methods used when generating presenteeism monetary estimates. A systematic review of studies that have reported presenteeism costs is provided in Chapter Four of this thesis to contribute to the debate in this area.

The valuation of presenteeism has involved an extension of the HCM as used in the valuation of absence. The debate around whether to use the FCA appears to have remained on a theoretical level. Applications of the FCA when valuing presenteeism are investigated as part of the review of presenteeism studies in Chapter Four. To promote further applications of the friction cost approach, the valuation of presenteeism using this friction cost approach is explored in Chapter Seven of this thesis.

Similarly, the debate on incorporating of compensation mechanisms and multiplier effects assessing the impact of productivity loss on team production has attracted few empirical studies (Jacob-Tacken et al., 2005; Zhang et al., 2012; Krol et al., 2012). However, further evidence is needed on these aspects of productivity cost valuation, as ignoring them could lead to over estimation or underestimation of productivity costs. The application of full productivity costs incorporating multiplier effects alongside economic evaluation is explored in Chapter Seven of this thesis.

The friction cost approach offers benefits to economic evaluation in generating more realistic productivity loss estimates compared with the human capital method. This is because the friction cost approach values productivity costs only during the time it takes to replace a given worker. However it remains unclear how often this approach has been applied in economic evaluation. Whether and how productivity costs are valued can have a significant impact on whether an intervention is cost-effective, particularly in the clinical area of back pain. If productivity loss is reduced by using the friction cost approach, then current interventions in this area could potentially be misrepresented. Further research is therefore needed to extend this approach in current economic evaluation studies within relevant settings. The application of the friction cost approach in a UK setting therefore forms the major foci of the thesis.

2.10.3 Overall conclusion

It is critical to have reliable measurement and valuation approaches for absenteeism and presenteeism if the application of these costs within economic evaluation is to be realised. Given a lack of consensus on a preferred approach for valuation of absenteeism and the recent developments in measurement and valuation of presenteeism, this chapter has discussed the theoretical and methodological debates and issues in relation to estimation and incorporation of lost productivity in economic evaluations towards an increased research agenda in this area. Before moving on to empirical work suggested by this chapter, it is important to understand more fully how the friction cost approach has been applied in the UK and how presenteeism has been measured in economic evaluation studies and these issues are the subject of the next two chapters.

CHAPTER THREE ASSESSMENT OF THE FRICTION COST APPROACH IN PRACTICE: A SYSTEMATIC REVIEW

3.1 Introduction

Chapter Two provided an overview of the theoretical foundations and practical methods for economic evaluation, while it also gives an insight into the role of productivity costs in these studies. The aim of this chapter is to systematically review applications of the friction cost approach (FCA) in practice. It focuses on assessing whether and how the method has been applied in the UK and internationally.

One of the main criticisms of the FCA as discussed in Chapter Two of this thesis is that the approach is difficult to apply mainly because of a general lack of specific country context parameters for its use. One of the key parameters when using the FCA is the length of a friction period, which is the transition period required to replace an absent sick worker in an organisation. The other important aspect is the elasticity value which compensates for work losses during normal working hours. It has been argued that such data are not always available in country-specific settings, which might potentially contribute to limited application of this approach (Tranmer et al., 2005). Although the advantages and limitations of this method have been clearly highlighted, the number of economic studies that have attempted to employ this method, and the country settings of these studies, remains largely unknown.

The aim of this review was therefore to identify, synthesise and interpret the findings of current economic costing research on applications of the friction cost approach to inform policy decision making in healthcare and current practice. The focus is on original cost-of-illness and economic evaluation studies in the UK and internationally, on the basis that this review will provide insights into parameters used when applying the friction cost approach in different settings. These studies are used to assess *whether* and *how* the friction cost approach has been used to assess productivity costs in practice worldwide.

3.2 Methods

3.2.1 Search strategy

Based on a protocol designed for this thesis, systematic literature searches were conducted to identify relevant studies that have valued productivity costs in the literature using the friction cost approach. Articles including friction cost related costs were identified in two ways: 1) database searches 2) searching bibliographies of identified papers. Searches were conducted in the following health bibliographic databases: MEDLINE, EMBASE, PsycINFO, ISI Web of Science and the specific health economics database NHS Economic Evaluation (NHS EED).

The database searches were initially conducted during November 2010, covering the period from 1996 to November 2010, and were limited to those in English for pragmatic reasons. This period was chosen in order to reveal the most recent literature on applications of friction cost approach. All searches were subsequently updated until 30th April 2013.

The search strategies used for the database searches were based on the following key pre-defined search keywords: ‘friction cost’ OR ‘friction cost approach’ OR ‘friction cost method’

OR ‘friction period’. Where relevant, Medical Subject Headings (MeSH) containing the words, friction costs, friction cost approach, and friction approaches were exploded. A bibliographic review of retrieved papers, and relevant systematic review papers eligible for inclusion was performed to identify any additional studies. Full details of the search strategy used in this review are presented in Appendix A.1.

3.2.2 Inclusion and exclusion criteria

Studies were included in this review if they were cost-of-illness, cost-utility, cost-effectiveness, or cost-benefit studies. Articles were included if they discussed, described or reported:

1. Monetary estimates of productivity loss using the FCA
2. Monetary estimates of productivity costs using both the FCA and HCM

and were excluded if they:

1. Gave estimates of productivity using either only the HCM or WPA.
2. Were methodological papers not reporting any monetary costs.
3. Were articles discussing methods of valuing productivity costs
4. Were not full papers (such as conference abstracts, editorials and letters).

3.2.3 Data selection and extraction

Relevant studies were initially identified by their titles and abstracts. After excluding reviews and articles not deemed relevant, the next stage involved reviewing full articles to determine their relevance. Each study was evaluated to determine its importance and relevance for the review based on a set of exclusion and inclusion criteria. Relevant data were extracted, recorded and tabulated from each article based on the extraction criteria below:

Part 1: Overview of the study

- First author
- Country setting of study
- Year of study
- Analytic technique (cost-of-illness, cost-effectiveness, cost-benefit analysis, cost-of-illness)
- Condition
- Data context/Duration

Part 2: Specific methodological valuation approaches

- Elasticity value/Source,
- Friction period values/Source
- Cost of labour measure used
- Whether compensation mechanisms or multiplier effects were included and how
- Whether recruitment and training costs were incorporated
- Macroeconomic effects adjustments
- Productivity costs reported
- Comments on the study

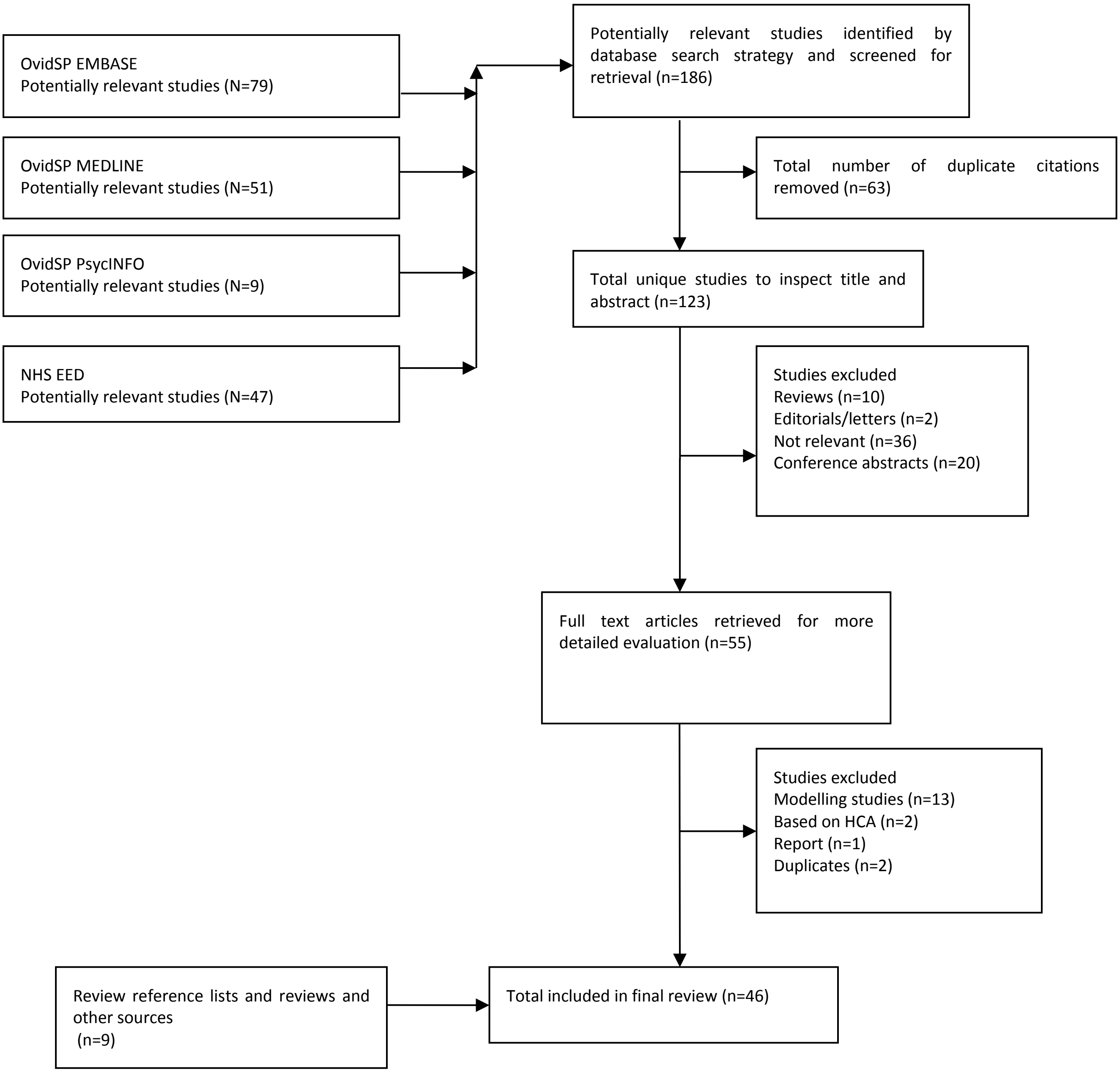


Figure 3:1 Flow diagram showing study selection for review of FCA application

3.3 Results

3.3.1 Study selection

The literature search identified 186 papers. Details of the selection process are illustrated in Figure 3.1. As shown in Figure 3.1, 63 were duplicates resulting in a total of 123 papers. Of the remaining articles screened for titles, abstracts and where appropriate full-text, 10 were reviews, 2 were editorials or letters, 36 were not relevant to the friction cost approach, and 20 were conference abstracts. All of these papers were excluded. The full text of the remaining 55 potentially relevant articles was obtained. A further 18 articles not meeting the study criteria were subsequently excluded. The citations within the remaining articles were checked in order to identify any additional costing and economic evaluation studies meeting the inclusion criteria. This procedure resulted in the inclusion of 9 additional studies. As a result, 46 studies met the inclusion criteria for this review after further exclusion (Figure 3.1). A full list of the studies included is provided in Appendix A.2.

3.3.2 Overview of the Studies

Table 3.2 provides general information about the studies included in the review.

3.3.2.1 Country setting

The majority (28) of the studies were conducted in the Netherlands. Seven studies were set in the United Kingdom, two in Germany, and one in each of Australia, Ireland, Sweden, Canada, Spain, Denmark, Austria, Norway and Greece (Table 3.2). One study conducted an economic evaluation based on a multinational clinical trial setting and used country specific costing (Rutten-van Molken et al., 2007). Another study conducted economic evaluations in Austria and the Netherlands (Van Tubergen et al., 2002) (Table 3.2).

3.3.2.2 Condition

Table 3.1 provides a summary of the disease conditions identified in this review. The largest group of studies were back pain problems followed by depression and mental health related disorders.

Table 3:1 Summary of disease conditions identified

Disease area	No of Studies
Back pain	8
Depression/mental health	6
Chronic obstructive pulmonary disease	4
Limb disorder	4
Neck pain	3
Rheumatoid Arthritis	2
Arthritis	2
Cardiovascular heart disease	2
Cancer	2
Ankolysing spondylitis	1
Haemorrhage	1
Hernia	1
Diabetes	1
Congenital malfunctions	1
Lung transplantation	1
Tubal pregnancy	1
Obesity	1
Physical activity	1
Breast and prostate cancer	1
Stroke	1
Haemodialysis	1
Gestational hypertension	1

3.3.2.3 Economic analysis approach and data context

Of the 46 studies, 33 used information obtained through randomised clinical trials (RCT), nine were national prevalence based studies, one was an observational study, one a national survey and two were economic costing studies based in a hospital setting. Thirty-one studies were cost-effectiveness analyses, of which five incorporated cost-utility analysis. Thirteen were economic cost-of-illness studies, and one used both CEA and CBA (Table 3.2).

Table 3:2 Summary of the studies included in this review

	First author	Year	Country	Perspective	Condition	Analytical technique	Data Context/Setting	Study Duration
1	Borghouts et al	1999	Netherlands	Societal	Neck pain	COI	Prevalence based	1 year
2	Bosmans et al	2000	Netherlands	Societal	Major depression	CEA/CUA	RCT/Treatment	1 year
3	Brouwers et al	2007	Netherlands	Societal	Minor mental disorders	CEA	RCT	2 years
4	Brunenberg et al	2005	Netherlands	Societal	Hip& knee replacement	CEA/CUA	RCT/Treatment	12 months
5	De Brujin et al	2007	Netherlands	Societal	Shoulder complaints	CEA	RCT	26 weeks
6	Dirksen et al	1998	Netherlands	Societal	Hernia	CEA	RCT/Treatment	5 years
7	Van Eijsden et al	2009	Netherlands	Societal	Upper limb no specific disorder	CEA	RCT	1 year
8	Fautrel et al	2007	Canada	Societal	Rheumatoid Arthritis	COI	Convenience sample/survey	1 year
9	Gallefos and Bakke	2002	Norway	Societal	COPD	CEA, CBA	RCT	1 year
10	Goossens et al	1998	Netherlands	Societal	Low back pain	CEA	RCT	3 years
11	Hakkaart-van Roijen et al	2006	Netherlands	Societal/NHS	Depressive disorder	CUA	RCT/Treatment	1.5 years
12	Hakkaart-van Roijen et al	2004	Netherlands	Societal	Bipolar disorder	COI	Costing/ Treatment	3 years
13	Hanley et al	2012	Ireland	Societal/employer	Breast and Prostate cancer	COI	COI	1 year
14	Huscher et al	2006	Germany	Societal	Rheumatic arthritis (RA), ankolysing spondylitis, psoriatic arthritis (AS), systemic lupus (SL).	COI	National treatment database	1 year
15	Hutubessy et al	1999	Netherlands	Societal	Back pain	COI	Prevalence based	1 year
16	Jellema et al	2007	Netherlands	Societal	Low back pain	CEA	C-RCT	1 year
17	Kaitelidou et al	2005	Greece		Haemodialysis	COI		1 year
18	Korthals- de Bos et al	2003	Netherlands	Societal	Neck pain	CEA	RCT	1 year
19	Lewis et al	2007	United Kingdom	Societal/NHS	Neck pain	CEA/CUA	RCT/Hospital	0.5 years
20	Liem et al	1997	Netherlands	Societal	Hernia	CEA	RCT/Treatment	1 year
21	Luengo Fernandez et al	2006	United Kingdom	Societal	Cardiovascular disease	COI	Prevalence/National study	1 year
22	Liu et al	2002	United Kingdom	Societal/NHS	Coronary heart disease	COI	Prevalence/National study	1 year
23	Luijsterburg et al	2007	Netherlands	Societal	Sciatica	CEA	RCT	1 year

	First author	Year	Country	Perspective	Condition	Analytical technique	Data Context/Setting	Study Duration
24	Maniadakis and Gray	2000	United Kingdom	Societal/NHS	Back pain	COI	Prevalence study/national study	1 year
25	Mc Eachan et al	2011	United Kingdom	Societal	Physical activity	CEA	RCT	1 year
26	Mol et al	1999	Netherlands	Societal	Tubal pregnancy	CEA	RCT	3 years
27	Neovius	2012	Sweden	Societal	Obesity	COI	costing	38 years
28	Nikken et al	2005	Netherlands	Societal	Acute peripheral join injury; Knee, Wrist and Ankle	CEA	RCT/University hospital setting	2 years
29	Oliva et al	2005	Spain	Societal	Cervical and Breast cancer	COI	National data	1 year
30	Poley et al	2001	Netherlands	Societal	Congenital anorectal malfunctions	CEA	RCT	1 year
31	Ponto et al	2013	Germany	Societal	Graves orbitopathy	COI	Observational study	4 years
32	Rivero-Arias et al	2010	United Kingdom	Societal/NHS	Aneurismal subarachnoid haemorrhage(Asah)	COI	National/Cohort/prevalence based	1 year
33	Rutten-Van Molken	2007	Netherlands	Societal/NHS	COPD	CEA	RCT/Multicentre	12months
34	Saka et al	2009	United Kingdom	Societal	Stroke	COI	Prevalence-based	1 year
35	Smit et al	2006	Netherlands	Societal	Depression	CEA	RCT	12 months
36	Soegaard et al	2007	Denmark	Societal	Low back pain	CUA	RCT	4- 8 years
37	Stant et al	2003	Netherlands	Societal	Schizophrenia	CEA	RCT/Treatment	18 months
38	Steenstra et al	2006	Netherlands	Societal	Low back pain	CEA/CUA	RCT	2 years
39	Steuten et al	2007	Netherlands	Societal	Diabetes	CUA	Hospital setting/Treatment	2 years
40	Steuten et al	2006	Netherlands	Societal	Asthma/COPD	CEA	Pre-post study	24 months
41	Van de hout	2009	Netherlands	Societal	Rheumatoid Arthritis	CUA	RCT/Treatment	2 years
42	Van de Roer et al	2008	Netherlands	Societal	Low back pain	CEA	RCT	52 weeks
43	Van Enckevort	1997	Netherlands	Societal/life-time	Lung transplantation	CEA	Cohort study	12 months
44	Van Schayck et al	2009	Netherlands	Societal	Constructive obstructive pulmonary disorder	CEA	RCT/Treatment	1 year
45	Van Tubergen et al	2002	Austria/Netherlands	Societal	Ankylosing Spondylitis	CEA/CUA	RCT	10 months
46	Vijgen et al	2010	Netherlands	Societal	Pre-eclampsia	CUA	RCT/Treatment	12 months

3.3.2.4 Analytical Horizon

In terms of the length of the study, 27 studies were carried out over a period of one to two years, 9 a period of two or more years, and the rest of the studies for a period shorter than one year (Table 3.2).

3.3.2.5 Study perspective

In this review, all studies adopted a societal perspective, of which nine also used a healthcare perspective and one also used an employer perspective (Hanly et al., 2012) (Table 3.2).

3.3.3 Methodological aspects of valuing productivity costs

3.3.3.1 Length of friction period considered

The key parameter when valuing productivity loss using the FCA is the friction period. Overall, 35 (76%) studies explicitly reported the length of the friction period used, of which only 13 (31%) stated the source of this estimate. The reported length of a friction period from the studies identified was between 2 and 6 months (Table 3.3).

The length of the friction period has mostly been assessed as a single value, yet some have argued that there are potential differences between individuals that could lead to an estimation of inaccurate societal costs for interventions (Koopmanschap and Rutten, 1996). The review found only one study that attempted to use disaggregated friction period estimates (Hutubessy et al., 1999). This study, from Ireland, used estimates of 2.8 and 3.3 months for individuals having received basic education and intermediary education respectively, with these estimates being taken from a previous piece of methodological work in the Netherlands (Koopmanschap et al., 1995).

Additionally, the friction period depends on the level of unemployment (availability of labour) and the ability of the labour market to match demand and supply of labour, factors that differ across different countries. It is therefore important to have a length of friction period relevant to the particular setting. In this investigation, the majority of studies from the Netherlands provided clear sources for the friction period value used. Where reported, the source of the length of friction period originated from the Netherlands, with the exception of one study (Maniadakis and Gray, 2000) that used friction period estimates from the UK, one from Germany (Ponto et al., 2013), and one from Ireland (Hanly et al., 2012).

3.3.3.2 Friction costs to the employer

Friction costs to employers are the costs of filling a vacancy and training new personnel if an employee is to be replaced permanently. There is generally a dearth of information on the friction costs to employers. From the review, the majority of studies did not attempt to estimate friction costs to the employer. Only one study from Denmark provided estimates of the employers' friction costs, reporting employer replacement costs of $1670 (Soegaard et al., 2007) (Table 3.3).

Table 3:3 Detailed extractions of friction cost approach methodological details

	First author, year	Country	Friction period	Friction period source	Elasticity Value/Source	Employers friction costs included	Macroeconomic effects included	Absence/presence source	Labour measure used	Compensation mechanism/ multiplier effects
1	Borghouts et al.,1999	Netherlands	90 days	Koopmanschap, 1994	0.8/ Koning and Tuyl, 1984	NI	NI	No of sick days/Source	Mean daily wages	NI
2	Bosman et al., 2007	Netherlands	123 days	NI	NI	NI	NI	Lost days/cost diaries	Mean age-gender income	NI
3	Brouwers et al.,2006	Netherlands	154 days	Lamers, 2005	0.8/ Oostenbrink J et al., , 2004.	NI	NI	Number of hours of absenteeism/Questionnaire	Mean income by age and gender.	NI
4	Brunenberg et al., 2005	Netherlands	NI	NI	NI	NI	NI	Absence days/questionnaire	Standard friction cost per hr – Oostenbrink, 2000	NI
5	De Brujin et al.,2007	Netherlands	154 days	Koopmanschap et al., 1995.	NI	NI	NI	Absence days/Cost diaries	NI	NI
6	Dirksen et al.,1998	Netherlands	NI	NI	NI	NI	NI	No of absent days/NI	age-gender gross-wage	NI
7	Van Eijsden et al.,2009	Netherlands	154 days,	Koopmanschap et al, 1995.	NI	NI	NI	Number of days absent/Questionnaire	Average wage	NI
8	Fautrel et al., 2007	Canada	NI	NI	NI	NI	NI	No of sick leave days	Gender-matched wages; Average hourly wages	NI
9	Gallefos and Bakke., 2002	Norway	NI	NI	0.7/Source no provided	NI	NI	Number of days absent from work/questionnaire	National average daily wage rate	NI
10	Goosens et al.,1998	Netherlands	3 months	Koopmanschap et al, 1992	NI	NI	NI	Number of days absent from work/Cost diary	National average gross hourly wage	NI
11	Hakkaart-van Roijen et a., 2004	Netherlands	5 months	NI	NI	NI	NI	Absence days/reduced productivity/HLQ Questionnaire	Average Income/wages	NI
12	Hakkaart-van Roijen	Netherlands	154 days	NI	NI	NI	NI	Absence/questio	Standard friction	NI

	First author, year	Country	Friction period	Friction period source	Elasticity Value/Source	Employers friction costs included	Macroeconomic effects included	Absence/presence source	Labour measure used	Compensation mechanism/ multiplier effects
	et al., 2006							nnaire-TiCP	cost per hr – Oostenbrink, 2000	
13	Hanley et al.,2012	Ireland	11.3weeks	Average duration	1.0 (0.8 sensitivity analysis).	NI	NI	Absence from work/questionnaire	Age-gender gross wage-rates	NI
14	Huscher et al.,2006	Germany	58 Days	Federal labour office	NI	NI	NI	No of sick leave days	Average daily wage	NI
15	Hutubessy et al.,1999	Netherlands	3 months, (basic education and an intermediary education 2.8 and 3.3 months)	NI	0.8/ de Koning and Tuyl, 1984).	NI	Yes	No of sick days/ Social Insurance Council in the Netherlands	Wage by age, sex and education level	NI
16	Jellema et al.,2007	Netherlands	154 days	NI	NI	NI	NI	Sick leave days/prospective cost diaries	Mean average income by age and sex.	NI
17	Kaitelidou et al.,2005	Greece	NI	NI	NI	NI	NI	Absenteeism/presenteeism/questionnaire	Mean gross hourly wage	NI
18	Korthals- de Bos et al., 2003	Netherlands	122 days	NI	NI	NI	NI	Days of inactivity	Mean income by age and sex	NI
19	Lewis et al., 2007	United Kingdom	6 months	NI	NI	NI	NI	Days off work/questionnaire	Average daily wage	NI
20	Liem et al., 1997	Netherlands	75 days	NI	NI	NI	NI	Sex-and age-dependant friction costs	Sex-and age-dependant friction costs	
21	Luengo Fernandez et al.,2006	United Kingdom	90 days	Liu et al., 2002	NI	NI	Yes (economic activity and unemployment)	Average duration of each spell on incapacity	Average annual earnings	NI

	First author, year	Country	Friction period	Friction period source	Elasticity Value/Source	Employers friction costs included	Macroeconomic effects included	Absence/presence source	Labour measure used	Compensation mechanism/ multiplier effects
22	Liu et al., 2002	United Kingdom	90 days	Maniadakis and Gray, 2000	NI	NI	Yes (economic activity and unemployment)	Mean certified days off work/ department of employment.	Average annual earnings of workers	NI
23	Luijsterburg et al.,2007	Netherlands /societal	154 days,	Koopmanschap, 1995.	NI	NI	NI	Number of days absent/Questionnaire	Mean income of Dutch population	NI
24	Maniadakis and Gray.,2000	United Kingdom	90 days	NI	NI	NI	NI	Days of certified incapacity/Department of social security	Average wage	NI
25	McEachan et al., 2010	United Kingdom	NI	NI	0.8	NI	NI	Days off work/questionnaire	Median UK Wage rates	NI
26	Mol et al.,1999	Netherlands	10 weeks	NI	NI	NI	NI	Absence days/ Prospective questionnaire	Age- and sex-stratified income data	NI
27	Neovius et al., 2012	Sweden	6 months	Van de Hout et al 2009	NI	NI	NI	No of days lost/register	Average annual salary	NI
28	Nikken et al.,2005	Netherlands	6 months	Oral communication, MA Koopmanschap, 2002.	NI	NI	NI	Number of days absent from work/questionnaire	Age and sex friction cost data Oostenbrink et al 2000	NI
29	Oliva et al.,2005	Spain	75 days (2.5 months) Sensitivity analysis (105 days).	NI	NI	NI	NI	NI	Mean income by age and sex	NI
30	Poley et al.,2001	Netherlands	NI	NI	NI	NI	NI	Number of days absent from work/questionnaire	NI	NI
31	Ponto et al., 2013	Germany	58 days	Regional employment office	1.0, 8.0, 0.3	NI	NI	Sick-leave data/questionnaire	Average wage-rate	NI

	First author, year	Country	Friction period	Friction period source	Elasticity Value/Source	Employers friction costs included	Macroeconomic effects included	Absence/presence source	Labour measure used	Compensation mechanism/ multiplier effects
32	Rivero-Arias et al., 2010	UK	90 days	NI	NI	NI	NI	.Incapacity spells data/From UK trial database	Average annual earnings	NI
33	Rutten-Van Molken et al., 2007	Netherlands	154	NI	NI	NI	NI	Days /hours /questionnaire	Weighted average of gross hourly earning full-time	NI
34	Saka et al.,2009	UK	NI	NI	NI	NI	Yes (economic activity and unemployment)	Mean annual certified days off work from stroke	Mean daily earnings	NI
35	Smit et al., 2006	Netherlands	5 months	NI	NI	NI	NI	No of absence days/Number of 'work cut-back' /questionnaire TiCP	Age and sex friction cost data Oostenbrink et al 2000	NI
36	Soegaard et al.,2007	Denmark	3 month friction period	NI	NI	Replacing an employee =USD 1670;	NI	Absenteeism and disability /Social science research register.	National average Gross income in age and gender	NI
37	Stant et al., 2008	Netherlands	5 months	NI	NI	NI	NI	Absence days/questionnaire	NI	NI
38	Steenstra et al.,2006	Netherlands	122 days	Oostenbrink et al., 2002.	NI	NI	NI	Days on sick leave/ postal questionnaires	Age and sex dependent loss per day.	NI
39	Steuten et al., 2006	Netherlands	NI	NI	NI	NI	NI	No of sick days off/questionnaire	NI	NI
40	Steuten et al., 2007	Netherlands	NI	NI	NI	NI	NI	sick-leave days/questionnaire	Age dependent Friction costs	NI
41	Van de hout et al., 2009	Netherlands	6 months	NI	0.8	NI	NI	Absence days/ cost diaries	Age-dependent hourly costs	NI
42	Van de Roer et al ., 2008	Netherlands	NI	NI	NI	NI	NI	Hours of absenteeism	Mean wage by age and gender.	NI
43	Van Enckevort et al.,	Netherlands	3 months	Average	NI	NI	NI	Absence	Average wage	NI

	First author, year	Country	Friction period	Friction period source	Elasticity Value/Source	Employers friction costs included	Macroeconomic effects included	Absence/presence source	Labour measure used	Compensation mechanism/ multiplier effects
	1997			vacancy duration				days/questionnaire		
44	Van Schayck et al., 2009	Netherlands	NI	NI	0.8	NI	NI	Lost working hours/cost-diaries	Age-group gender productivity costs per hour	NI
45	Van Tubergen et al.,2002	Austria/Netherlands	4 months	NI	NI	NI	NI	Number of days of illness-related absence from work/Cost diary	Gross earnings of patients	NI
46	Vijgen et al., 2010	Netherlands	10 weeks	NI	NI	NI	NI	Absence days/questionnaire	Age-gender wage rate	NI

*NI Not included in study

3.3.3.3 Labour elasticity value

Another key component of the FCA relates to the elasticity value. The identified studies were evaluated to determine if the elasticity of labour time versus labour productivity value (elasticity value), often used to adjust for short-term work compensations, was applied within the studies. The studies were also assessed for inclusion of multiplier effects that have been reported to have a significant impact on overall productivity costs (Krol et al., 2012). In this review, there were no studies that included costs resulting from negative impacts of absenteeism or presenteeism on teamwork, substitutability or overall productivity (multiplier effects). Furthermore, none of economic evaluation studies included specific compensation mechanisms that have been shown to significantly reduce the impact of productivity costs on total societal costs (Jacob-Tacken et al., 2005).

In relation to the elasticity of labour value, the review found that only 9 out of the 46 studies reviewed explicitly reported results using the elasticity correction factor (Brouwers et al., 2007; Borghouts et al., 1999; Hutubessy et al., 1999; van den Hout et al., 2009; Hanley et al., 2003; Gallefoss and Bakke, 2002; Ponto et al., 2013; Van Schayck et al., 2009; McEachan et al., 2011). Of these, seven used an elasticity value of 0.8 originating from the Netherlands (Koopmanschap et al., 1995), while one study (Gallefoss and Bakke, 2002) used a value of 0.7 without providing a clear source. One study from Germany used varying elasticity values (1, 0.8, and 0.3) for varying sick leave periods, but provided no clear source for these values (Ponto et al., 2013). The majority of the studies included in the review, however, did not include any workplace adjustments used to reduce short term production losses (Table 3.3).

3.3.3.4 Medium-term macro-economic effects

Another important factor when using the FCA is the occurrence of medium term macroeconomic consequences outside the friction period resulting from sickness, disability and treatment (Koopmanschap et al., 1995). A macro-econometric model has previously been recommended as a way of estimating macro-economic consequences of work absence and disability (Koopmanschap et al., 1995). The economic evaluation studies identified focussed on generating short-term friction costs, and the majority did not explicitly estimate a medium term effect or report using the model by Koopmanschap and colleagues (Table 3.3). Only three cost-of-illness studies from the UK accounted for medium-term impact by adjusting productivity loss for rate of economic activity and current unemployment alongside the friction period (Saka et al., 2009; Luengo-Fernandez et al., 2006; Liu et al., 2002).

3.3.3.5 Valuation of productivity costs

The valuation of productivity costs using the friction cost approach should consider the duration of absence in relation to that of the friction period. Often, productivity losses of absence periods shorter than the friction period are adjusted with the labour elasticity value. Productivity costs of absences longer than the friction period are assumed to be equal to the full cost during the friction period (Koopmanschap et al., 1995).

In this review, the majority of studies did not provide enough details to identify whether this adjustment was carried out, and neither was the actual estimation approach used clearly documented. Such a lack of detail makes it difficult to compare the calculations used by the different studies. Only one study (Ponto et al., 2013) was identified as having explicitly applied elasticity values for varying sick-leave periods (Table 3.3).

3.3.3.6 Cost components and productivity loss data context

Indirect costs broadly include a variety of costs such as productivity losses (absenteeism and presenteeism), and informal costs such as family or relative costs. All 46 studies reviewed included the estimation of lost time due to absence from paid working time, with one study (Smit et al., 2006) incorporating reduced productivity loss using the FCA. None of the studies incorporated leisure related costs. Fifteen studies estimated monetary benefits of unpaid labour, while four included patient travel related costs (Table 3.3).

The majority (38 of the 46 studies) used days/hours absent from work as the productivity losses component whose cost was assessed (Table 3.3). Four of the studies used certified days of incapacity as the productivity loss component (Maniadakis and Gray, 2000; Saka et al., 2009; Rivero-Arias et al., 2010; Liu et al., 2002). The rest of the studies did not provide details of the productivity metric used to estimate the reported productivity costs. Only two studies (Rutten-van Molken et al., 2007; Hakkaart-van Roijen et al., 2004) reported recording start and end dates of each absence spell (which is required to successfully apply the friction cost approach). The data sources for time loss were: national department dataset (four studies), cost-diaries (seven studies), and questionnaires (21 studies), with the remaining studies not clearly detailing the data sources (Table 3.3).

3.3.3.7 Value of lost labour

A variety of measures have been used to assess the value of foregone earnings. The following measures were used in the majority of studies: a wage for the relevant age-sex dependent group (19) and an average wage for all groups (18). The exceptions were a weighted average of gross wage (Rutten-van Molken et al., 2007), median wage-rate (McEachan et al., 2011), and gross earnings of study participants (Van Tubergen et al., 2002). Seven studies reported using the age-group gender based productivity cost per hour from the Dutch costing manual (Brunenberg et al., 2005; van Roijen et al., 2006; Liem et al., 1997; Smit et al., 2006; Van Schayck et al., 2009; Steuten et al., 2007). The value of foregone earnings could not be established in the remaining four studies (Table 3.3).

3.4 Discussion

The aim of this systematic review was to assess economic costing studies in which the friction cost approach has been used in valuing productivity costs in the UK and internationally. It focused on cost-of-illness and original economic evaluation studies. The systematic review identified several studies internationally and a very small number in the United Kingdom, and provides valuable information to understand methodological applications and shortcomings of the friction cost approach. The findings showed wide variations in the friction cost approach valuation methodology, highlighting the need to increase transparency in the way in which the method is applied and to generate country specific economic parameters. The results were presented in two major categories: (i) an overview of the studies and, (ii) a detailed review of methods used to generate friction cost estimates.

The systematic review results showed wide variations in friction period estimates from different countries, an overall lack of transparency in data sources of these estimates, and how aggregated productivity costs were calculated. Moreover, few applications of disaggregated friction periods have been generated, with only one study being identified. Estimates of the labour elasticity factor and aspects of compensation mechanisms, multiplier effects and medium-term macro-economic effects were largely ignored in the studies reviewed. Overall, costing studies incorporating the friction costing approach are still rare, which could perhaps be attributed to the lack of standardised country context economic parameters.

The findings from this review are in agreement with previous literature assessing the estimation of productivity costs in practice (Pritchard and Sculpher, 2000). In their comprehensive review, Pritchard and Sculpher identified 40 economic evaluation studies of which only seven used the friction cost approach. The findings from their review indicated that in a number of studies, the methods used to estimate productivity costs were not clearly stated, and they advocated improvements to the reporting of productivity costing methods. This was particularly the case for studies applying the friction cost approach and the US Panel cost-effectiveness approach. Nevertheless, the systematic review in this chapter differs from this study in two major ways. Firstly, Pritchard and Sculpher (2000) assessed applications of the three productivity cost valuation methods. The systematic review in this chapter focusses in greater detail on applications of the friction cost approach and hence complements the findings from their review. Secondly, they review only economic evaluation studies from the Health Economic Evaluation Database. This review considers cost-of-illness and economic evaluation studies from various databases and therefore provides a more comprehensive assessment of the current state of the literature.

Currently, there is very limited evidence on reported friction period estimates within the United Kingdom. The reported friction period values for the United Kingdom did not include explicit details of relevant data sources (Maniadakis and Gray, 2000; Lewis et al., 2007; Saka et al., 2009; Rivero-Arias et al., 2010). In addition, few attempts have been made to disaggregate friction periods according to different population groups (Koopmanschap et al., 1995; Hutubessy et al., 1999). Koopmanschap et al. (1995) previously reported friction period estimates disaggregated by education level in the Netherlands. No such estimates have been generated in the United Kingdom.

Overall, the systematic review results presented and discussed here highlight a need for clearer, structured approaches on how studies apply and report methods of the friction cost approach in current practice. Currently, gaps exist in identifying up-to-date valid estimates of important factors such as a length of friction periods and elasticity factor relevant for specific country settings. Applications of the friction cost approach have routinely relied on economic parameters not clearly reported or specific to their country settings. Only one study from the review used the FCA in valuing presenteeism and none was found in the UK.

Strengths and limitations

The systematic review has some strengths and limitations. One strength is the diverse range of databases from which studies were identified. In addition, the review comprehensively assesses cost-of-illness and original economic evaluation studies to provide important information on how data relevant for the friction cost approach has been generated. Although care has been taken to include all relevant studies, it should be noted that in searching the literature, some studies that estimated productivity costs using the friction cost approach may have been missed. In the search, specific terms to the method such as “friction period” and “friction cost approach” were used to identify studies. Although more general terms such as “productivity costs” could have been used, in the context of this study (for studies in all disease areas estimating productivity costs), these could have significantly increased the number of potential abstracts. Thus, a decision was made to go for a search strategy that was more specific to the friction cost approach. This strategy may have missed studies that used the friction cost approach, but did not use any of these key words in the abstract and were not registered in the chosen databases. Moreover, the review limited the chosen studies to original economic evaluations with modelling studies excluded. This is because the review aimed to assess how friction cost approach related data is practically collected and generated in practice. Additionally, potentially relevant studies could have been excluded due to the search strategy employed in the review protocol that excluded non English articles and was initially based on reviewing abstracts and titles of papers. However, attempts were made to comprehensively review relevant databases and citations of relevant literature in this area.

Implications for policy and research

Given the limited number of economic evaluation studies that have applied the friction cost approach in the UK, these findings provide valuable information for researchers and policy makers. Irrespective of the methodological challenges identified, additional economic evaluation studies applying the friction cost approach are required in the UK to provide further evidence that will enable the comparison of cost estimates with other valuation methods and further knowledge on standardised approaches of using the friction cost approach.

The general lack of explicit economic parameters necessary for applications of the friction cost approach in the UK context necessitates prospective studies estimating the length of a friction period, and how this potentially varies in different subgroups. Other potential areas of research include assessing labour elasticity values relevant for the United Kingdom. Analysis of friction costs to the employer in training and recruiting replacement workers should also be considered in such studies.

3.5 Conclusion

In summary, information on applications of the friction cost approach in economic evaluations in the UK and internationally is scarce and the reported methods where applied are unclear. Given a lack of consensus among economists on the most preferred approach for valuing productivity costs, this review has critically reviewed applications of the friction cost approach in cost-of-illness and economic evaluation studies. It has revealed important gaps, requiring improvements in reporting methodological aspects of the friction cost approach. One of the most important issues is the lack of any evidence of appropriate friction periods for the UK, and the absence of information about how the friction periods vary by occupation. Obtaining better evidence on this is thus one of the major foci of the thesis.

CHAPTER FOUR LITERATURE REVIEW OF REDUCED PRODUCTIVITY AT WORK (PRESENTEEISM) MEASURES AND THEIR INCLUSION IN ECONOMIC EVALUATION STUDIES

4.1 Introduction

Chapter Three comprised a systematic literature review assessing productivity costs in practice using the friction cost approach. The aim of this chapter is to report on a further systematic review, in this case, of the assessment of presenteeism costs in current practice. The chapter also reports on a scoping review identifying currently existing instruments developed for the purpose of measuring reduced productivity at work.

The chapter begins by providing a background to the systematic review. This is followed by the methods used to identify the articles considered for this review and the approaches used in extracting the necessary information. The literature obtained on the methodological issues around assessing presenteeism in practice is then discussed. Finally, an overall discussion and conclusion is provided, highlighting implications of the findings, and gaps in the current research.

4.2 Background to review

Most national economic guidelines, adopting a societal perspective, currently recommend inclusion of absenteeism costs, but not presenteeism (Knies et al., 2010). One of the main criticisms with considering presenteeism in costing studies is a lack of consensus about the methodological approaches that should be used to generate monetary estimates from existing measures (Johns, 2010; Brooks et al., 2010). In spite of the growing debate on the appropriateness of including these costs in economic evaluation, whether and how presenteeism costs are estimated in economic studies remains largely unknown. The systematic review aims to contribute to the debate by providing an overview of existing presenteeism instruments and evaluating whether and how presenteeism has been assessed in economic studies.

The research objectives for this review were:

(i) To identify current instruments used in measuring health-related reduced productivity. This was investigated via a review of systematic literature review studies

(ii) To assess presenteeism costs in practice. This was assessed by systematically reviewing and identifying scientific publications that have reported monetary estimates of presenteeism. The review will contribute to current research questions in this growth area.

4.3 Methods

4.3.1 Search strategy

The review presented in this chapter involved a two stage process. In the first stage, an evaluation of existing systematic reviews reporting presenteeism instruments was undertaken from a variety of databases as part of a scoping review exercise to identify and populate a list of existing presenteeism instruments. The scoping review was also used to assess the breadth of literature on the estimation of presenteeism from these databases. Subsequently, a structured systematic review was undertaken to identify economic studies that have reported presenteeism costs based on a protocol designed for this thesis. This was also to assess whether and how the existing presenteeism instruments identified in the first stage have been used in the measurement of presenteeism within economic evaluation practice. Papers relevant to this review were identified in two ways: (i) bibliographic database searches (ii) searching bibliographies of papers identified. The searches were conducted in three key databases spanning the medical sciences (MEDLINE) and behavioural sciences (PsycINFO) commonly used in the systematic reviews identified in the scoping exercise, and in health economics (NHS EED). The NHS EED database was included to provide a more comprehensive assessment of the costing literature.

The database searches were conducted in January 2011, with no starting date limitation, in order to provide an in-depth analysis of existing instruments and costing studies limited to those in English. However, all searches were updated to 30th April 2013. The search strategies used for the database searches were based on the following pre-determined key words: used in the main search were 'presenteeism' or 'reduced productivity', or 'productivity costs' or 'lost productivity' or 'work limitations' or 'work productivity' or 'work performance' and subsequently in conjunction with the terms 'cost and cost analysis' or 'cost-effectiveness analysis' or 'cost-utility analysis'. Where relevant, MeSH headings were exploded. A similar search strategy was used to search for relevant papers in the PsycINFO and NHS EED. A bibliographic review of all retrieved papers, other electronic sources, and systematic review studies relevant to this review was performed to identify any additional studies. Full details of the search strategy used in this review are presented in Appendix B.1.

4.3.2 Inclusion and exclusion criteria

Two broad categories of studies were considered for this review: The first involved including all relevant systematic review studies detailing rating scales of presenteeism. In the systematic review, articles were included if they discussed, described or reported:

1. Systematic review studies detailing rating scales measuring presenteeism.
2. Economic evaluations or cost-of-illness studies reporting presenteeism estimates.

and were excluded if they were:

1. Studies reporting absenteeism only costs
2. Methodological papers not providing any monetary estimates of presenteeism
3. Letters and reviews not relevant to this investigation.
4. Not full papers (such as conference abstracts, editorials and letters).

4.3.3 Data extraction strategy

Relevant information for this systematic review was initially identified by reviewing abstracts of all studies and/or by reviewing their titles. After excluding articles that were not found to be relevant, full articles for the remaining studies were downloaded and assessed for relevance to the review based on the pre-set inclusion and exclusion criteria. Relevant information was then extracted, and tabulated based on the extraction criteria detailed below.

Part 1: Overview of the study

- First author/year
- Country of setting
- Analysis technique (cost-of-illness, cost-effectiveness, cost-benefit analysis, cost-of-illness)
- Condition

Part 2: Specific methodological valuation approaches

- Instrument used/recall period
- Monetary conversion method
- Productivity related domains included
- Whether compensation mechanisms were included
- Cost components reported

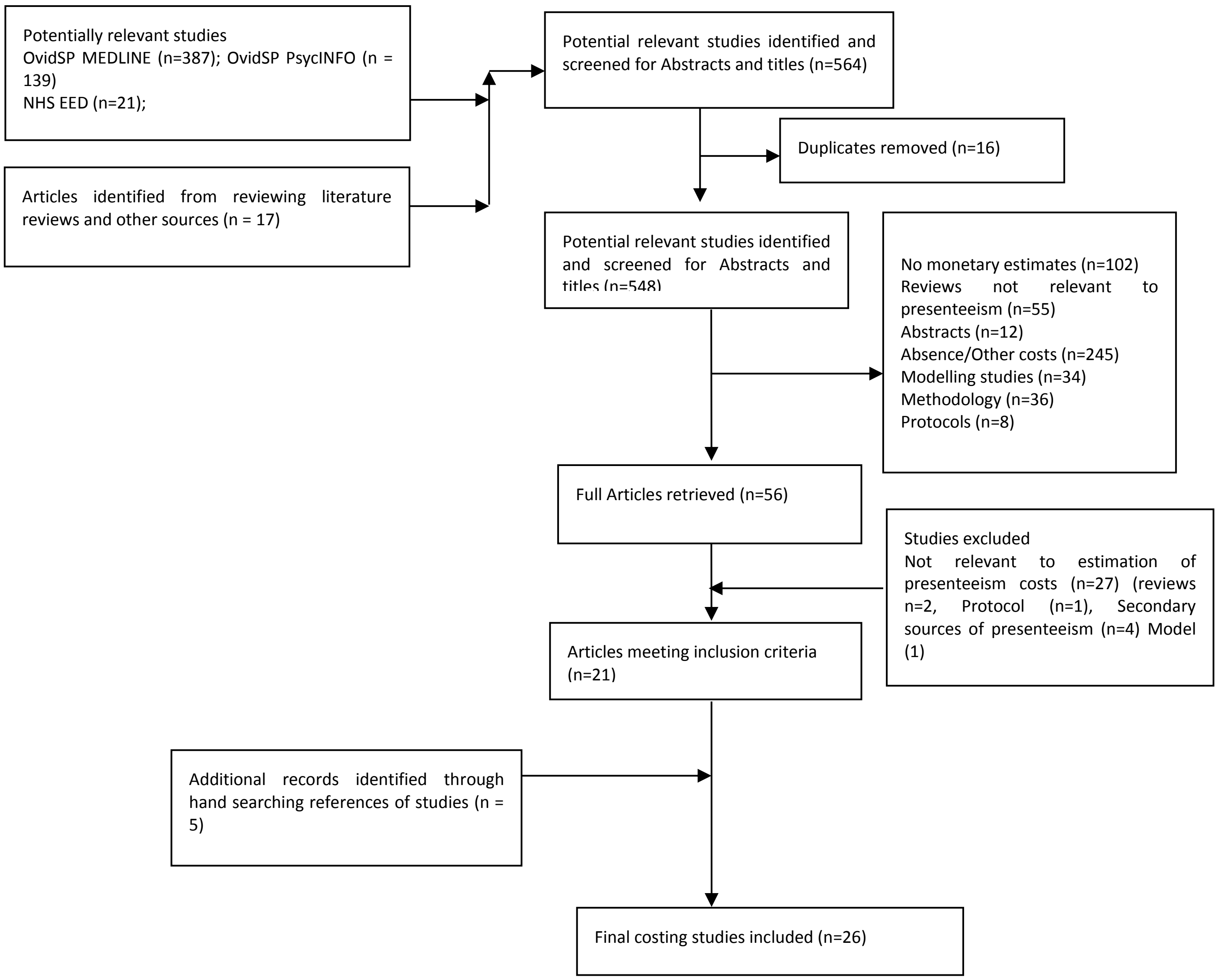


Figure 4:1 Flow diagram showing study selection of assessing presenteeism in economic studies.

4.4 Results

The literature search identified a total of 564 potential articles from the three databases (387 from Medline, 139 from PsycINFO, 21 from NHS EED and 17 through hand searching). As shown in Figure 4.1, 16 of these were duplicates resulting in a total of 548 articles. Details of the selection process are illustrated in the flow chart (Figure 4.1). The screening of the titles and abstracts against the exclusion and inclusion criteria resulted in a total of 56 potentially full articles. The full texts of the remaining 56 potentially relevant papers were obtained and assessed. Of these, 21 articles met the inclusion criteria for the review. Five additional articles were identified through reviewing citations within the identified papers resulting in a total of 26 articles meeting the inclusion criteria. A full list of the studies included is provided in Appendix B.2.

An initial search that was not systematic also identified 12 systematic literature review studies that have evaluated various aspects of presenteeism measurement instruments. From these, a comprehensive list of currently existing presenteeism measures was generated (Table 4.1).

Table 4:1 Summary of productivity loss instruments from literature review

Instrument	Full title	Brooks et al.,2010	Beaton et al.,2009	Lofland et al .,2004	Loeppke et al 2003	Mattke et al.,2007	Nieuwenhuijsen et al.,2010	Prasad et al ., 2004	Ozminkowski et al.,2004	Roy et al.,2011	Schultz et al .,2009	Schultz et al., 2007	Zhang et al.,2010
ALWQ	Angina related limitation at Work Questionnaire	X		X		X					X		
EHCA	Employer Health Coalition of Tampa assessment Instrument						X						
EWPS	Endicott Work productivity Scale	X	X	X		X	X	X			X		
HAQ	Health assessment questionnaire	X									X		
HLQ	Health and Labour Questionnaire	X	X	X		X	X	X			X		X
HPQ	World Health Organisations health and work performance questionnaire	X	X		X	X	X	X			X	X	X
HRPQ-D	Health Related productivity Questionnaire diary	X	X			X	X				X		
HWQ	Health and Work Questionnaire	X	X	X		X	X	X			X		

Instrument	Full title	Brooks et al.,2010	Beaton et al.,2009	Lofland et al .,2004	Loeppke et al 2003	Mattke et al.,2007	Nieuwenhuijsen et al.,2010	Prasad et al ., 2004	Ozminkowski et al.,2004	Roy et al.,2011	Schultz et al .,2009	Schultz et al., 2007	Zhang et al.,2010
MIDAS	Migraine Disability Assessment Questionnaire	X				X		X			X		
MWPLQ	Migraine Work and Productivity Loss Questionnaire	X		X	X	X		X			X		
ORQ	Occupational role questionnaire		X				X						
Osterhaus(OST)	Osterhaus technique	X	X	X		X	X						
Q-Q (Prodisq)	Quantity and Quality method		X				X						
SPS (6,13)	Stanford Presenteeism Scale	X	X	X	X	X	X		X	X	X		
WALS	Work activity limitations scale		X				X			X			
WHI -APA	Work and Health Interview – The American productivity audit	X	X				X		X		X	X	
WIS	Work Instability Scale	X								X			
WRF/WL-26/	Work role and functioning measure scale		X	X			X						

Instrument	Full title	Brooks et al.,2010	Beaton et al.,2009	Lofland et al .,2004	Loeppke et al 2003	Mattke et al.,2007	Nieuwenhuijsen et al.,2010	Prasad et al ., 2004	Ozminkowski et al.,2004	Roy et al.,2011	Schultz et al .,2009	Schultz et al., 2007	Zhang et al.,2010
WLQ	Work Limitations Questionnaire	X	X	X	X	X	X	X	X	X	X	X	X
WPAI	Work productivity and Activity impairment questionnaire	X			X	X	X		X		X	X	X
WPI	Worker productivity index			X		X	X						
WPSI	Work productivity Short Inventory	X	X			X	X		X		X		
WPS-RA	RA – Specific Work productivity survey		X										
WRFQ	Work Role Functioning Questionnaire									X			

4.4.1 Characteristics of the scoping review

4.4.1.1 Frequency

The evaluation of systematic reviews reporting on presenteeism instruments identified 24 instruments (Table 4.1). Each of these instruments was then assessed for frequency of appearance in at least one systematic review. The findings indicated that the most commonly cited presenteeism instrument was the Work limitations Questionnaire (WLQ), appearing in all 12 literature review studies. This was closely followed by the Health and Work Performance Questionnaire (HPQ), Work productivity and Activity impairment questionnaire (WPAI), Health and Labour Questionnaire (HLQ), and Health and Work Questionnaire (HWQ) which appeared in 9, 8, 8, and 7 of the review studies respectively (Table 4.1).

4.4.1.2 Generic versus Disease specific

Three of the instruments including the Migraine Disability Assessment Questionnaire (MIDAS), Angina-Related Limitations at Work Questionnaire (ALWQ), and Migraine Work and Productivity Loss Questionnaire (MWPLQ) were disease specific questionnaires, while the other 21 measures were all generic measures used to assess productivity loss for various health conditions. The Work Productivity Activity Impairment Questionnaire (WPAI) was created as WPAI general health (WPAI-GH) or WPAI specific health (SH) with similar templates only varying in response to the health status being considered for various disease conditions. For this review, the specific questionnaires are all reported as part of the original version of the questionnaire (Table 4.1).

4.4.1.3 Presenteeism versus Absenteeism measurement

The findings showed that only two of the identified instruments, the Work Limitations Questionnaire (WLQ) (e.g. (Lerner et al., 2008)) and Stanford Presenteeism Scale (SPS) (Collins et al., 2005) exclusively measure presenteeism.

4.4.2 The assessment of presenteeism costs in practice

The findings presented thus far have provided an overview of existing instruments measuring presenteeism in the scientific literature. The following section presents and describes the results on how often they have been used in empirical studies.

4.4.2.1 Country setting

Table 4.2 shows an overview of the studies included in the systematic review. These studies originated from eight countries. The majority of the studies identified were from the United States (US). The others were from the Netherlands (Boonen et al., 2010; Uegaki et al., 2011), from Canada (Zhang et al., 2008; Daley et al., 2009; Li et al., 2006), the United Kingdom (Finkelstein et al., 2010), Sweden (Hellgren et al., 2010), and Thailand (Thavorncharoensap et al., 2010).

Two multi-country based studies were also identified. One study reported costing estimates from Australia, US and the United Kingdom (UK) respectively (Hilton et al., 2008). Another study reported cost-estimates from 8 European countries including Germany, Italy, Lithuania, the Netherlands, Luxembourg, Austria, France, and Spain (Linde et al., 2012).

Table 4:2 Summary of economic costing studies from review

No.	Study	Country	Study type	Instruments used/recall	Monetization Method	Productivity Metrics	Primary measure reported	Findings (% - percentage of total)
1	Boonen et al.,2010	Netherlands	COI/Ankylosing spondylitis	HLQ/ 2 weeks	Average wage	Presenteeism and Absenteeism	Extra work hour's needed to compensate for inefficient hours.	Annual presenteeism costs: €967; Absenteeism €1832 per patient per year. Percentage of total not provided.
2	Braakman-Jansen et al .,2012	US	COI/RA	PRODISQ, QQ, WPAI	Average wage-rate per hour/HCM	Presenteeism and Absenteeism	WPAI: Degree of problems affecting work productivity past 7 days on a scale of 0 to 10. PRODISQ/QQ: Quantity and quality of work on an 11-point NRS from 0 to 10)	WPAI; Annual presenteeism costs: 318(73%) and 72(92%) for intervention and control. Annual absenteeism costs: 116(27%) and 6(8%). PRODISQ/QQ: Annual presenteeism: 299 (71%) and 154 (95%) for the intervention and control. Annual absenteeism costs: 120 (29%) and 9 (5%).
3	Burton et al.,2002	US	COI/migraine	Global presenteeism question from interview question	Daily Wage rates/HCM	Presenteeism only	Work days of reduced productivity	Annual presenteeism: $21.5M (60%) Annual Absenteeism: $24.4M (40%).
4	Burton et al.,2005	US	COI/Various health risks	Modified WLQ /2 weeks	Not clear	Presenteeism only	% of time the respondent was limited in performing a specific dimension of job tasks	Annual Presenteeism costs: $1392 to $2592 per employee per year. Annual Extrapolated to $99M to $185M entire population.
5	Cisternas et al., 2003	US	COI/Asthma	Global presenteeism question from survey	Mean Hourly Wage from Census Survey/HCM	Presenteeism only	Reduced work hours due to asthma	Annual costs: $4912, Indirect costs: $1732 (35%). Presenteeism (28%).
6	Collins., 2005	US	COI/ Chronic conditions	SPS and WOS/ 4 weeks	National average wage-rates per job type/ HCM variant	Presenteeism only	Percentage of "usual" productivity not achieved in the 4-week period.	Annual costs per employee: $661 absenteeism, and $6721 Presenteeism. 10% of total labour costs. 6.8% presenteeism.
7	Daley et al., 2009	Canada	COI/Insomnia	Global presenteeism question	Age-gender mean salaries/HCM	Presenteeism only	Extent to which insomnia is responsible for reduced productivity on a 0-10 scale.	Absenteeism $970.6 million Presenteeism $5 billion (76%) Total cost $6.6 billion

No.	Study	Country	Study type	Instruments used/recall	Monetization Method	Productivity Metrics	Primary measure reported	Findings (% - percentage of total)
8	Finkelstein et al., 2010	United Kingdom	COI/obesity	WPAI/1 week	Age-gender specific wage/HCM	Presenteeism and Absenteeism	% reduction in productivity and estimate of time lost during past 7 days.	Presenteeism ($555 to $3792); Absenteeism ($85 to $1026)
9	Fishman and Black ,1999	US	COI/headache	Global presenteeism question /6 months	Age-gender working, educational, mental status specific wage /HCM	Presenteeism only	Degree to which headache affects normal activities on a scale of 0 to 10.	Presenteeism greater than absenteeism. Monetised findings not included.
10	Goetzel et al.,2004	US	COI/10 conditions	Global questions in survey/2weeks, 3 months, 12 months	National Hourly wage rates/ HCM	Presenteeism only	Number of days when the employee was at work but not feeling well, missed hours at work. Rate at which performance was reduced because of health problems.	Annual presenteeism: 61% of total cost in 10 conditions.
11	Goetzel et al.,2010	US	COI/Obesity	WLQ/2 weeks	National average wage-rates/ HCM variant	Presenteeism only	% of time the respondent was limited in performing a specific dimension of job tasks due to obesity.	Annual absenteeism and presenteeism ($2596). Direct costs ($2842). % monetised values not incorporated.
12	Hellgren et al.,2010	Sweden	COI/ Allergic rhinitis and common cold.	Global question from HRA	Self-reported Salary/HCM variant	Presenteeism only	Number of days at work with rhinitis and self-reported productivity while at work during the last month/year	Annual: € 2.7 billion. Presenteeism (37%), Absenteeism (44%).
13	Henke et al.,2000	US	COI/ PUD and GERD	General question as part of interview Questionnaire/ 3 months.	Self-reported Salary/HCM variant	Presenteeism only	Reduced productivity because of PUD or GERD	Presenteeism: Annual PUD costs per year $205 (28%), Annual GERD $72 (27%).
14	Hilton et al.,2008	US, UK, AUSTRALIA	COI//psychological distress	HPQ/4 weeks	Mean Wage-rates ONS from UK and Australia; / HCM	Presenteeism and Absenteeism	Performance of an average person working in a similar job to patient as well as patients own on a self-anchoring scale of performance of 0 to 10 (worst to best).	Annual total costs USD$11.1 billion. Presenteeism % not provided.
15	Lamb et al., 2006	US	COI/Allergic Rhinitis	WPSI/Not reported	Standard hourly wage/HCM	Absenteeism/Presenteeism	Number of unproductive hours spent at work during the recall period.	No monetary details provided of total costs.
16	Lerner et al., 2008	US	CEA/Fibroids	WLQ/2 weeks	Average wage/HCM	Presenteeism only	% of time the respondent was limited in performing a specific dimension of job tasks (%) due to Fibroids.	Annual Presenteeism: $2341 for intervention group; $836 for control group. Annual Absenteeism: $2044 for intervention group and $540.2 for

No.	Study	Country	Study type	Instruments used/recall	Monetization Method	Productivity Metrics	Primary measure reported	Findings (% - percentage of total)
								control group. % of total not provided.
17	Li et al., 2006	Canada	COI/Arthritis	WLQ/2 weeks	Annual-average wage-rate/HCM	Presenteeism only	% of time the respondent was limited in performing a specific dimension of job tasks (%) due to Arthritis.	Total Annual costs: $11,553 Presenteeism: $4724 (41%)
18	Linde et al., 2012	8 European countries	COI/Headache	General presenteeism question	Average–gender specific wage-rate/HCM	Absenteeism and Presenteeism	days at work when the amount done was ≥ 50% reduced productivity counted as 1 day of reduced productivity).	Annual cost per person: £ 1222; Absenteeism: £371 Presenteeism: £ 765 (63%).
19	Ricci and Chee, 2005	US	COI/Obesity	WHI	Self-reported salary /HCM variant	Presenteeism and Absenteeism	average amount of time between arriving at work and starting work on days when a worker is not feeling well and the average frequency of engaging in five specific work behaviours (ie, losing concentration, repeating a job, working more).	Annual cost: $11.70 billion per year. Presenteeism 67% total costs.
20	Smit et al., 2006	Netherlands	CEA/Depression	Global questions	Age-gender wage-rate	Presenteeism only	number of work cut-back days as the number of days actually worked when ill, multiplied by a self-reported inefficiency score, which ranged between 0 and 1 (0, as efficient as when in good health; 1, totally inefficient).	Annual presenteeism, intervention: €2232(33%); Annual total costs: €6766; Annual presenteeism ,control: €3175(39%); Annual total costs: €8614;
21	Stewart et al.,2003a	US	COI/Headache, back pain, Arthritis.	WHI	Self-reported Salary/HCM variant	Presenteeism and Absenteeism	Reduced work hours on days at work during the recall period quantified based on responses to 6 questions.	Total cost $61.2 billion, absenteeism: $14.4 bn, presenteeism: $46.9 bn (76.6%). Presenteeism for Arthritis (84.4%) and Back pain (69.7%).
22	Stewart et al.,2003b	US	COI/Headache, back pain, Arthritis.	WHI	Self-reported Salary/HCM variant	Presenteeism and Absenteeism	Reduced work hours on days at work during the recall period quantified based on responses to 6 questions.	Total productivity costs: $225.8 billion per year. Most costly conditions included pain (headache, low back pain, Arthritis). On average, Presenteeism 71% of total costs.
23	Thavorncharoensap et al., 2010	Thailand	COI/ Alcohol consequences	Questions from WPAI 1 week	Average income per year	Presenteeism and Absenteeism	The degree to which the health problem affected regular activities evaluated on a scale of 11 points, ranging from 0 (no effect on work) to 10 (health problem prevent person from working).	Annual Total costs: $ 9,627 million Annual Presenteeism: $ 2,804 million (29% total). Mortality costs: $6,422 million.
24	Uegaki et al., 2011	Netherlands	COI/Maternity	HPQ/2 weeks	Not included/ HCM	Presenteeism and	Decreased work performance due to a health problem at 18-, 24- and 52-weeks	Annual presenteeism, intervention: €765(40%); Annual total costs: €1911;

No.	Study	Country	Study type	Instruments used/recall	Monetization Method	Productivity Metrics	Primary measure reported	Findings (% - percentage of total)
						Absenteeism	post-partum.	Annual presenteeism, control: €655 (38%); Annual total costs: €1734. Overall costs, Indirect costs (37%) presenteeism (52%).
25	Wilson et al.,2010	US	COI/Acute attacks/ Hereditary angioedema (HAE).	WPAI-GH included in survey/7 days	Self-reported Gross Salary/HCM variant	Presenteeism and Absenteeism	Hours actually worked; degree illness affected productivity while working from 0 (no effect) to 10 (maximum impairment); and the degree to which illness affected regular activities (from 0-10).	Total costs: $41,992 Indirect Costs: $16,108.Absenteeism: $3402.Presenteeism $5,750. Presenteeism (14% total costs).
26	Zhang et al.,2008	Canada	COI/RA	HLQ-2 weeks/HPQ- 4 weeks/WPAI-1 week, WLQ/2 weeks.	Age-gender employee type specific wage-rate /HCM	Presenteeism and Absenteeism	**HLQ**: single question asking how many extra hours individuals would have to work to catch up on tasks they were unable to complete in normal working hours due to health problems in the past 2 weeks **WPA**I: impairment while working due to health problems in the past 7 days are measured on a scale of 0–10. **WLQ**: a 25-item questionnaire asking about the frequency of difficulty over the past 2 weeks over 4 domains of work: time management, physical demands, mental-interpersonal, and output demands. HPQ: overall performance on the days they worked during the past 7 days on a scale of 0–10, with 0 indicating total lack of performance during time on the job and 10 indicating no lack of performance during time on the job.	$30.03, $83.05, $284.07, and $285.10 (HLQ, WLQ, HPQ, WPAI) over 2 a period of weeks.

4.4.2.2 Disease area

In general, a range of diseases were covered by the studies and the articles varied from national survey based costing studies covering various conditions to cost estimates from specific disease conditions. The most common conditions included obesity (Ricci and Chee, 2005; Goetzel et al., 2010; Finkelstein et al., 2010), rheumatoid arthritis (Zhang et al., 2008; Li et al., 2006; Stewart et al., 2003b), migraine (Linde et al., 2012; Fishman and Black, 1999; Burton et al., 2002) and insomnia (Daley et al., 2009). Details of the diverse set of conditions included are reported in Table 4.2.

4.4.2.3 Economic analysis approach and data context

The majority of the studies were cross-sectional in design, but two were clinical trials (Smit et al., 2006; Uegaki et al., 2011). Three of the studies were cost-effectiveness analyses (Snedecor et al., 2009; Smit et al., 2006; Uegaki et al., 2011) while the other twenty-five were all cost-of- illness studies (Table 4.2).

4.4.2.4 Type of instrument

In total, nine presenteeism instruments were identified from the final 28 studies included in the second stage of the systematic review (See Table 4.3). Reduced productivity loss was measured by a self-constructed questionnaire or visual analogue scale or standardised questionnaire. The most commonly used standard questionnaire was the WLQ (Li et al., 2006; Goetzel et al., 2010; Lerner et al., 2008), followed by the WHI (Ricci and Chee, 2005; Stewart et al., 2003a; Stewart et al., 2003b), and the WPAI (Thavorncharoensap et al., 2010; Finkelstein et al., 2010; Wilson et al., 2010). Other currently used multi-question instruments included the SPS (Collins et al., 2005) , HLQ (Boonen et al., 2010), PRODISQ (QQ) (Braakman-Jansen et al., 2012) and HPQ (Hilton et al., 2008). The remaining studies used a self-constructed global presenteeism question to identify the value of reduced productivity, based on a global response 0 – 10 scale (Linde et al., 2012; Fishman and Black, 1999; Daley et al., 2009; Burton et al., 2002; Goetzel, 2004; Cisternas et al., 2003; Hellgren et al., 2010; Henke et al., 2010) (Tables 4.2,4.3). One study used a modified version of the WLQ (Burton et al., 2005)

Table 4:3 Summary of productivity loss instruments identified from the systematic review assessing presenteeism costs in practice

No.	Instruments used/recall	Productivity metrics assessed	Study
1	Health and Labour Questionnaire (HLQ)	Presenteeism and Absenteeism	Boonen et al.,2010; Zhang et al.,2008
2	Quantity and Quality method (PRODISQ (Q-Q))	Presenteeism and Absenteeism	Braakman-Jansen et al .,2012
3	Work productivity and Activity impairment questionnaire (WPAI)	Presenteeism and Absenteeism	Braakman-Jansen et al .,2012; Finkelstein et al., 2010; Thavorncharoensap et al., 2010; Wilson et al.,2010; Zhang et al.,2008
4	Stanford Presenteeism Scale (SPS)	Presenteeism only	Collins., 2005
5	Work Limitations Questionnaire (WLQ)	Presenteeism only	Goetzel et al.,2010; Lerner et al., 2008; Li et al., 2006; Burton et al.,2005
6	World Health Organisations health and work performance questionnaire (HPQ)	Presenteeism and Absenteeism	Hilton et al.,2008; Uegaki et al., 2011; Zhang et al.,2008
7	Work productivity Short Inventory (WPSI)	Absenteeism and Presenteeism	Lamb et al., 2006
8	Work and Health Interview (WHI)	Presenteeism and Absenteeism	Ricci and Chee 2005, Stewart et al.,2003a; Stewart et al.,2003b
9	Global presenteeism question	Presenteeism only	Burton et al.,2002, Cisternas et al., 2003, Daley et al., 2009; Fishman and Black ,1999; Goetzel et al.,2004; Hellgren et al.,2010; Henke et al.,2000; Smit et al., 2006

4.4.2.5 Productivity loss conversion and monetisation approaches.

Measurement approaches to presenteeism varied among the identified studies. Productivity metrics used by the different studies included estimations of percentage reduced productivity, hours needed to compensate for reduced productivity, and overall performance assessed using a 0-10 scale. Studies using the WPAI, HLQ and HPQ instruments generated productivity loss estimates that were directly translatable into productivity loss, while the estimates from the SPS, WLQ, WHI, and QQ could be transformed into reduced productivity loss after making some assumptions. In general the approaches used by the different studies are summarised in three approaches including direct estimation of productivity loss in hours, estimation of perceived percentage loss, and comparative approaches of estimates between individuals and employees.

The direct approach generates productivity loss values in a similar way to the approach used in obtaining absenteeism productivity loss. A typical question involves asking respondents to estimate “the total hours you were unproductive because of an illness”. The duration considered as unproductive can sometimes incorporate different work-related aspects. For example using the WHI, presenteeism has been estimated from a combination of questions such as the average number of hours with low concentration at work, when working more slowly than usual, when feeling fatigued at work, and as the time in between arriving at work and starting work on the days when an employee is sick (Ricci and Chee, 2005). Alternatively respondents are asked to estimate the extra hours that would be needed to compensate for inefficient hours as used by the HLQ (Zhang et al., 2008; Boonen et al., 2010). The obvious ease of this approach is that it generates a directly usable productivity metric in lost hours. Comparisons with other presenteeism measures suggests this approach underestimates values (Zhang et al., 2010b).

The second approach involves asking respondents to provide a perceived overall estimate of how much illness has hindered or affected their performance at work. This was the most common approach (17 (80%) studies). With these studies the productivity metrics used included asking respondents to provide a percentage loss of productivity at work due to illness (Braakman-Jansen et al., 2012; Goetzel et al., 2010). For example articles using the WLQ, obtained an estimate of the percentage presenteeism loss (or gain) from respondents compared to a baseline or benchmark value for each Individual (Lerner et al., 2008; Goetzel et al., 2010). An alternative version of this approach involved asking respondents to provide an estimate of how illness has affected their performance at work on a 0 –10 scale which was then converted into a percentage productivity loss (Finkelstein et al., 2010; Wilson et al., 2010). Studies using the WLQ (Li et al., 2006; Goetzel et al., 2010; Lerner et al., 2008) and SPS (Collins et al., 2005) also assessed perceived limitations in different work function domains and for different work aspects. The output from these different domains is often then summarised to generate an index which is interpreted as a percentage loss attributed to reduced productivity. Conversion of outputs from presenteeism measures assessing productivity loss for different work domains remains a complex process. For example, the process involved in translating multi-item scores from the WLQ and SPS into an overall score involves converting the range of responses from the different domains into percentages using algorithms that are not always very clearly reported (Lofland et al., 2004; Brooks et al., 2010).

In the remaining studies an estimate of perceived reduced productivity was estimated using non-standard stand-alone single-item questions. Such questions were often used as part of a wider questionnaire with a global question asking respondents to either estimate perceived impairment on a scale of 0-10 or percentage reduction at work due to illness (Linde et al., 2012; Fishman and Black, 1999; Daley et al., 2009; Burton et al., 2002; Goetzel, 2004; Cisternas et al., 2003; Hellgren et al., 2010; Henke et al., 2010). However, the validity of such questions when used as part of a general survey questionnaire of the respective disease conditions has not been established.

The third approach from the HPQ questionnaire, found in three studies (Hilton et al., 2008; Uegaki et al., 2011; Zhang et al., 2008), involves comparing presenteeism estimates of a respondent with those of a colleague in a similar work role or of the respondent in full health. Normally, this will be done on a on a scale of 0 (worst performance at work) to 10 (best performance).

4.4.2.6 Monetisation approaches

The generation of presenteeism monetary estimates from all the studies identified was based on salary conversion approaches. More specifically, with the exception of Smit et al. (2006) who used the friction cost approach, the remaining 25 studies explicitly or implicitly applied the human capital method using the standard formula: $time\ lost\ \times wage\ rate$.

A variety of measures were used to assess the value of foregone earnings, and these included: an average wage for all groups (Lerner et al., 2008; Goetzel et al., 2010; Collins et al., 2005; Boonen et al., 2010; Hilton et al., 2008; Cisternas et al., 2003; Braakman-Jansen et al., 2012; Li et al., 2006; Thavorncharoensap et al., 2010), age-sex dependent wage-rates (Daley et al., 2009; Finkelstein et al., 2010; Fishman and Black, 1999; Linde et al., 2012), and a self-reported gross salary (Ricci and Chee, 2005; Henke et al., 2000; Stewart et al., 2003b; Wilson et al., 2010) (Table 4.2). The wage-rates used differed in terms of being hourly, annual or daily wage-rates.

None of the studies adjusted presenteeism costs for aspects of compensation mechanisms, multiplier effects, or included any additional costing elements in the presenteeism calculation.

4.4.3 Monetary estimates of presenteeism in current practice

The level of detail provided on presenteeism cost estimates varied across studies (Table 4.2). In total, findings from the review studies show that economic costs of presenteeism were reported either in monetary terms or as a percentage of the overall total cost of the illness. The results demonstrate that presenteeism is a major economic burden in many health conditions. In most studies presenteeism costs accounted for a greater percentage of total costs than absenteeism or sick leave, and comprised a significant proportion of the total costs of illness reported.

Seventeen studies reported productivity costs as a percentage of the overall total costs of the disease condition investigated (Ricci and Chee, 2005; Smit et al., 2006; Henke et al., 2000; Fishman and Black, 1999; Collins et al., 2005; Cisternas et al., 2003; Hellgren et al., 2010;

Stewart et al., 2003a; Stewart et al., 2003b; Wilson et al., 2010; Goetzel, 2004; Burton et al., 2002; Daley et al., 2009; Braakman-Jansen et al., 2012; Li et al., 2006; Linde et al., 2012; Thavorncharoensap et al., 2010; Uegaki et al., 2011). In the remaining articles, it was not possible to extract reduced productivity costs as a percentage of overall costs due to a limited level of detail for the cost items. As can be seen in Table 4.2, productivity costs reflected on average 52% (ranging 28% to 85%) of the total costs of the disease conditions investigated.

In studies that provided a limited level of detail for percentage productivity costs in relation to overall total costs, five included costs that enabled a comparison between absenteeism and presenteeism (Lerner et al., 2008; Collins et al., 2005; Boonen et al., 2010; Finkelstein et al., 2010; Linde et al., 2012). In these studies, productivity loss estimates from presenteeism were greater than absenteeism costs.

The three cost-effectiveness studies within this review included productivity losses related to presenteeism, but did not provide enough detail to assess the impact of presenteeism on incremental costs (Smit et al., 2006; Uegaki et al., 2011; Lerner et al., 2008).

4.5 Discussion

The present systematic review was designed to identify studies that have assessed presenteeism costs in practice. Presenteeism measures that have been used to generate presenteeism costs were identified from both cost-of-illness studies and economic evaluations. The systematic review sought to provide insights into existing presenteeism instruments, and to assess variations and gaps in the estimation of presenteeism costs.

The results from this systematic review show variations in approaches for estimating monetary estimates in existing presenteeism instruments. Another important finding was that a number of studies applied global, single-item, standalone presenteeism questions. The lack of validity information from the studies using global presenteeism questions shows the importance of investigating whether such questions provide valid responses when used outside standardised questionnaires. When reported, the economic burden associated with presenteeism was clearly highlighted with significant costs associated with various illnesses. In most cases presenteeism costs were greater than those for absenteeism and comprised the majority of the total costs in most studies.

The systematic review provides further insights into the most commonly applied instruments in the economic studies identified. With regards to measurement, the current systematic review found that there was variation in the approach used to generate presenteeism loss estimates, and in recall periods. Studies used different quantification approaches, and wage-rate values to generate monetary estimates. It is interesting to note that, given the significant costs associated with presenteeism related multiplier effects and possible compensation mechanisms, no studies considered these costs. Moreover, the majority of studies used the human capital method when valuing presenteeism. Only one study from the Netherlands used the friction cost approach, but did not provide any detail of how this application was applied (Smit et al., 2006).

Nine standardised presenteeism instruments used to assess presenteeism costs were identified from the review. Each of the instruments identified is associated with its own advantages and limitations. For example, direct estimation of time loss enables a direct translation of time loss output into productivity costs, but creates challenges in identifying appropriate ways of accurately deriving time loss related to aspects of unproductive time and extra-hours. In terms of practical application, the approach of deriving perceived time loss while working has often been used because of the ease in obtaining estimates from respondents. The drawback with this approach is that the presenteeism loss estimates generated cannot be compared against any benchmark values as in the case of comparative presenteeism measures. Moreover, multi-item and multi-dimensional presenteeism instruments in particular are associated with increased responder burden, time and costs which could potentially limit consideration of presenteeism costs, suggesting that a single-item questionnaire may be of value. The validity of non-standardised single-item presenteeism questions as identified in this systematic review has not been explicitly reported in the studies in which they are normally used.

The findings of this systematic review further support the idea of the general lack of consensus about the most appropriate presenteeism measure. The findings of this review are consistent with those of Schultz et al., (2009) who found wide variations in the monetary valuation approaches for on-the-job reduced productivity. Few studies (Ozminkowski et al., 2004; Meerding et al., 2005; Zhang et al., 2010b) have attempted to cross compare measurement instruments within the same population, and these have found varying presenteeism estimates from these instruments. Little is therefore known about the most appropriate method for generating monetary estimates of presenteeism.

The systematic review here generated results from the valuation of presenteeism that corroborated the findings of a previous methodological systematic review that found no attempt to apply the friction cost approach to US studies (Mattke et al., 2007). Relatively similar findings were reported in the systematic review in this chapter, as only one study from the Netherlands generated presenteeism costs using the friction cost approach. The remaining studies all generated presenteeism costs using the human capital method. The strengths and limitations associated with this method were discussed in Chapter Two. Notably, however, the method ignores internal work structures aimed at mitigating productivity losses, as well as multiplier effects resulting from effects of presenteeism on overall team productivity. In such cases, the human capital method might generate either over or underestimated productivity costs. Empirical research exploring how to extend the friction cost approach to valuing presenteeism is necessary.

To date, little evidence exists on presenteeism cost estimates in the UK. The majority of studies included in this review were cost-of-illness studies, indicating gaps in research on economic evaluation studies incorporating presenteeism costs and their impact on overall cost-effectiveness results. These findings could possibly be explained by a recent review of national economic guidelines clearly showing that the inclusion of presenteeism and work compensation costs in economic analysis is not recommended for the majority of national economic guidelines (Knies et al., 2010).

Strengths and limitations

The systematic review presented and discussed in this chapter has some strengths and limitations. One strength is that it provides a comprehensive review of studies that have assessed presenteeism costs in practice, highlighting how often presenteeism instruments have been used in empirical studies. Moreover, the review is based on databases of health economics, medical and behavioural science disciplines, hence providing a comprehensive overview of existing cost-of-illness and original economic evaluation studies in the literature. It provides important information on how presenteeism costs are assessed in practice. A limitation of the review is that although major databases were comprehensively searched, it is possible that eligible studies may have been omitted. In addition, modelling studies and studies not in English were excluded from this review, which may potentially have contributed to the small number of economic evaluation studies included in the review. These studies were excluded because the focus of the review was to assess how reduced productivity data is collected, measured and valued in primary research.

Implications for policy and research

Given the on-going debates in the measurement and valuation of presenteeism in economic evaluations, these findings provide valuable information for researchers and policy makers.

One of the issues that emerge from this review is that there is a scarcity of economic evaluation studies incorporating aspects of presenteeism, the effects of presenteeism on team productivity, and effects of work compensation mechanisms. In light of the significant costs associated with reduced productivity, economic evaluation results could be severely underestimated where effects of reduced productivity and associated work-related compensation mechanisms are not taken into account. The results indicated that productivity costs ranged from 28% to 85% of the total costs associated with various diseases, and therefore deserve more attention. Regardless of the lack of a standardised method estimating presenteeism related costs, empirical economic evaluation studies including presenteeism costs are needed. Increased transparency in the calculations and methods used in quantifying and generating these monetary estimates should be encouraged to ensure sound and rigorous methods are used.

The results from this review provide some support for exploring applications of single-item global presenteeism questions as they may offer potential advantages in collecting presenteeism estimates at reduced costs, with more simplicity and reduced burden. However, more research on these measures needs to be undertaken as their validity as standalone measures has not been identified. No attempts have been made to compare output from commonly used standardised questionnaires and standalone single-item presenteeism questionnaires. Further research on this area is also recommended.

A further key issue that emerges from this review is the limited application of the friction cost approach in valuing presenteeism. This is an important issue for further research to foster consistency of methods used to estimate absenteeism. Studies investigating applications of the friction cost approach in valuing presenteeism could take into account the impact of incorporating compensation mechanisms and multiplier effects due to presenteeism. Other potential areas of research, although beyond the scope of this thesis, include investigating the impact of different recall periods on presenteeism estimates when measuring presenteeism and extrapolation approaches for productivity loss estimates obtained in short-term periods.

4.6 Conclusion

In conclusion, there is potential for further methodological and empirical developments in measuring and valuing presenteeism. Currently, the number of studies including presenteeism costs in economic evaluation remains small and monetary generation approaches vary widely. The findings here have provided a systematic review assessing presenteeism costs in practice. They have revealed important gaps in measurement, valuation and incorporation of these costs in economic evaluations and the significant economic burden associated with presenteeism. One of the most important issues is the lack of any evidence on the validity and responsiveness of a single-item presenteeism question in specific disease conditions, and the general lack of the impact of presenteeism related costs on cost-effectiveness results. Obtaining better evidence on these areas is thus one of the major foci of this thesis.

CHAPTER FIVE AN OVERVIEW OF THE CLINICAL AREA

5.1 Introduction

This chapter provides an overview of the clinical area of low back pain. The chapter explores the nature, aetiology, prevalence and incidence of the condition as well as the risk factors associated with its occurrence and development. The approaches used to manage non-specific low back pain are discussed, with reference to current national guidelines. A review of literature on the management of non-specific low back pain and evidence for the cost-effectiveness of existing interventions is then provided and discussed. The economic burden of the illness is presented, with a particular focus on total societal costs in the UK and internationally. Finally, an overview of the data sets to be used in this thesis is provided.

5.2 What is Low back pain?

Back pain is a common disorder affecting nearly five million people each year within the UK and is one of the most common reasons for general practice consultations (Hay et al., 2008; Foster et al., 2010). The condition is associated with significant resource use, absenteeism, and long-term disability, with serious implications for individuals, families and society at large (van Tulder and Waddell, 2005).

Low back pain is defined as pain, discomfort, stiffness in the lower back region with or without referred leg pain (van Tulder et al., 2002; NICE, 2009a). Low back pain is classified into either non-specific (simple) or specific back pain. Non-specific low back pain is pain, soreness, tension or stiffness in the lower back that cannot be associated with any specific identifiable cause (Weiner and Nordin, 2010; NICE, 2009a). Most often, non-specific low back pain is characterised by disability and pain (Koes et al., 2006). Simple back pain normally lasts for a period of between 6 weeks and 12 months. On the other hand, specific low back pain is associated with particular identifiable conditions such as hernia, osteoporosis, and rheumatoid arthritis (Krismer and van Tulder, 2007; Cohen et al., 2008). The majority of low back pain sufferers experience some form of non-specific low back pain (Koes et al., 2006). This thesis is based on two datasets reporting the outcome and cost data on low back pain sufferers experiencing non-specific low back pain within a primary care setting.

5.3 Epidemiology of low back pain

5.3.1 Types of non-specific low back pain (NSLBP)

There are three types of non-specific low back pain: acute, sub-acute and chronic. This categorisation relates to the duration of the condition (Krismer and van Tulder, 2007). Acute non-specific LBP is pain that lasts for less than 6 weeks while sub-acute non-specific LBP lasts for between 6 weeks and 12 weeks (van Tulder et al., 2002). Back pain that persists beyond 12 weeks is classified as chronic LBP (van Tulder et al., 2002). Chronic non-specific LBP sufferers, if not managed properly, can end up with significant disabling long-term pain and disability (Adams, 1997). In most cases, however, non-specific low back pain resolves quickly, with 90% of patients ceasing consultations within 4 - 12 weeks (Croft, 1998).

5.3.2 Prevalence of low back pain

Low back pain is a very common problem among adults in all age groups and prevalence rates tend to increase with age and among working age adults (van Tulder et al., 2002;

Dionne et al., 2006). The prevalence however decreases in population groups after retirement (Dionne et al., 2006).

There is no general consensus on the prevalence, although many studies have been conducted to assess prevalence rates, with similar findings from different countries and settings. Differences in prevalence rates arise from the definition of low back pain used, the study design, and the way in which the data was collected (Jones and Macfarlane, 2005). Moreover, there are inherent challenges in estimating incidence rates of low back pain since the cumulative incidence of first back pain episodes is often high by early adulthood (Dionne et al., 2006). As such, the majority of studies have investigated prevalence of low back pain as either a period (e.g., one, two years or life time) or point estimates (i.e., at a specific time point) (Hoy et al., 2010b).

Estimates of point, one year and life-time prevalence vary considerably between studies and countries. International comparisons show point prevalence rates ranging from 12% to 30%, 12-month prevalence ranging from 50% to 67%, and life-time prevalence ranging from 10% to 80% (Cassidy et al., 1998; Koes et al., 2006; Mortimer et al., 2006; Ihlebaek et al., 2006; Hoy et al., 2012). Low back pain has been reported to be more prevalent among females and adults aged 40 years and over (Hoy et al., 2012). Prevalence rates also tend to vary by country. For example, the point prevalence was 50% in France (Rossignol et al., 2009), 68% in Australia (Walker et al., 2004), and 34-39% in the UK (Hillman et al., 1996). The life-time prevalence was estimated at 61% in the UK (Hillman et al., 1996), 20-69% in Canada (Cassidy et al., 1998), 61% in Norway and 70% in Sweden (Ihlebaek et al., 2006).

In the UK population one-year prevalence was on average 40% (Hillman et al., 1996; DOH, 1999), one month prevalence was 29% (Macfarlane et al., 2012), and point estimates 19% to 40% (Hillman et al., 1996; Macfarlane et al., 2012). Two UK studies suggest there are no significant differences in prevalence rates between men and women, with prevalence rates of 29% and 27% (Macfarlane et al., 2012), and 34% and 37% (Croft, 1998) respectively. The high prevalence rates from the various studies and countries demonstrate the importance of the condition.

5.3.3 Incidence of low back pain

The estimates from literature show that one year incidence rates among patients reporting a first ever episode of low back pain ranged from 6% to 15%, while those for any episode ranged from 2% to 36% (Hoy et al., 2012). Studies specific to the United Kingdom have reported first ever episode incidence rates ranging from 3% to 6% (Papageorgiou et al., 1996; Croft, 1998; Waxman et al., 2000).

5.3.4 Identification of low back pain

One of the main challenges in back pain management is the uncertainty related to the diagnosis of low back pain patients, particularly in the case of non-specific low back pain (Deyo and Phillips, 1996). The assessment of back pain involves the use of a classification system focussed on screening for 'red flags' (indicators for potential serious conditions), neurological problems, and identification of possible psychosocial and workplace risk factors (van Tulder et al., 2002). This classification system enables the distinction between specific or non-specific low back pain conditions in order to determine appropriate management strategies of low back pain and to identify the most appropriate options for treating individuals (Weiner and Nordin, 2010). Box 5.1 below provides a summary of the most commonly identified red flags in low back pain sufferers.

Box 5.1: Red flag conditions indicating possible underlying spinal pathology or nerve root problems

Red flags

- Non-mechanical pain (unrelated to time or activity)
- Non-responsive to care
- Thoracic pain
- Widespread neurological symptoms
- Weakness in limbs
- Pain at multiple sites
- Previous history of carcinoma, steroids, HIV
- Feeling unwell
- Unexplained weight loss
- Structural spinal deformity
- Osteoporosis

Indicators for nerve root problems

- Unilateral leg pain
- Severe lower extremity pain
- Radiates to foot or toes
- Abnormal deep tendon reflexes
- Numbness and paraesthesia in same distribution
- Straight leg raising test induces more leg pain
- Localised neurology (limited to one nerve root)

Modified from Koes et al., 2006 and Cohen et al., 2008

5.3.5 Aetiology of low back pain

In addition to prevalence, incidence and identification, epidemiological researchers have provided evidence examining the risk factors associated with occurrence and chronicity of low back pain (Macfarlane et al., 2006). Current models attempting to identify the cause of low back pain involve an interaction of various factors that can be classified into individual, psychosocial and occupational factors as presented and discussed below.

5.3.5.1 Individual factors

Although results from studies tend to vary, research demonstrates that the most common individual factors associated with development of low back pain include social demographic factors such as age, gender, occupation, and education status. Other individual factors reported in the literature include smoking, levels of physical activity, obesity, pain radiating in leg, and poor strength of back muscles (Macfarlane et al., 2006; van Tulder et al., 2002). The initial occurrence of LBP often takes place between the ages of 20 and 40 (Adams, 1997).

The literature indicates that the evidence relating to smoking is inconclusive with some finding weak associations between smoking and the occurrence of low back pain (Leboeuf-Yde, 1999) and others finding strong associations, particularly among young adults (Shiri et al., 2010). Similarly, the evidence supporting gender differences in LBP sufferers is relatively mixed with some epidemiological studies reporting higher LBP estimates among women (McIntosh et al., 2000; Macfarlane et al., 2006), and others finding weak evidence to support gender differences among LBP sufferers (Turner et al., 2000; Hayden et al., 2005).

5.3.5.2 Psychosocial factors

The relationship between back pain, psychological and psychosocial factors has been investigated in epidemiological studies investigating predictors of low back pain onset or poor outcomes (Breen et al., 2005; Macfarlane et al., 2006). Studies show that the most common psychosocial risk factors leading to LBP occurrence and long-term back pain conditions include avoidance of activities based on fear (fear avoidance), depression, beliefs about pain, emotional instability, anxiety, cognitive dysfunction, mental distress and long-term multiple physical disorders (somatisation). Fear avoidance has been found to be a key psychological factor significantly associated with disability in long- term back pain conditions (McIntosh et al., 2000; Pincus et al., 2002; Grotle et al., 2010).

There is also evidence showing that depression and distress increase the risk of long-term LBP disability, especially in the early stages (Pincus et al., 2002). Moreover psychological distress has been linked to the persistence and development of chronic back pain (Linton and Nordin, 2006), as well as to delays in returning to work among low back pain sufferers (Crook et al., 2002). Further evidence among primary care consulters in the UK has shown that having negative perceptions about the low back pain condition, depression and distress are significant predictors of poor low back pain outcome in the long run (Foster et al., 2008).

5.3.5.3 Occupational factors

Low back pain injuries have been associated with disability, lost work days and significant cost to the work place (Freburger et al., 2009). As a result, researchers have focused on identifying various aspects of work and the work environment that are likely to influence the risk of back pain development. The most common occupational risk factors associated with individuals developing low back pain include physical jobs that involve frequent bending, twisting, physically heavy manual work, sedentary work, static postures and vibrations (Punnett and Wegman, 2004). The evidence shows increased risk of long-term absence from work resulting from physical factors of lifting, carrying, and pulling heavy weights and loads within the workplace setting (Harkness et al., 2003; Thomas et al., 2006).

Occupational psychological related factors such as job stress, job dissatisfaction, unemployment and lack of social support at work are also associated with an increased risk of the future occurrence of back pain as well as persistence in low back pain (Hoy et al., 2010a; Shaw et al., 2009). Moreover, persistent low back pain has been strongly associated with high job demands, dissatisfaction at the work place (Thomas et al., 1999; Macfarlane et al., 2009), and with unemployment among low back pain sufferers (Dunn et al., 2011).

5.4 Management of low back pain

5.4.1 Introduction

The management of low back pain is a key aspect of reducing the economic burden of illness to society. Currently, there is no clear cut approach for the treatment and management of non-specific low back pain, particularly in individuals with pain and disability lasting more than 6 weeks (NICE, 2009a). There is, however, more clarity about the management of acute low back pain (i.e. pain lasting less than six weeks) among back pain patients in the existing national treatment guidelines. In most cases, the strategy employed aims to ensure prevention and targeted management of the condition in order to prevent long-term disability.

The 'stepped up' care model has been adopted into most of national clinical guidelines (Koes et al., 2006; Airaksinen et al., 2006; NICE, 2009a). This model was proposed by Von Korff as a system for the management of low back pain in primary care (von Korff, 1999). The theory underlying this model is that patients are first introduced to the low-intensive, low-cost interventions, and then moved onto to more intensive, costly or complex interventions depending on their response to treatment (Von Korff and Moore, 2001).

5.4.2 An overview of current national treatment guidelines

The development of national and international guidelines for clinical management of low back pain (Airaksinen et al., 2006; van Tulder et al., 2006a; Chou et al., 2007; NICE, 2009a; Koes et al., 2010) has been guided by evidence from the numerous randomised clinical trials, and cost-effectiveness studies that have been conducted on the management of the illness. In most of these guidelines, the recommendations for managing each condition have involved applications of the Von Korff 'stepped up' care model (Von Korff and Moore, 2001). In brief, the current available treatments for non-specific low back pain from the national guidelines cited above can be summarised as:

- Early management which involves advising back pain sufferers to stay active and the provision of evidence-based educational materials for self-management of LBP;
- In addition, individuals are also provided with medications such as paracetamol, non-steroidal anti-inflammatory drugs (NSAIDs), opioids as well as antidepressants and muscle relaxants if deemed necessary;
- For persistent non-specific low back pain, alternative interventions such as spinal manipulation, exercises, physiotherapy, mobilisation, cognitive behaviour therapy and acupuncture are made available.

The consensus in most national guidelines in managing chronic or long-term LBP is to consider interventions such as combined behavioural therapy, invasive procedures and, in some cases, surgical treatment (van Tulder et al., 2006a; Airaksinen et al., 2006; NICE, 2009a; Koes et al., 2010). The following section presents and compares the existing treatment and management strategies for low back pain, with a consideration of clinical and cost-effectiveness evidence for these interventions.

5.4.3 Currently available interventions for managing non-specific low back pain

5.4.3.1 Information and education materials for self-care

Early management involves providing patients with education and information in relation to the nature, causes and impact of the condition to promote self-management strategies. The Back Book is one such patient educational material that was produced in the UK to provide patients with specific advice on how to cope with back pain in daily life (Burton et al., 1999). Such education resources as the Back Book are recommended because they provide complementary advice to clinicians and are relatively inexpensive.

There is evidence to suggest that the provision of educational materials alongside other interventions such as exercising improves overall patient beliefs and clinical outcomes (Burton et al., 1999; Little et al., 2001; Chou et al., 2007). Furthermore, evidence on the effectiveness of Back schools and other educational interventions shows that these are more effective than interventions such as spinal manipulation, and advice when provided in an occupational setting and alongside exercise programmes (Heymans et al., 2004; May, 2010). Others have, however, found that advice and educational programmes may influence aspects of behaviour but are not effective in improving low back pain function in patients with chronic low back pain (Roberts et al., 2002; Engers et al., 2008).

The advice to avoid bed rest and stay active has been associated with long-term improvement of outcomes, particularly among acute patients (van Tulder et al., 2006b). Advice to rest in bed compared to staying active has been found to have no effect on low back pain and in some instances has been associated with small harmful outcomes for acute low back pain patients. (Hagen et al., 2000; Dahm et al., 2010).

The economic evidence on education and advice interventions is largely inconclusive. Advice and educational material interventions have been reported as being more cost-effective when provided alongside cognitive behavioural therapy (Lamb et al., 2010b), in a combined multidisciplinary programme including acupuncture, exercises, dietary advice, and a back care book (Herman et al., 2008) than when used alone. However, there is some evidence indicating that advice to patients is a more cost-effective strategy than physiotherapy when provided in routine primary care practice (Rivero-Arias et al., 2006).

5.4.3.2 Pharmacological interventions

Pharmacological interventions are often considered as a means of providing acute and chronic low back pain patients with short-term pain relief alongside advice to stay active (NICE, 2009a; Koes et al., 2010). Such medications include paracetamol, non-steroidal anti-inflammatory drugs, antidepressants, muscle relaxants, and opioids. National guidelines from the UK (NICE, 2009a), the USA (Chou et al., 2007) and the Netherlands (van Tulder et al., 2006a) all recommend use of opioids, muscle-relaxants and anti-depressants in addition to commonly prescribed NSAIDs.

The evidence shows the effectiveness of NSAIDs and opioids when compared to paracetamol in providing short-term relief in patients with non-specific chronic low back pain, but are associated with more adverse side-effects in comparison to paracetamol (Roelofs et al., 2008). Moreover, there is no significant difference in pain relief between NSAIDs and paracetamol, although paracetamol tends to have fewer side effects (Roelofs et al., 2008). Similarly, muscle relaxants should be used with caution as they are associated with significant adverse effects yet are effective in the management of non-specific low back pain (van Tulder et al., 2003).

Economic evidence comparing the cost-effectiveness of pharmacological interventions in the management of low back pain remains scarce in the UK and internationally. However, a study in the US assessing the resource utilisation patterns of chronic low back pain patients found that opioids comprised a major cost component of the health plan for chronic low back pain patients (Vogt et al., 2005).

5.4.3.3 Interventions involving exercise programmes and advice

Various forms of exercise are often recommended alongside advice for patients to stay active particularly for people experiencing persistent non-specific low back pain (NICE, 2009a). The evidence shows programmes involving physical exercises alone are associated with improvements in pain and functionality outcomes (Hayden et al., 2005). Muscle strengthening and stretching exercises were in particular found to be the best type of exercise therapy. On the other hand, advice to stay active through exercises is associated with only small benefits in pain and functional improvement when applied to acute low back pain suffers (Dahm et al., 2010).

The economic evidence shows that a combination of GP prescription exercises and the Alexander technique sessions are a cost-effective strategy when compared to massage therapies (Hollinghurst et al., 2008). Furthermore, exercise programmes have been found to be cost-effective when provided in addition to current practice (U. K. Beam Trial, 2004), and cost-effective when provided in combination with education and cognitive therapy in chronic low back pain patients (Johnson et al., 2007).

5.4.3.4 Alternative manual and physiotherapy treatments

A wide range of physical treatments such as spinal manipulation or mobilisation, physiotherapy, acupuncture and massage have been recommended for treating patients with persistent low back pain (Koes et al., 2006). The evidence concerning the effectiveness of these interventions appears to be diverse. There is evidence showing massage therapy is more effective when compared with joint mobilisation, acupuncture, physical therapy and self-care education, particularly if massage treatments are combined with exercise and education strategies (Furlan et al., 2009). On the other hand, spinal manipulation and mobilisation treatments are associated with some short-term functional improvement in treatment of acute low back pain patients (BEAM, 2004) but not for chronic low back pain sufferers (Bronfort et al., 2004; Rubinstein et al., 2011).

Studies investigating the effectiveness of acupuncture have found insufficient evidence of the effectiveness of acupuncture interventions among acute non-specific low back pain patients, and some evidence of short-term pain relief among non-specific chronic low back pain sufferers (Furlan et al., 2009). However, acupuncture therapy showed some evidence of effectiveness in providing pain relief and improved outcome when provided alongside other treatment interventions in chronic but not in acute low back pain patients (Furlan et al., 2005).

There is economic evidence comparing strategies involving physiotherapy, pain management, spinal manipulation, acupuncture, exercise and behavioural counselling. The literature shows that physiotherapy treatment is cost-effective for treating low back pain when compared to pain management programmes (Whitehurst et al., 2007) but less cost-effective when compared to advice from a physiotherapist (Rivero-Arias et al., 2006). In treating patients with persistent chronic low back pain, traditional acupuncture care alongside conventional primary care interventions was found to be a cost-effective strategy in reducing long-term pain compared to usual care from a GP (Thomas et al., 2005).

Further work on spinal manipulation shows that treatment has been found to be cost-effective when provided alone or as an add on strategy to current practice within primary care (U. K. Beam Trial, 2004). In contrast, massage interventions have been found not be cost-effective when compared to advice , exercise or conventional care (Hollinghurst et al., 2008). However, a combination of massage, exercise and behavioural counselling improved the cost-effectiveness of the massage intervention after a one year period (Hollinghurst et al., 2008).

Treatment strategies in the management of persistent non-specific low back pain in the United Kingdom involving acupuncture have been found to be highly cost-effective from a societal perspective when compared to usual GP care (Ratcliffe et al., 2006) and, in Germany, relatively cost-effective when provided in addition to routine practice (Witt et al., 2006). In both studies, the additional costs associated with acupuncture were worth the benefits accrued for chronic non-specific low back pain patients.

5.4.3.5 Psychological treatments

Psychological treatment interventions involve a combination of cognitive behavioural interventions and self-management programmes (NICE, 2009a). Cognitive behavioural therapy in particular has often been used in treating patients with chronic non-specific low back pain (van Tulder and Waddell, 2005). Social and psychological factors have been found to increase functional disability in low back pain sufferers, hence the need for behavioural therapy and multidisciplinary interventions (Henschke et al., 2010). Recent clinical guidelines have therefore stressed the importance of providing psychological interventions in combination with other interventions (NICE, 2009a). Cognitive Behavioural Treatment (CBT) includes a combination of cognitive and behavioural techniques with the aim of enabling patients to better manage and respond to their pain and condition (Weiner and Nordin, 2010) . CBT has been found to be particularly effective for chronic low back pain (Koes et al., 2006).

In general, there appears to be minimal difference in functional status and behavioural outcomes between behavioural therapy interventions and commonly used therapies such as group exercises for pain in chronic low back pain patients, particularly in the long-term (Henschke et al., 2010). Similar evidence was reported in a clinical trial within a general practice setting offering CBT alongside education and active exercise interventions (Johnson et al., 2007). However, a large trial involving 56 general practices found both short and long-term benefits associated with CBT both in acute and chronic low back pain patients (Lamb et al., 2010a).

Combined treatments involving CBT aspects have been found to be superior to short-term pain relief interventions such as Back schools, advice and physiotherapy (Henschke et al., 2010). Moderate evidence in short-term and long-term effectiveness has been found when behavioural therapy interventions are combined with exercises and back schools (Ostelo et al., 2005).

In addition to the clinical evidence, short-term and long-term cost-effectiveness of CBT compared to active management has been shown among chronic low back pain patients (Lamb et al., 2010a).

Additional health economic evidence on psychological interventions found that a pain management programme incorporating cognitive-behavioural techniques was a highly cost-effective alternative to conventional outpatient physiotherapy or spinal manipulation intervention (Critchley et al., 2007). Additionally, cognitive behavioural interventions were found to be highly cost-effective when compared to current practice offering advice to low back pain patients in primary care (Lamb et al., 2010b). However, a physical therapy

(McKenzie approach) intervention was found to be more cost-effective than a brief physiotherapy pain management programme using cognitive behavioural principles (Manca et al., 2007). Thus, cognitive behavioural therapy appears to be a viable option for managing patients with chronic low back pain when compared to some of the common primary care interventions, but this might not be the most cost-effective option.

5.4.3.6 Summary

In summary, the literature shows that pain management programmes, cognitive behavioural approaches, exercise, spinal manipulation, acupuncture, and physiotherapy offer cost-effective options to health service providers particularly when combined together. The evidence on advice and educational interventions when provided alone is inconclusive and these appear to be best offered alongside other interventions. However, massage therapies appear not to be cost-effective, especially when provided as a separate intervention. More health economics evidence is needed on the cost-effectiveness of pharmacological interventions when provided alongside other acute and chronic pain management strategies.

5.4.4 Subgrouping of low back pain patients

Individuals developing low back pain have often been treated as a homogenous group without taking into consideration differences in underlying individual patient risk factors and characteristics (Wand and O'Connell, 2008). Clinical studies evaluating effectiveness of alternative treatments of low back pain patients often report small mean health benefits associated with non-specific low back pain interventions (BEAM, 2004; van Tulder et al., 2006b; Hay et al., 2008).

This observed, small, mean benefit could possibly be explained in light of recent advancements in the subgrouping of non-specific low back pain patients. It has been suggested that the assumptions of homogeneity in treating non-specific low back pain could potentially lead to the allocation of treatments that are not appropriate for patients diagnosed in this way (Boersma and Linton, 2005; Foster et al., 2008). Researchers have therefore developed ways to identify subgroups of patients who are likely to benefit from targeted treatment interventions as well as subgrouping tools as potential guides to decision making (Brennan et al., 2006; Hill et al., 2008; Kamper et al., 2010). These tools and measures have been implemented in clinical trials to assess the effectiveness of low back pain treatment interventions among patient subgroups.

For example, a recent randomised trial conducted within primary care general practices in the UK assessed the effectiveness of stratified care compared with current practice that does not involve stratification of low back patients (Hill et al., 2011). Using a validated subgrouping tool (Hill et al., 2008) – the STarT Back Screening Tool - patients in the trial were classified into one of three groups: 'low risk', 'medium risk' and 'high risk' based on a various factors associated with the development of back pain chronicity. The findings demonstrated that the stratified intervention approach which involved screening for risk factors and targeting treatment for specific subgroups of patients resulted in higher health gains and significant improvements in the primary outcome (disability) for patients with back pain when compared to the non-stratified approach. The study showed that without screening approaches to guide treatment allocations for back pain patients, patients in the 'medium risk' and 'high risk' groups could be denied access to the most appropriate treatments.

Similar research has been carried out to assess the implementation of a subgrouping and targeted treatment system within routine care practice producing similar findings (Foster et al., 2010).

In summary, it appears that the clinical assessment and management of non-specific low back pain should involve a stepped up care model approach in order to differentiate between simple low back pain patient conditions and potential serious back pain conditions. In this approach, management for low risk patients uses simple, low cost interventions such as advice and education and some medication and then moves on to more complex treatments such as exercises, massages, spinal manipulation, acupuncture, and cognitive behavioural therapy with increasing risk.

5.5 Assessing the economic burden of back pain

Low back pain is a common and costly condition according to the reported point prevalence rates ranging from 19% to 40% in the UK (Hillman et al., 1996; Macfarlane et al., 2012). Moreover, it has been estimated that 80% of working adults will experience back pain at some stage in their life (Palmer et al., 2000). The condition imposes a significant socioeconomic burden on the UK and internationally. Identifying the cost burden of low back pain through cost-of-illness studies provides important information for policy makers when planning (Dagenais et al., 2008).

Back pain has been linked with costs of over £1 billion in treating the different types of back pain each year within the NHS (NICE, 2009b). A significant proportion of the economic burden of low back pain is associated with costs to society through sickness absence from work as well as reduced productivity at work due to back pain illness. Work productivity loss due to back pain illness has been estimated at around 4.9 million working days per year (HSE, 2003/2004). Moreover, because of high prevalence rates, back pain affects many working age individuals hence contributing to economic loss in the society. The following section therefore provides an overview of cost of illness estimates for back pain in the UK and internationally in order to provide greater understanding of the economic burden of low back pain.

5.6 Cost of low back pain

5.6.1 Economic burden of low back pain internationally

Internationally, back pain is linked with significant costs wherever estimates of the condition have been quantified. A number of cost-of-illness studies reporting associated economic costs of low back pain in various countries have been identified internationally, with the results consistently showing relatively high indirect costs compared to direct costs (Dagenais et al., 2008; Maetzel and Li, 2002). There are variations in the studies, which makes it difficult to compare costs across multiple countries. Costing approaches from these countries tend to vary by perspective (whether healthcare or societal), costing method (top down approach or bottom up approach), indirect cost method (human capital or friction cost approach), sources of data and the time when the study was conducted. Table 5.1 provides a summary of cost-of-illness studies that have reported full societal costs including both direct and direct costs of low back pain from different countries. Costs were converted to UK pound sterling (£) using original exchange rates (http://www.hmrc.gov.uk/exrate/). The original price years of the estimates were recalculated to 2005 prices using CPI index data from the World Bank (http://data.worldbank.org/indicator/FP.CPI.TOTL?page=1).

Table 5:1 Selected cost-of-illness national total cost estimates for low back pain

Author/study year	Country	Price year	Total costs (In billions)*	Total Cost (2005 prices)**	% indirect costs
(Boonen et al., 2005)	Netherlands	2002	1.80	1.89	34%
(Jonsson and Husberg, 2000; Ekman et al., 2005; Hansson and Hansson, 2005)	Sweden	1994	2.13	2.39	87%
		2001	1.16	1.21	84%
		1995	2.78	3.06	93%
(Maniadakis and Gray, 2000)	United Kingdom	1998	6.7 to 12.3	7.35, 13.56	41% to 87%
(Shinohara et al., 1998)	Japan	1994	0.04	0.04	55%
(van Zundert and van Kleef, 2005)	Belgium	1999	0.78	0.88	84%
(Walker et al., 2003)	Australia	2001	3.29	3.66	89%
(Wieser et al., 2011)	Switzerland	2005	2.2 to 2.9	2.2, 2.9	46% to 61.4%

**Costs were expressed in 2005 price year costs using CPI data from the World Bank

* costs reported in original price years and expressed in UK pound sterling (£)

5.6.2 Costs of low back pain internationally

Cost-of-illness studies reporting both direct and indirect cost estimates of low back pain were identified from Australia, Belgium, Japan, Netherlands, Korea, Sweden, and Switzerland (Table 5.1). In the nine societal cost-of-illness studies identified, mean indirect costs accounted for 34% (Netherlands) to 93% (Sweden) of total disease burden costs, suggesting that indirect costs comprise the majority of total societal LBP costs. The largest societal costs of LBP were observed in the United Kingdom, with costs ranging from £7bn to £14bn based on estimates from the friction cost approach and the human capital method. This was followed by the study in Australia that found the societal costs of LBP to be £3.7bn in the 2005 estimated costs.

The corresponding estimates in Sweden ranged from £1.2bn to £3.1bn. Studies in Switzerland (Wieser et al., 2011), the Netherlands (Boonen et al., 2005), and Belgium (van Zundert and van Kleef, 2005) reported total societal costs ranging from £0.88bn to £2.9bn. The international comparison showed that the societal costs of LBP in Japan were smaller in comparison to costs reported from all other countries with reported costs of £0.04bn. The study by Wieser et al., (2011) estimated total costs in Switzerland of £2.9bn when using the human capital method and only £2.1bn using the friction cost approach, assuming a friction period of 22 weeks. This estimate was between 1.6% and 2.3% of the GDP. Similarly, costs from the Netherlands were calculated using the friction cost approach, assuming a friction period of three months.

5.6.3 Cost of back pain in the UK

Few attempts have been made to estimate the cost-of-illness of low back pain in the UK. These demonstrate clearly that low back pain poses a major economic burden to the UK, with total societal costs ranging from £1.2 to £12.3bn. For example, a prevalence based study report estimated the cost of back pain to the NHS at between £265 million and £382 million in 1992 (Klaber-Moffett et al., 1995). The corresponding costs incorporating productivity losses were estimated to range between £1.2bn and £1.74bn. The estimates reported in the study could however have been underestimated as the work absence days were based on a percentage of certified incapacity days (Klaber-Moffett et al., 1995).

Another UK study assessed the economic burden of LBP imposed on patients, the healthcare system and society as a whole in the UK using a combination of prevalence estimates, epidemiological, service utilisation and unit cost data from various sources (Maniadakis and Gray, 2000). This was a prevalence based study, estimating the annual cost of back pain in the UK in 1998. The findings indicated that the overall cost of back pain in 1998 was £12.3bn based on the human capital method (HCM) and £6.7bn using the friction cost approach (FCA). The total direct costs associated with care and treatment of back pain were £1.63bn, 35% of which comprised services provided by the private sector. Indirect cost of lost productivity due to incapacity to work was estimated at £3.44bn and £9.1bn using the FCA and the HCM respectively. The direct cost was estimated at £1.6 billion, while the cost of private care services was estimated at £565m. Estimates from this study were updated to 2005 costs in order to facilitate international comparisons (Table 5.1).

5.6.4 Comparison of low back pain with costs of other diseases

The important finding from the cost-of-illness studies in LBP is that indirect costs heavily outweigh the direct costs of healthcare. A comparison of LBP costs with other disorders in the UK shows similar findings. It is interesting to note that the back pain is more costly than other disorders such as depression, mental health, obesity and COPD. A lack of common methodology and cost components, however, makes comparisons across disease areas difficult. Nevertheless, the economic and social costs associated with mental health conditions in England have been estimated at around £77bn in 2003 (Sainsbury Centre for Mental Health, 2003). The corresponding amount in 2009 was estimated at £105bn with productivity costs comprising 29% of the total costs (Sainsbury Centre for Mental Health, 2010). In another study, the total costs of obesity to the UK were estimated to be over £2bn, with productivity costs accounting for 80% of the total costs (Vlad, 2003). Another investigation showed that chronic obstructive pulmonary disease places a significant burden on the healthcare system with total direct costs of £800 million and indirect costs of around £3bn (NICE, 2011). In the clinical area of depression, studies have reported total societal costs of £9bn, with productivity loss comprising over 90% of the total costs (Thomas and Morris, 2003) .

5.7 The study data sets

To address the question of how to incorporate productivity costs within economic evaluation studies, the empirical work of this thesis is based on studies designed to compare a stratified primary care management intervention with current practice in low back pain patients. In brief, this comprised two studies (Table 5.2). The first, investigated the effectiveness of a stratified management intervention compared with a non-stratified current practice in reducing long-term disability from low back pain in a randomised controlled trial (STarT Back; (Hill et al., 2011)). The second investigated the impact of introducing and managing the stratified management treatment intervention approach in primary care based on a prospective, population based, before and after study design in routine practice (IMPaCT Back study; (Foster et al., 2010)). Across both studies, patients consulting for low back pain in general practices were invited to attend an assessment clinic from which eligible patients were identified. Assessments were carried out by means of questionnaires. The STarT Back study was granted ethical approval by the North Staffordshire Local Research Ethics

Committee, while ethical approval for the IMPaCT Back study was obtained from the Cheshire Local NHS Research Ethics Committee.

Table 5:2 Summary of the data sets and study design

Study Design	Aim
Intervention study 1: STarT Back RCT (n=851)	To assess the effectiveness of a stratified primary care back pain management intervention with matched treatment intervention against a non-stratified current best practice.
Intervention study 2: IMPaCT Back Study (n = 927)	To implement a stratified management intervention care system for the evaluation and management of LBP patients in primary care and assess the impact on patient outcomes and practitioners.

5.7.1. Intervention Study 1: The STarT Back trial

5.7.1.1 Aims and objectives

This first intervention study which provided effectiveness data had the following aims and objectives. The main aim of the study was to assess the overall effectiveness of the stratified primary care back pain management intervention with targeted treatment options against non-stratified current best practice.

Secondary objectives of the study included:

1. To test whether, for the low-risk patients, minimum treatment provided non-inferior clinical outcomes to current best practice care
2. To test whether, for medium-risk patients, systematic referral to physiotherapy led to better clinical outcomes than did current best care
3. To test whether, for high-risk patients, systematic referral to psychologically augmented physiotherapy led to better clinical outcomes than did current best care.

Full details of the study methodology are reported in the study protocol and clinical study (Hay et al., 2008; Hill et al., 2011). An overview of the methods is provided in the next section.

5.7.1.2 Study population

The STarT Back study comprised low back pain consulters, aged 18 years and over, with a consultation of low back pain, identified from 10 general practices within the Keele General Practice Research Partnership, England (Hill et al., 2011). Eligibility was based on patients consulting for non-specific low back pain (i.e. excluding potentially serious pathology such as inflammatory arthritis, malignancy, and serious co-morbidity, pregnancy-related low back pain, personality disorders, individuals receiving care for a current back pain episode and those having had spinal surgery in the last 6 months (Hay et al., 2008).

Patients were recruited between the period of June 2007 and November 2008, with low back pain sufferers identified from searching GP patients' electronic data records. All patients identified were mailed a letter requesting them to make contact for an appointment at an assessment back pain clinic. Appointments were set up for patients who contacted the study team and details of the trial were explained during the appointment. Those consenting to participate in the study were taken through a process of completing written consent for participation in the study. 1573 individuals were identified as being eligible from the 10 general practices and, of these, 851 were included in the study. Patients were randomised to the interventions by a computer generated stratified block randomisation. Stratification was carried out based on the STarT Back screening tool risk subgroup (Hay et al., 2008).

5.7.1.3 Study design

The design of the study was a pragmatic randomised controlled trial over a period of 12 months; a total of 851 patients were randomised to either receive stratified primary care management (intervention, n = 568) or current best practice (control group, n = 283, in line with the standard physiotherapy usually provided). Consequently study participants were followed up at 3 months, 6 months and 12 months during the study period after the baseline assessment.

5.7.1.4 Study Interventions

The STarT Back trial interventions included an initial 30-minute physiotherapy evaluation with advice and education on physical activities and exercises for all patients. In the control group, the decision about additional treatment was determined by the physiotherapist, without having any knowledge of the patient group classifications. In the intervention group (stratified management), however, decisions about additional treatment were based on the STarT Back tool risk group classifications. High-risk patients were provided with additional psychologically inclined physiotherapy interventions aimed at addressing psychosocial factors affecting back pain recovery. Medium-risk patients were assigned to physiotherapy interventions aimed at addressing pain-related symptoms and physical functions; no further treatment was provided for the low-risk group (Hay et al., 2008; Hill et al., 2011).

5.7.1.5 Outcome measures

The main clinical outcome measure was the 24-item Roland Morris Disability Questionnaire (RMDQ; scores ranging from 0-24; high scores indicating severe disability) (Roland and Morris, 1983). Utility was assessed using QALYs derived from the EuroQol EQ-5D measure (Dolan, 1997). Work-related outcomes included days off work in the past 12 months, and reduced productivity at work in the past 30 days due to low back pain illness.

5.7.2. Intervention Study 2: The IMPaCT Back study

5.7.2.1 Aims and objectives

The second intervention study was the IMPaCT Back study (IMplementation study to improve Patient Care through Targeted treatment for Back pain). This was a quality improvement study that was designed to extend the stratified management approach within routine care in a primary care setting. The primary aim of the IMPaCT Back study was to assess whether the STarT Back model of stratified care for LBP could be implemented in routine primary care (Foster et al., 2010).

Secondary objectives of the study included:

1. Improve patients' clinical outcomes at 2 and 6 months follow-up;
2. Provide evidence for the sustainability of this care system.

5.7.2.2 Study population

The IMPaCT Back study recruited patients from 7 GP practices in the UK. The practices were selected with a balance of urban, semi-rural, rural and small, medium sized general practices.

Eligibility for the study included adult patients aged 18 years and over consulting for non-specific low back pain. Similar to the STarT Back Trial, exclusion of study patients was based on specific identifiable causes ('red flags') such as inflammatory arthritis, pregnancy, malignancy, recent surgery on back pain and patients already receiving physiotherapy treatment for the current back pain episode (Foster et al., 2010).

5.7.2.3 Study design and interventions.

A detailed account of the study interventions is reported in the study protocol (Foster et al., 2010). Briefly, the study was a prospective observational cohort study with a before and after study design involving three phases in routine primary care. A summary of the interventions from the study is provided below.

Phase 1 included provision of usual primary care for all patients, which included an option for referral to other healthcare services such as physiotherapy. Subgrouping for targeted treatment for low back pain patients was not included. In phase 1, clinical outcomes of patients in a baseline period of 6 months were obtained as well as attitudes and behaviours of general practitioners (GP's).

In the 2nd phase of the study, a diverse set of interventions aimed at improving the quality of the health system including educational courses, mentoring support, regular feedback sessions, and systems to support the subgrouping for targeted treatment in primary care was implemented. The multi-component quality improvement intervention included incorporating the subgrouping tool in physiotherapy and GP sessions within the primary care system in order to classify patients as being of low-risk, medium-risk or high-risk of poor outcome.

Phase 3 involved the same procedure as at baseline for patients and GPs recruited during a 12 month period following the roll-out of the new subgrouping and targeted care system. Data were collected to assess changes in patients' clinical outcomes and healthcare resource use, and healthcare professionals' back pain-related attitudes and behaviours after implementation of the stratified care and subgrouping intervention. In this phase, healthcare practitioners used the STarT Back tool as introduced in the phase 2 implementation of the quality improvement intervention.

5.8 Conclusion

This chapter has provided an overview of low back pain, presenting the nature, prevalence and management of the condition. The literature shows that non-specific low back pain is a common disorder, impacting patient morbidity and leading to significant productivity costs to individuals and society. This chapter has also reviewed economic evaluation studies of low back pain interventions and cost-of-illness studies in low back pain sufferers. The cost-of-illness literature demonstrates the magnitude of the economic burden of low back pain to society for various countries.

Policy makers worldwide have acknowledged the importance of low back pain by developing clinical guidelines (van Tulder et al., 2006a; Chou et al., 2007; Airaksinen et al., 2006; NICE, 2009a). The novel approaches to subgrouping and stratified management care discussed in this chapter have been found to improve the clinical management of low back pain, and they could potentially help in reducing the economic burden. The low back pain clinical area and study datasets discussed in this chapter will be used as case studies for the empirical work conducted later in this thesis. Before that, however, the next chapter focuses on generating empirical estimates for friction period estimates specific to the UK context.

CHAPTER SIX ESTIMATING A FRICTION PERIOD FOR THE UK

6.1 Introduction

The overall aim of this chapter is to further develop the friction cost approach for use in the UK. Two areas are covered. The first focuses on generating specific data on vacancy durations, hence friction periods, that can be used in practical applications of the friction cost approach in valuing productivity costs within economic evaluations in a UK context. The second is to compare methods for capturing time off work. The next section of this chapter covers the rationale for the work. This is followed by information about the data and background. The methods and results of (i) investigations into estimating friction periods for the UK and (ii) comparability between sickness certification records and self-report data in estimating time off work are then explored. Finally, the discussion focuses on the implications of these findings and their application for work in the forthcoming chapters.

6.2 Rationale

The literature reviews presented in Chapters Three and Four show that few studies have applied the friction cost approach within economic evaluations in the UK. This is mainly attributed to the absence of reliable relevant data necessary for its application, with particular reference to a lack of information about the friction period, for which there are no reliable estimates within the UK (Koopmanschap et al., 1995; Pritchard and Sculpher, 2000). The duration of a vacancy has been suggested as a proxy for estimating a friction period (Brouwer and Koopmanschap, 2005). This is the period employers take to successfully fill an employment vacancy (van Ours and Ridder, 1991). However, few studies have empirically investigated vacancy durations, particularly within the UK. To date, only four studies in the

UK, (Beaumont, 1978; Roper, 1988; Adams et al., 2002; Andrews et al., 2008) have explored different aspects of vacancy data, and these do not provide the sort of information required to estimate the friction period for use in productivity cost valuations. Of these studies, Andrews et al., (2008) analysed computerised career service data using econometric modelling (Andrews et al., 2008). Their findings showed an average vacancy duration of four weeks; however this was limited to a youth population. In a duration of four weeks, a model by Adams et al., (2002), that considered employees mostly from the financial and services industries, found that 63 per cent of vacancies had been filled, and only 8 per cent of the vacancies were unfilled within a period of 12 weeks. Elsewhere, using vacancy data from UK survey of employers, Roper (1988) assessed the effect of recruitment methods on vacancy durations. Their results showed that the method of recruitment affects the length of vacancy durations. Neither of these studies however provides the detailed specific vacancy duration information required to estimate a mean friction period necessary for valuing lost productivity when using the friction cost approach. Moreover the data used in these studies are relatively old (1985-1992 for Adams et al., (2002), 1996-1997 for Andrews et al., (2008) and data from the 1970's for both Beaumont (1978) and Roper (1988)) and do not accurately illustrate the prevailing labour market conditions. Clearly this is an under-researched area within the UK.

Economic evaluations often use a single average friction period when valuing productivity costs. However, the use of an average friction period ignores differences between types of employees in the labour market, which could potentially lead to inaccurate estimation of productivity costs (Koopmanschap and Rutten, 1996; Koopmanschap et al., 1995). The impact of using more detailed friction periods therefore remains uncertain. A stratification of friction periods by individual subgroups based on education status has been conducted (Koopmanschap et al., 1995), but as yet, there has not been a stratification based on occupational status. However, occupational status is often collected in economic evaluation studies and is clearly more directly relevant to vacancy durations and hence friction periods (Andrews et al., 2008). Estimation of friction periods by occupational status would clearly provide an advance on the information currently available.

6.3 Data and Background

To accurately estimate more detailed friction periods for the UK, data on the average length of time to fill a vacancy were surveyed from four sources for the period 2007-2011. The initial data search involved surveying the office of National Statistics (ONS) using a questionnaire, as has been done in previous countries, for example the Netherlands when obtaining vacancy duration statistics. This was followed by a survey from the Chartered Institute of Professional Development (CIPD), the largest human resource professional body in the UK. From the CIPD library and information services department, a third data source, i.e. DLA Piper LLP that collects relevant data on vacancy duration data was also identified. Finally, vacancy duration data were purposively collected from the University of Birmingham as part of a primary data collection exploration process. In each case, the data collected comprised the time taken to fill a vacancy, subgrouped by the relevant occupational categories.

These data sources are briefly described below:

1) DLA Piper LLP (DLA)

DLA Piper LLP is a global services legal organisation. It works with various local, national and global organisations and also operates throughout the public sector, working with central governments, local governments, universities and housing associations. The organisation has offices throughout Asia, Europe, Australia, the Middle East and the United States. It offers a human resource (HR) measurement and benchmarking service that collects performance data on a range of indicators from over 400 contributing organisations.

2) Chartered Institute of Personnel and Development (CIPD)

The Chartered Institute of Personnel and Development is the world's largest chartered HR and professional body. It is an internationally recognised organisation that supports and develops individuals responsible for managing and developing people within human resources. The organisation collects HR vacancy statistics data through annual surveys.

3) Office of National Statistics (ONS)

The Office of National Statistics is a key source of national statistics in the UK. It is the UK's largest independent producer of official statistics, and is also the recognised national statistical institute of the UK. ONS works with various organisations in the UK and provides a leading role in the development of national and international good practice in the production of official statistics. The organisation collects a range of work performance indicators over a diverse range of periods.

4) University of Birmingham (UoB)

The University of Birmingham is one of UK's leading universities founded in 1900 and associated with excellence in research and teaching. The organisation is a very large employer with over 5000 staff from a broad range of occupations. The university maintains a database with HR vacancy statistics through the Human Resources statistics department.

Data from these four sources were used to estimate detailed friction periods stratified according to different occupational level categories. Each of the data sources included, however, slightly different questions, categorised occupational levels differently and provided data from different time points which necessitated adjustments to make the data more comparable. A detailed description of each data source and the approaches for dealing with differences in the dataset used are described in the methodology.

6.4 Methodology

The methodology is organised as follows: section 6.4.1 defines the meaning of vacancy duration for the purpose of the research, and explains how adjustments were made to ensure the same definition was used across all data sources. Section 6.4.2 then describes the relationship between the occupational titles of participants in the study datasets and an established international standard occupational classification. In section 6.4.3, a detailed description of the data sources and the underlying matching processes with the standard occupational classification is described. A description of the necessary adjustments carried out for each source to make it generalisable at firm and national level is then provided in section 6.4.4. Finally, sections 6.4.5 and 6.4.6 provide an overview of the work-related questionnaire used in the study and the process of identifying a friction period occurrence in the datasets.

6.4.1 Standardising vacancy duration definition

Data from the different sources were collected at firm-level based on slightly different vacancy duration definitions and occupational categorisations. The standard period considered for the survey was the average period taken to fill a vacancy, from when a vacancy is raised to when an individual starts working. To ensure comparability, an additional period, in line with previous literature, was added to the values from those data sources that did not cover the survey duration from when a vacancy is raised to when an individual starts work. Since there is no known estimate for the time lags involved before a vacancy is raised as well as the extra time lags for a successful replacement employee to start work in the UK, an average period of 4 weeks was adopted from a study in the Netherlands that used this period (Koopmanschap et al., 1995).

6.4.2 Occupational categories

In the STarT Back trial (SBT) and IMPaCT Back study (IBS), respondents were asked to report whether they were employed at baseline and follow-up as well as to record their current job title. The job title captured for each individual was used to assign responders to specific categories according to shared characteristics based on the Standard Occupational Classification 2000 (SOC, 2000). As the thesis was focused on valuing productivity costs of individuals in paid employment, final classifications used in categorising the friction period were based on the SOC 2000 which is limited to employed individuals. The SOC 2000 is a classification that can be applied to all paid jobs performed by economically active individuals in the UK (SOC, 2000). Other standard classifications widely used in the United Kingdom include the Standard Industrial Classification (SIC) and National Statistics Socio-economic Classification (NS-SEC). The SOC 2000, which focuses on employed individuals, was used here to link categories of employed individuals with the Annual Survey of Hours and Earnings (ASHE) carried out by the Office of National Statistics (ONS). The ASHE is a labour market and earnings survey based on a sample of UK employees that provides information about the earnings and hours paid for employees within different occupations in industries by age groups and regions (ASHE, 2008). The ASHE is based on the SOC classification system and so was a particularly useful classification to use.

The following major groups of the SOC 2000 were considered:

a) Managers and senior officials

b) Professional occupations

c) Associate professional and technical occupations

d) Administrative and secretarial occupations

e) Skilled trade occupations

f) Personal service occupations

g) Sales and customer service occupations

h) Process, plant and machine operatives

i) Elementary occupations.

As the different data sources asked slightly different questions, and were categorised differently, the following final occupational category levels were identified and summarised for this thesis following adjustments as described in the descriptions of the data sources.

i) “Managerial and professional” or “operations and support” for DLA Piper LLP

ii) “Managers and senior managers” or “professional occupations”, “Administrative, secretarial and technical professions” or “services” (including sales, customer, and personal staff) or “elementary occupations” for the three remaining data sources including the ONS, CIPD and UoB.

6.4.3 Measuring the average duration to fill a vacancy – Data sources

To estimate friction periods for the UK, vacancy duration statistics were obtained from a number of sources described in the next section. The following data sources were identified and are described in more detail including the questions asked, the duration covered, the nature of data collected and how it was categorised.

6.4.3.1 DLA PIPER’S HR Bench marker Statistics (2009-2011)

The DLA Piper UK LLP data analysed in this study were collected by annual vacancy surveys as part of the organisation’s HR bench marker research (DLA PIPER, 2012). The HR bench marker is a human resource service involving over 400 participating organisations, which enables organisations to compare their performance on indicators such as vacancy duration and absence rates. For this analysis, recruitment data obtained were used to generate trends data for the period 2009 to 2011. Data from the survey comprised both public and private sector organisations with the latter comprising slightly over 40%. The data were collected based on two categories of employees: i) managerial/professional ii) operational/support. Measures of average, median, upper and lower quartiles were then estimated from the reporting organisations for each category. Respondents in the survey were asked to estimate the number of working days taken to fill a vacancy (from vacancy notification to job offer)(DLA PIPER, 2012).

The SOC 2000 classification was collapsed into the two levels for which vacancy duration data were available from DLA piper according to Figure 6.1. The 9 classes represent the original SOC 2000 classification. To standardise the duration from when a vacancy is raised to when an individual reports to their work place, the data were adjusted by adding in a period of one month (representing the period from the offer date to the start date).

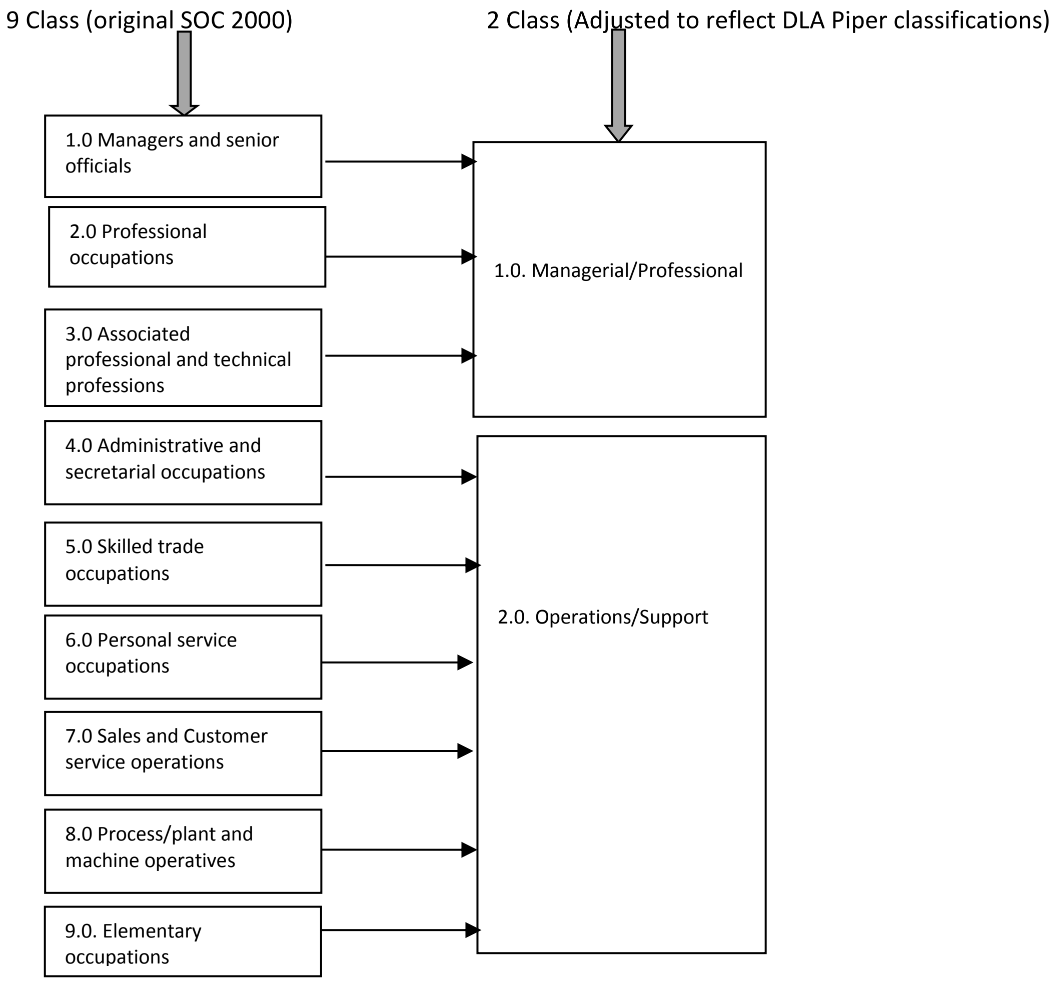


Figure 6:1 Collapsing the Standard Occupational Classification 2000 (SOC 2000) into the DLA Piper Dataset friction period

6.4.3.2 Chartered Institute of Personnel and Development (CIPD) - Annual surveys (2007 - 2009)

The CIPD data analysed in this research were collected by annual surveys assessing the average time and costs required to fill a vacancy in the UK. Specifically, the data were obtained from the Recruitment, Retention, and Turnover Survey Reports that show trends of how organisations recruit and maintain personnel within the UK (CIPD, 2009; CIPD, 2008; CIPD, 2007). In the surveys, questionnaires were sent out to HR professionals in private, public and voluntary sectors.

The specific wording used in the CIPD surveys is presented below:

“What is the average **number** of weeks to fill a vacancy?”

Data were available only for the period 2007-2009 and subsequent surveys from 2010 no longer collect statistics on filling job vacancies. The SOC 2000 classification system was categorised into five levels based on the data obtained from CIPD according to Figure 6.2.

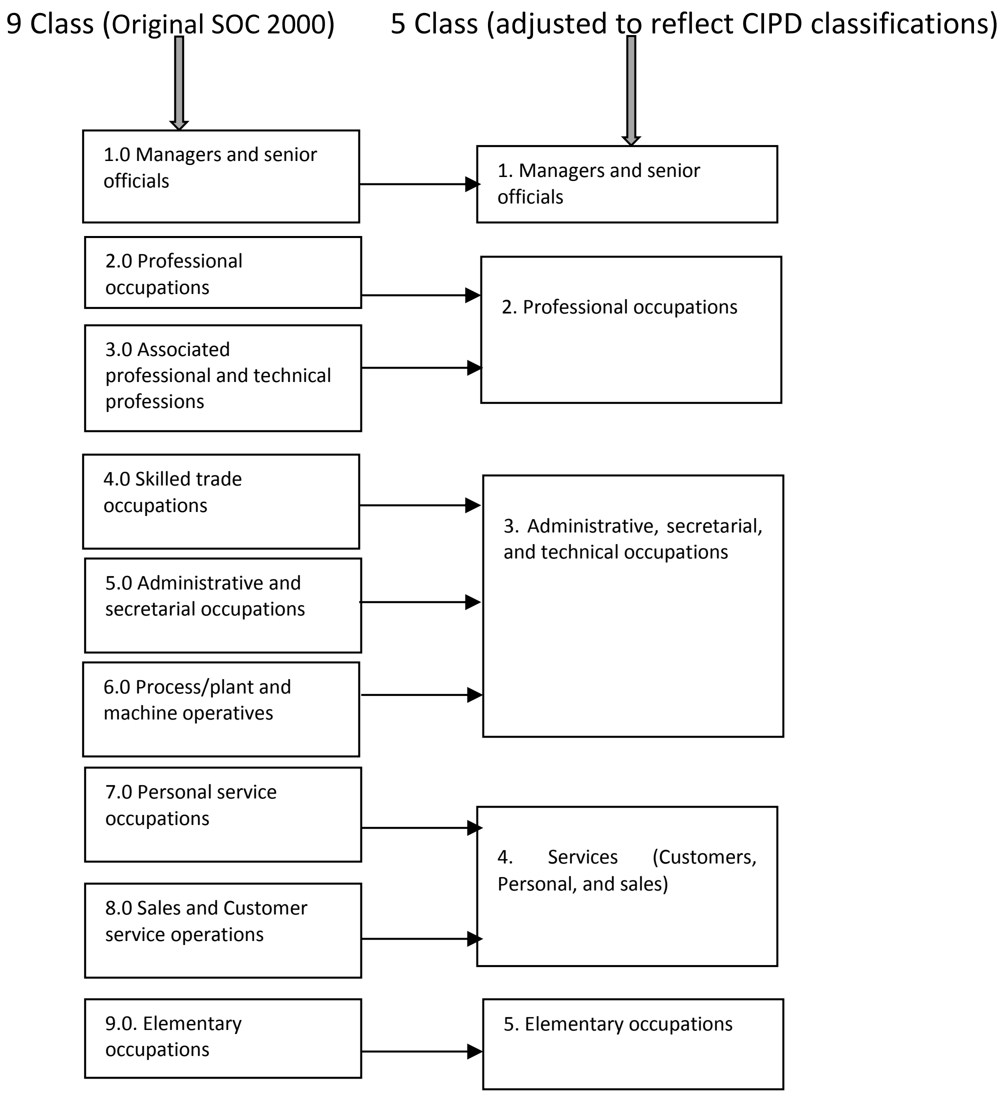


Figure 6:2 Collapsing the Standard Occupational Classification 2000 (SOC 2000) into the CIPD Dataset

6.4.3.3 Office of National Statistics (ONS)

Vacancy duration data were also obtained from the ONS, which generates independent information on the UK labour economy. These data mainly comprised vacancy outflows by duration and occupation, based on the Jobcentre Plus vacancy statistics. Jobcentre Plus, a public employment service attached to the Department of Work and Pensions (DWP), provides services that link people to work and helps employers to fill vacancies (Bukowski et al., 2010). The data were obtained through the ONS official labour statistics database known as Nomis (http://www.nomisweb.co.uk/default.asp) and summarised.

Nomis is a service provided by the Office for National Statistics (ONS), which provides free access to comprehensive and updated UK labour market statistics. The database provides various categories of data including the measure on vacancy outflows used here. Vacancy outflows are the number of jobs that have been filled or withdrawn from the market by the Jobcentre Plus during a specified period. A vacancy is considered as filled if a client is awarded a job as a result of the Jobcentre Plus recommending that client to the vacancy.

The sample used in this analysis covers the period January 2007 to October 2011 with the exception of data from September 2010 that were withdrawn due to inconsistencies in the data sources. Similarly to the DLA Piper data, these data were standardised to the duration from when a vacancy is raised to when an individual reports to their work place by adding in a one month duration, which covers the period from offer date to a starting date (Koopmanschap et al., 1995). Similarly, the SOC 2000 classification system was collapsed into five category levels by mapping their work characteristics onto the data obtained from ONS as in Figure 6.2.

6.4.3.4 University of Birmingham

As a final data source, which can be viewed as an exploratory exercise into primary data collection of vacancy duration statistics, data were also collected from an organisation within a research and academic context. This enabled the estimation of friction periods for academic and research occupations that are potentially not available through the other data sources used. These estimates were surveyed and obtained from the University of Birmingham, Human Resources (HR) statistics department. The University of Birmingham employs over 5000 staff. The vacancy duration statistics generated included the duration from when a vacancy is advertised to the actual start date. The SOC 2000 classification system was categorised into five levels based on the data obtained from the University as in Figure 6.2.

6.4.4 Weighting friction period estimates for national level data

The friction period values obtained from the sample data from the four data sources were weighted to adjust for the proportion of jobs in each occupational class at the population level, to make the overall friction period estimate representative of the UK population. Sample weights comprising the total number of jobs in each occupation class for each year, from the ONS covering the period 2007-2011, were used to adjust the friction period values from the sample data to values of the population. The ONS labour force survey (http://www.ons.gov.uk/ons/search/index.html?newquery=all+in+employment) data were considered appropriate for this adjustment as they provide the largest annual comprehensive national estimates of jobs across the entire UK economy. Results are presented for the un-adjusted and adjusted friction periods.

Weighting of the sample datasets involved a weighted average statistical analysis of the occupational categories data. Consider e different occupational categories, and let f_1, f_2…, f_e represent the friction period. The weighted average was based on equation 6.1.

Equation 6.1

$$\boldsymbol{Xw} = \frac{{}_{i=1}^{e} Wi\, fi}{{}_{i=1}^{e} Wi}$$

Here,

- w_i represents the weighting factor based on the total UK national vacancies by occupation
- f_i represents the estimated values of friction periods in the different occupational categories
- Xw corresponds to the weighted mean duration of a friction period for the UK

6.4.5 Work-related questionnaire used in the studies

In both the SBT and IBS, information on absenteeism and presenteeism was collected through a postal questionnaire. In addition to back pain and quality of life related questions, respondents were also asked a number of employment and demographic questions including gender, age, job title, and current job status. Specifically, respondents were also asked to report on whether they had experienced absenteeism from work during the study reporting period. If so, they were asked to record the duration of the absenteeism. Then, they were asked to report the extent to which back pain affected their performance at work (presenteeism).Information on the respondents’ current working situation was also obtained.

6.4.6 Identifying a friction period in the data sets

To determine the frequency of friction periods, data on work absence and disability were sought from the two study datasets. In accordance with previous research (Koopmanschap et al., 1995; Koopmanschap and Rutten, 1996), a friction period is considered in a situation where there is:

i) Immediate death

ii) Absence from work that resulted in an individual having to stop work due to a work-related disability and hence the need for them to be replaced at their workplace

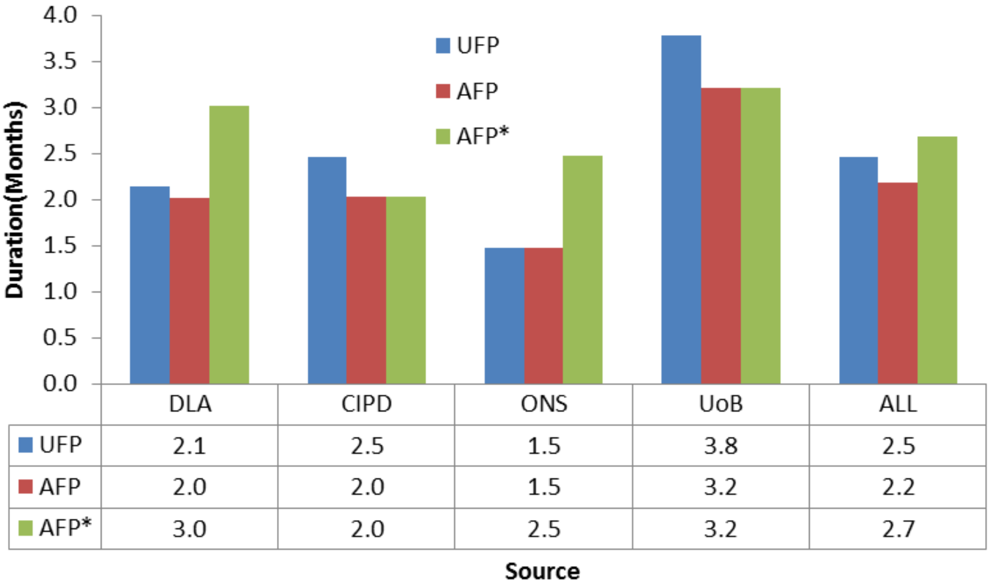
iii) Absence from work longer than the friction period.

However, because of the structuring of the work questions in the study datasets, the duration of specific individual episodes of work absence could not be obtained. A more accurate extraction of this information could have been achieved had the retrospective work questions been structured in a more specific way to capture the frequency of absence episodes and the duration of each episode or through the use of detailed patient diaries documenting absence days (Stewart et al., 2000). This informed part of the methodological work undertaken and is presented in the following subsection as part of this research – assessing the feasibility of using sickness certification records to identify length of sickness absence episodes using medical certification data in order to apply the friction cost approach.

6.5 Results: Estimating a friction period for the UK

6.5.1 Vacancy duration

The average vacancy duration for the period 2007-2011 ranged from 2 months to 3.2 months after adjusting the period to national level values (See Figure 6.3). The overall average adjusted friction period was 2.7 months.



	DLA	CIPD	ONS	UoB	ALL
UFP	2.1	2.5	1.5	3.8	2.5
AFP	2.0	2.0	1.5	3.2	2.2
AFP*	3.0	2.0	2.5	3.2	2.7

Figure 6:3 All organisations – Summary of friction period estimates for UK in months (number), by occupational level

DLA - DLA Piper LLP consultancy firm

CIPD Chartered institute for Professional Development

ONS Office of National Statistics

UoB - University of Birmingham

ALL - Average of all organisations

UFP – Unadjusted friction period, AFP Friction period adjusted for national vacancy stocks by occupation, AFP* Friction period adjusted for national vacancy stocks by occupation and duration to take up job vacancy for DLA and ONS.

Details of vacancy duration estimates for the individual data sources are shown in Tables 6.1 to 6.4 and Figures 6.3 to 6.7. More specifically, the results showed higher maximum vacancy durations in the UoB (range: 2.5 to 5.6) and CIPD (range: 1.5 to 4.2) individual data sets; relatively lower vacancy duration ranges were observed in the DLA piper LLP dataset (range: 3.0 to 3.6 – including additional correction for time lags in starting a new job) and the ONS dataset (range: 2.3 to 2.7 – including additional correction for time lags in starting a new job). The additional adjustment for these data sets was to standardise the data collection period to include the period when an individual accepts a job up to the point of actually starting work.

As Tables 6.1 to 6.4 show, there were differences between the friction period range values from the statistics generated from the different data sources.

Table 6:1 Friction period estimates for UK in months (number), by occupational level, for 2009-2011 – DLA Piper dataset

Year	**Occupational Classification Average (median; LQ - UQ)**		
	ALL	1	2
2009	2.3 (2.0;1.5-1.8)	2.6 (2.3;1.5-3.1)	1.9 (1.5;1.2-1.8)
2010	2.2 (2.0;1.4-2.7)	2.6 (2.4;1.6-3.1)	2.0 (1.7;1.3-2.3)
2011	2.3 (2.0;1.5-2.7)	2.5 (2.3;1.5-3.1)	2.0 (1.7;1.3-2.4)
Unadjusted values	2.2 (2.0;1.5-2.4)	2.5 (2.3;1.5-3.1)	1.9 (1.6;1.2-2.1)
Adjusted for national UK vacancy values	2.0	2.6	2.0
Overall adjusted*	3.0	3.6	3.0

ALL. All employees,

1. Managerial/Professional,

2. Operational/Support

* Includes average duration of one month for time taken for an individual to start work

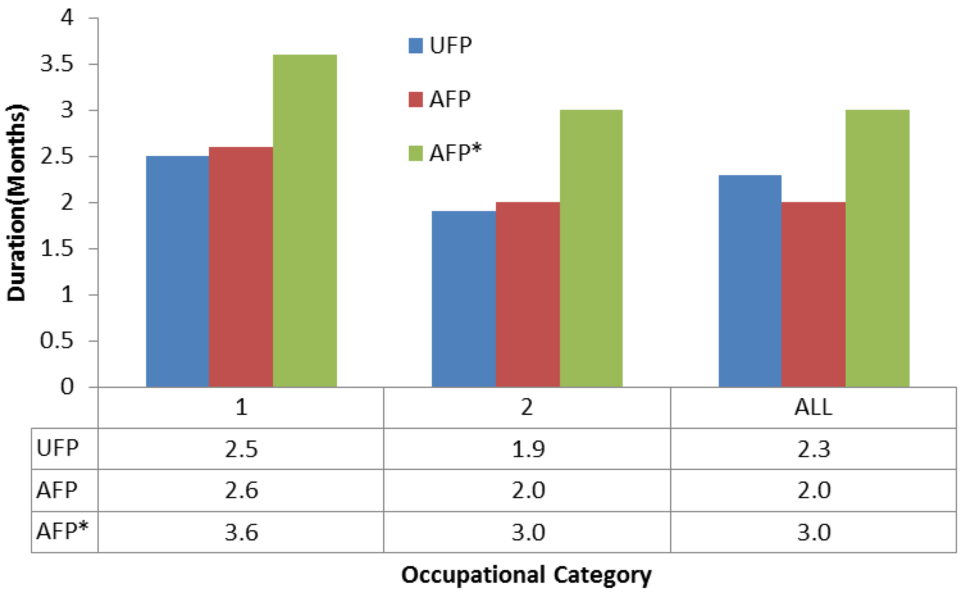


Figure 6:4 DLA Piper – Summary of friction period estimates for UK in months (number), by two occupational level categories (2009-2011)

1 Managerial/Professional

2 Operational/Support

ALL Average of all organisations

UFP – Unadjusted friction period

AFP Friction period adjusted for national vacancy stocks by occupation

AFP* Friction period adjusted for national vacancy stocks by occupation and average duration to take up job vacancy

Table 6:2 Friction period estimates for UK in months (number), by occupational level, for 2007-2009 CIPD

Year	Occupational classification Average(n)				
	1	2	3	4	5
2007	4.1 (761)	3.2 (814)	1.7 (797)	1.9 (678)	1.5 (611)
2008	4.1 (723)	3.0 (733)	1.7 (737)	1.9 (670)	1.5 (642)
2009	4.3 (703)	3.1 (720)	1.6 (717)	1.8 (657)	1.5 (633)
ALL	4.2	3.1	1.7	1.9	1.5
Adjusted for national UK vacancy values	4.2	3.1	1.7	1.9	1.5

1. Senior managers/directors
2. Managers and professionals
3. Administrative, secretarial and technical
4. Services (customer, personal, protective and sales)
5. Manual/craft workers

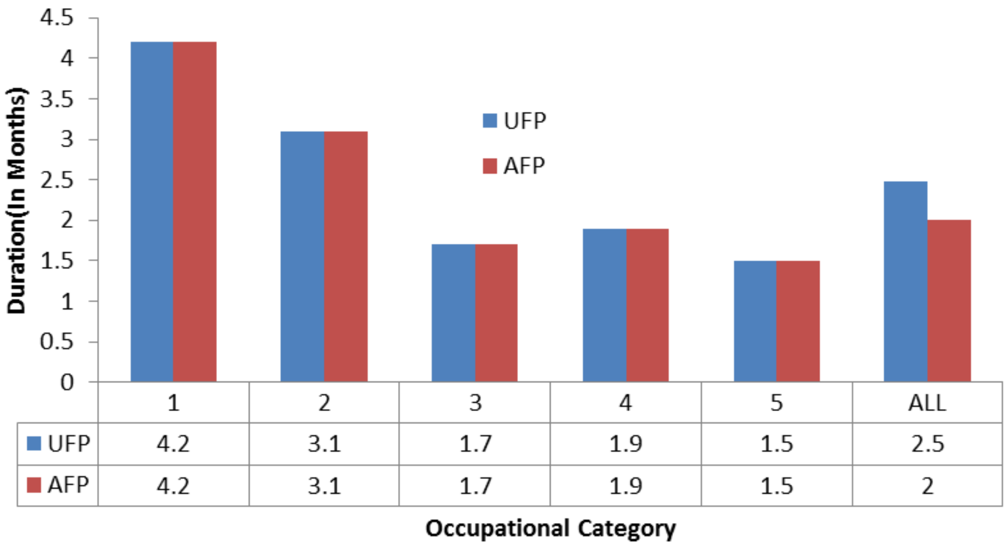


Figure 6:5 Summary of friction period estimates for UK in months (number), by two occupational level categories (2007-2009)

ALL Average of all organisations. UFP – Unadjusted friction period, AFP Friction period adjusted for national vacancy stocks

1. Senior managers/directors
2. Managers and professionals
3. Administrative, secretarial and technical
4. Services (customer, personal, protective and sales)
5. Manual/craft workers

Table 6:3 Friction period estimates for UK in months, by occupational level, for 2007-2011 (Based on vacancies filled by job centre plus and alternative recruitment channels) ONS

Year	Occupational level (Average)				
	1	2	3	4	5
2007	1.6	1.7	1.4	1.9	1.4
2008	1.6	1.9	1.4	1.8	1.4
2009	2.0	1.8	1.2	1.7	1.2
2010	1.4	1.5	1.2	1.6	1.3
2011	1.3	1.4	1.2	1.5	1.2
ALL	1.6	1.6	1.3	1.7	1.3
Adjusted for national UK vacancy values	1.6	1.7	1.3	1.7	1.3
Adjusted Overall*	2.6	2.7	2.3	2.7	2.3

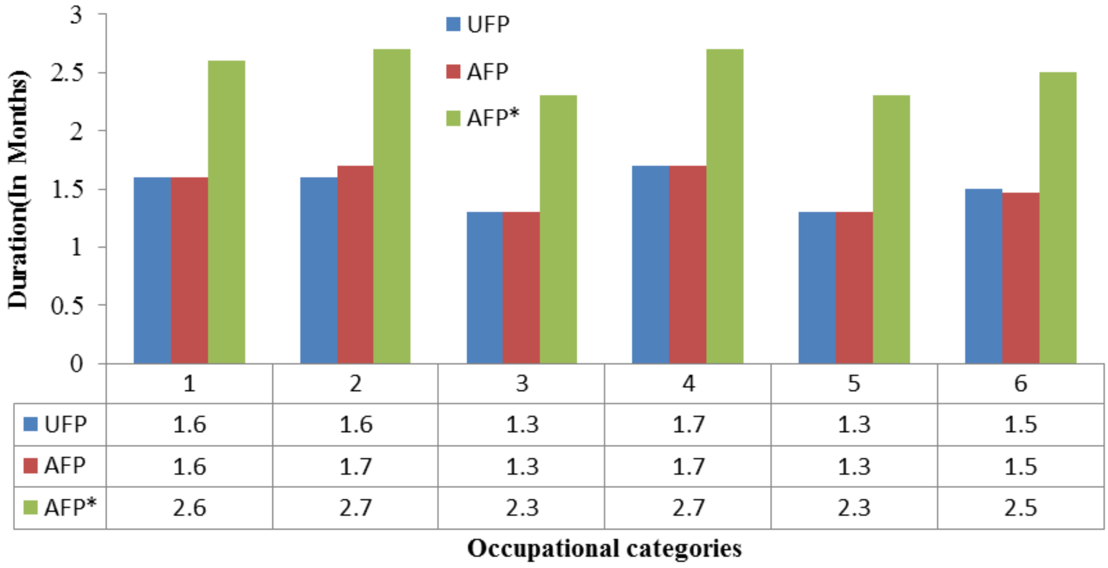
1: Managers and Senior Officials

2: Professional Occupations/Associate Professional and Technical Occupations

3: Administrative and Secretarial Occupations/Skilled Trades Occupations/ Process, Plant and Machine Operatives

4: Personal Service Occupations/Sales and Customer Service occupations

5: Elementary Occupations



	1	2	3	4	5	6
UFP	1.6	1.6	1.3	1.7	1.3	1.5
AFP	1.6	1.7	1.3	1.7	1.3	1.5
AFP*	2.6	2.7	2.3	2.7	2.3	2.5

Figure 6:6 Summary of friction period estimates for UK in months (number), by two occupational level categories (2007-2011)

ALL- Average of all organisations. UFP – Unadjusted friction period, AFP Friction period adjusted for national vacancy stocks.

Managers and Senior Officials

2: Professional Occupations/Associate Professional and Technical Occupations

3: Administrative and Secretarial Occupations/Skilled Trades Occupations/ Process, Plant and Machine Operatives

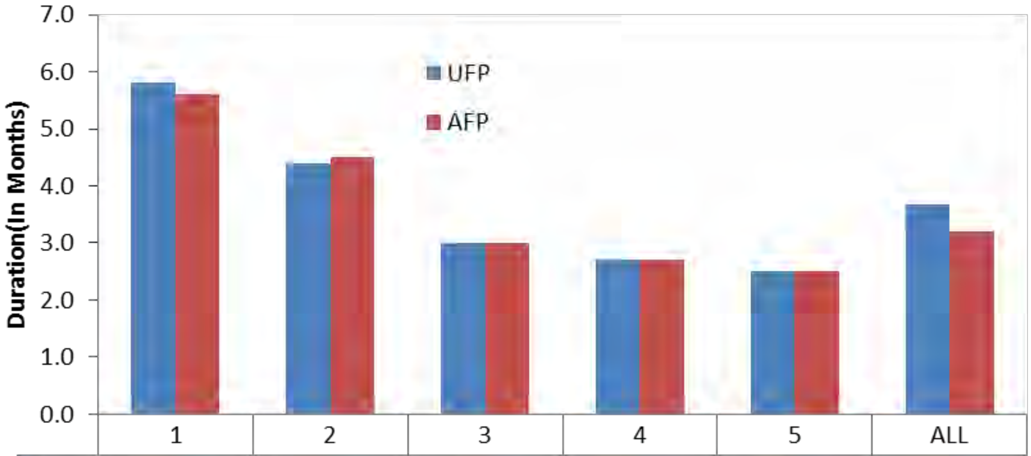
4: Personal Service Occupations/Sales and Customer Service occupations

5: Elementary Occupations. AFP* Friction period adjusted for national vacancy stocks by occupation and average duration to take up job vacancy for DLA.

Table 6:4 Estimated friction period in months, by occupational level, for 2009-2011 (Based on vacancies filled by the University of Birmingham) UoB

Year	Occupational Classification (average)				
	1	2	3	4	5
2009	5.0	3.9	2.9	2.9	2.1
2010	6.6	5.2	3.2	2.9	2.3
2011	6.0	4.3	3.0	2.4	3.0
ALL	5.8	4.4	3.0	2.7	2.5
Adjusted for national UK vacancy values	5.6	4.5	3.0	2.7	2.5

1. Managers, Senior officials, Directors
2. Professional and technical occupations
3. Administration, Skilled trade
4. Services
5. Elementary/Routine



	1	2	3	4	5	ALL
UFP	5.8	4.4	3.0	2.7	2.5	3.7
AFP	5.6	4.5	3.0	2.7	2.5	3.2

Occupational Categories

Figure 6:7 UoB – Summary of friction period estimates for UK in months, by five level occupational categories (2007-2009)

ALL - Average of all organisations. UFP - Unadjusted friction period, AFP - Friction period adjusted for national vacancy stocks.

1: Managers and Senior Officials,

2: Professional Occupations/Associate Professional and Technical Occupations,

3: Administrative and Secretarial Occupations/Skilled Trades Occupations/ Process, Plant and Machine Operatives,

4: Personal Service Occupations/Sales and Customer Service occupations,

5: Elementary Occupations.

AFP* Friction period adjusted for national vacancy stocks by occupation and average duration to take up job vacancy for DLA. UoB University of Birmingham

6.6 Feasibility of using sickness certification records to identify work absence in economic evaluation

In the UK, general practitioners (GPs) play a key role in the sickness certification process as part of their contractual role (Hussey et al., 2004). GPs will normally issue a fit note, formerly a doctor's sick note (sickness certificate) for an illness lasting more than one week based on their expert judgement of a patient's capacity to work and the probable duration of sick leave required (Morrison, 2011). However, patients could potentially return to work at an earlier date than agreed with the GP. The 'fit note' replaced the doctor's sick note on 6 April 2010 (DWP, 2013). Additionally patients can self-certify sickness for an absence period of 7 days or less as GPs are not required to issue a fit note before seven days of absence have elapsed (DWP, 2013). However, self-certified absence is not documented in GP certification databases. Prospective investigations exploring how sickness absence due to illness using GP issued sickness certificates (fit note) could provide additional data to estimate episodes of work absence due to sickness. Nevertheless, knowledge about the comparability of GP sickness certification records and self-reported absence methods remains very limited (Wynne-Jones et al., 2008).

A secondary analysis was carried out for a sample of respondents in the SBT and IBS datasets who consented to having data from their sickness absence records collected from the reviewed general practitioner database. This analysis was to assess the feasibility of using sickness certification records to estimate time off work due to sickness in economic evaluations. A secondary purpose was to explore the use of sickness certification records to identify the frequency and duration of work absence episodes when applying the friction cost approach. It was hypothesized that such data could be used to assess incidences and durations of back pain sickness episodes, in an economic evaluation setting, in the absence of detailed patient self-reported data that clarified the duration of individual detailed episodes of time off work.

This exploration involved comparing: i) incidence of sickness absence based on a self-report questionnaire data and GP certification records; ii) sickness absence duration data by combining the total duration of each episode listed in the medical certificates and self-reported absence days.

6.6.1 Population of the sample

Study participants (SBT n=295; IBS n=256) who gave their consent for review of their sickness certification records were included in the study. In addition to the self-report postal questionnaire including work absence completed by participants in the two datasets, back pain related sickness certificates were extracted from the primary care registry and consultations databases in the Keele GP research partnership. Data on certification information showing patients who received sickness certification were linked to time off work data from the SBT and IBS data sets using a unique individual identification number and Read codes. Read codes comprising levels of process, symptom and morbidity codes were used to identify the type of sickness certificate, alongside the corresponding consultation data (Hiscock and Ritchie, 2001; Wynne-Jones et al., 2009).

The certification database included information about patients who had sick leave spells of 7 days or more. Sickness certificates issued during these periods were matched with self-reported work absences.

6.6.2 Measurement and analysis

The main measures in this investigation were the occurrence of an absence episode and the associated length of absence in the self-report and sickness certification data sources. For each of these data sources, the reported time off and their associated lengths were reviewed. Agreement of sickness certificates with absence records was based on a set of criteria in Figure 6.8 below adopted from Wynne-Jones et al., (2008), pg 116.

Direct match
No reported work absence and no sickness certificate
A reported work absence >7days and a sickness certificate

Consistent match
A reported work absence of <7 days and a sickness certificate
A reported work absence of <7 days and no sickness certificate

Mis-match
No reported work absence and a sickness certificate
A reported work absence of >7 days and no sickness certificate

Figure 6:8 Criteria used in marching the self-report and GP sickness certification data sources

Further analysis on the primary self-reported and sickness certification data was carried out using the inter-rater agreement kappa statistic (Altman, 1995) to evaluate the magnitude of agreement between the two data sources. The kappa assesses the measure of agreement between two variables while adjusting for agreement by chance. The kappa statistic ranges from -1 to 1. In this study therefore, a higher kappa statistic reflected a higher degree of agreement between self-reported time off work and sickness certification records. It has been proposed that the strength of agreement measured by a Kappa statistic can be graded as < 0.20 = poor; 0.21 – 0.40 = fair; 0.41 – 0.60 = moderate; 0.61 to 0.80 = good; 0.81 – 1 = Very good (Altman, 1995).

A secondary analysis was carried out to evaluate the magnitude of agreement between the two data sources excluding self-reported days off work of less than 7 days. This was done to ensure that the same definition was used across the two data sources since sickness certification data does not include time off work episodes of less than 7 days.

6.7 Results

6.7.1 Characteristics of the data sources

The SBT included 836 patients who provided consent to have their records reviewed, of whom 516 (62%) were employed at baseline and 295 (35%) were employed at 12 months. The IBS dataset included 728 patients who provided consent to have their records reviewed, of whom 551 (59%) were employed at baseline and 256 (35%) were employed at 6 months. The mean age of the 295 in the SBT was 46 (range: 18-69); 56% female. Similarly, in the IBS, the mean age of the 256 in the SBT was 48 (range: 19-72); 59% were female.

6.7.2 Summary of self-reported absence and sickness certificates measures

As shown in Tables 6.5 and 6.6, approximately 25% of the respondents in both datasets reported having had an absence from work episode (26 % and 25 % for SBT and IBS respectively). On the other hand, few participants in both studies reported receiving a sickness certification during the study period (7 % and 7 % for SBT and IBS respectively).

Table 6:5 Self-reported absence from work for employment patients

Self-reported Absence	Number	Percent	Number	Percent
	STarT Back Trial (n=295) 12 months recall		IMPaCT Back Study(n=256) 6 months recall	
Yes	78	26%	67	25%
No	217	74%	191	75%

Table 6:6 Sickness certificates issued during study period

Sickness certificate	Number	Percent	Number	Percent
	STarT Back Trial (n=295) 12 months period		IMPaCT Back study (n=256) 6 months period	
Yes	21	7%	17	7%
No	274	93%	239	93%

6.7.3 Overall agreement between the electronic medical records and self-report

Table 6.7 presents data on matching of self-reported work absence periods with sickness certification records from the GP certification database. Relatively similar agreement levels were observed in the two study datasets. Among employed individuals, 73% and 72% (for SBT and IBS respectively) had direct matches, where a reported absence greater than 7 days was associated with a sickness certificate and no sickness certificate was recorded for records without a reported absence. However, a relatively high mismatch rate of 15% and 13% was observed for the SBT and IBS databases respectively.

A significant proportion of the inconsistent matches were found to be self-reported work absence days greater than seven days, which warranted a sickness certification but had no record of a sickness certificate. A hand search of the consultation free text showed the certificates were not documented in the electronic database and therefore could not be matched. The remaining minority of records (3% and 3% for the SBT and IBS) were reported sickness certification records for which no self-reported work absence was given. Thus the results seem to imply that agreement between self-reported and sickness certification records over long periods of time are poor.

Table 6:7 Agreement of self-reported work absence with sickness certificates for the study datasets

Sickness absence (SA) according to Self-report record	STarT Back(12 months recall) *n = 295*		IMPaCT Back(6 months recall**)** *n =255*	
	SA based on certification database - Sickness Certificate			
	Yes	No	Yes	No
Yes No	10 11	68 206	10 7	54 184
	Sickness certificates matching categorized self-reported work absence period			
< 7 days > 7 days No	0** 10* 11***	36** 32*** 206*	0** 9* 7***	29** 25*** 184*
	Percentage agreement			
Direct match Consistent match Inconsistent/Mis-match	73% 12% 15%		72% 11% 13%	

*Direct Match **Possible Match ***Inconsistent Match

6.7.4 Association between self-report and sickness certification absence episodes using the inter rater kappa

Kappa statistics obtained for the overall sample of employed patients were 0.085 (SE: 0.044; $p < 0.05$) and 0.158 (SE: 0.048; $p < 0.001$) for the SBT and IBS datasets respectively. The observed values demonstrated a poor ($k < 0.20$) level of agreement according to the classification system by Altman, (1995).

The corresponding kappa values excluding self-reported time off work values less than 7 days were 0.2074(SE: 0.570; $p < 0.001$) and 0.292 (SE:0.061; $p < 0.001$). The adjusted datasets demonstrated a fair ($0.21 < k < 0.4$) level of agreement between the two data sources. Overall, the additional analysis on the primary data variables shows further evidence of low comparability between sickness certification records and self-reports of work absence.

6.8 Discussion

6.8.1 Estimation of a friction period for the UK

This chapter aimed to contribute to knowledge on the application of the friction cost approach by estimating a length of friction period for the UK, and also more detailed friction periods stratified by occupational level categories. Vacancy duration estimates were collected from a number of organisations and used to estimate the length of a friction period in the UK. The overall average duration of a friction period incorporating all data sources was 2.7 months. National vacancy duration statistics estimated from DLA Piper LLP, the Office of National Statistics, Chartered Institute of Professional Development and the University of Birmingham in the UK, show friction periods for all occupations combined as ranging between 2.0 to 3.2 months. The findings showed wide variations in the friction period across occupational categories in some of the data sources (see Figure 6.3 to 6.7), highlighting the need to assess the impact of using more detailed friction period values in economic evaluation studies applying the friction cost approach. Friction period estimates were summarised in three formats including: i) an overall average friction period for all occupational categories adjusted to national-level estimates ii) a friction period for two levels of occupational categories, iii) a friction period for each dataset split into five occupational level categories. Such an analysis has not been performed before.

The results provide insight into the range of vacancy duration estimates from different data sources, and the variations in friction periods between occupational categories. The findings show differences in vacancy duration and hence friction periods between occupational classifications with higher occupational levels including senior managers, professional and technical occupations, showing longer friction periods compared to lower level occupations including skilled, services and elementary occupations. The results reflect higher unemployment for lower level occupational categories and a longer friction period needed to employ higher level occupational workers. Higher level occupational category jobs such as senior managers and professional occupations generally have a smaller pool of individuals to choose from, which might result in longer recruitment periods compared to lower category occupational jobs and this could have influenced this variation. In addition, such occupations might have longer notice to leave periods for their employees. Indeed the data set from the University of Birmingham showed higher job notice period estimates of six to twelve weeks for this category of employees.

Slightly lower estimates were observed in the unadjusted ONS and DLA piper LLP datasets which could be attributed to the absence of the period between a job offer and the actual start date from these datasets. Currently, there is no reliable information available in the UK about the average duration between individuals accepting a job offer and starting employment. Therefore, this estimate was sought from published literature in other countries. Previous researchers have used an average duration of four weeks to allow for this period (Koopmanschap and Rutten, 1996; Hanly et al., 2012). The additional findings from the University of Birmingham dataset agreed with previous estimates in the Netherlands showing broadly similar estimates of three to five weeks for the average period between accepting a job offer and starting work (Koopmanschap et al., 1995). In contrast, higher estimates (six to twelve weeks) were observed for occupations specific to academic and research related occupations.

To date, there is little evidence on friction period estimates within the UK. Studies that have used the friction cost approach have reported a friction period of about 90 days (Maniadakis and Gray, 2000) and 180 days (Lewis et al., 2007) without providing detailed information on the sources of the friction period estimates.

Few researchers have studied vacancy durations or provided comprehensive estimates of vacancy durations (Beaumont, 1978; Roper, 1988; van Ours and Ridder, 1991; Erdogan-Ciftci and Koopmanschap, 2011). Van Ours and Ridder used a statistical model analysing grouped vacancy duration data to estimate vacancy durations according to education level for two occupations (van Ours and Ridder, 1991). The data used in the model were based on a sample of data from a survey. They found that, in general, employers need more time to fill vacancies that require higher educational levels and vacancy durations are more responsive at higher educational levels. However, the duration data used in this model were limited only to the period of the survey.

Further work by Koopmanschap et al., (1995) in the Netherlands, followed van Ours' model by using completed vacancy durations by education level for 1988 and 1990 to estimate vacancy durations. The analysis was based on quarterly data on the number of vacancies and the uncompleted vacancy durations from a large number of Dutch firms (Koopmanschap et al., 1995). This study added an additional period to allow for the time taken in raising a vacancy as well as the time in between filling a vacancy and an individual starting work. A recent conference abstract reported estimated vacancy durations in 2009 for the Netherlands, Belgium, Germany, France, the UK, Norway and Sweden as ranging between

40-80 days using a combination of national aggregate stock and time series data on vacancies (Erdogan-Ciftci and Koopmanschap, 2011).

The research within this Chapter shows that the average friction period is broadly similar to the overall average friction period from these studies as well as from studies as discussed in Chapter Three of this thesis. Nevertheless, this research differs in a number of ways. Firstly, they study either a cross-section of vacancy durations or patterns of vacancy durations and vacancy flows. This research considers both approaches from a variety of data sets and hence the findings here complement and strengthen these studies. Secondly, they study either average vacancy durations or vacancy durations by education level. This research considers variations in vacancy stocks at UK national level and vacancy durations are split into occupational categories of the labour market which has not been performed before and therefore provides more detailed vacancy duration values and hence friction periods. Overall, the vacancy durations from the four sources provide estimates that can be used in the application of the friction cost approach within the UK. Consequently, if friction periods differ according to occupational levels as has been shown by the vacancy duration estimates presented here, then the current practice of using a single friction period for the labour market could potentially have an impact on overall friction related costs and on overall cost-effectiveness estimates.

The analysis presented here has some strengths and limitations. One strength of the current study was the diverse range of data sources from which vacancy duration data for various levels of occupational categories were collected for the UK. In addition, the datasets were adjusted to national level estimates for the different occupational categories based on data from the office of National Statistics. A limitation of the study is that data were obtained from different sources that applied different methods in collecting the vacancy duration estimates. Ideally, a comprehensive survey covering the various occupational categories would provide more rigorous friction period estimates but this was beyond the scope of this work and the available resources. However, the selection of data sources included in this study span a range of occupational categories at different levels.

Moreover, the data were adjusted to national level data using aggregate vacancy statistics comparable to national UK values. Another drawback of this study is that the friction period results were categorised based on a broadly classified occupational classification system. A more detailed analysis could have been presented if the respondent's occupational status had been coded at a more detailed occupational level classification. This could have been achieved through more precise structuring of work questions on the individual questionnaires in order to obtain extra information such as whether the respondent is self-employed, is an employee, the size of the organisation, and supervisory status. However, obtaining these extra data could potentially come at a cost of reduced response rates.
A further limitation with the data sources used is that they only presented limited summary data thus limiting the presentation of more detailed information such as standard deviation and other measures of dispersion. However, limited data on medians, Lower quartile and Upper quartile ranges was available from the DLA Piper data source.

Given that very few economic evaluation studies have applied the friction cost approach, mainly due to uncertainties in estimating a friction period (Hanley et al., 2012), these findings provide estimates that will guide future work and research in the application of the friction cost approach in the UK. Furthermore, in current practice of economic evaluations using the friction period, an overall average friction period is commonly used and often depends on the availability of data. The results from this study can be used to further develop standard friction period estimates for different occupational categories within the UK. In the current practice and reporting of economic evaluations, it is not often possible to identify the validity and source of friction period estimates used. Finally, the lack of readily available empirical data could potentially be one of the main reasons the friction cost approach has rarely been used in the United Kingdom. Data presented here and methodological aspects discussed, can therefore help to overcome some of these drawbacks as well as stimulate further research.

6.8.2 Feasibility of using sickness certification records in identifying friction periods

A second aim was to contribute to the scarce research on comparability between sickness certification data from GP records and self-report work-related data. This investigation is one of the first comparisons of self-reports and sickness certification records in a clinical trial and cohort analysis setting within the United Kingdom. The motivation for this investigation was to explore how to use retrospective sickness absence data from datasets that have not collected durations of individual episodes in order to estimate the occurrence of friction periods within economic evaluation studies. This is especially necessary when valuing productivity costs using the friction cost approach.

The results show that, although self-reports and electronic sickness certification records are comparable, the discrepancies found between the two data sources were much higher than were typically found in other studies. For example, the only study identified that compared self-reports with sickness certification records by Wynne-Jones et al., (2008) reported a 95% match of sickness certification medical records with self-reported records. However, this may be due to the recall period considered in that study. In comparison to the investigation in this thesis that compared self-reported absences over a period of 6 and 12 months, their study compared sickness certification records with self-reported time off work over a much shorter recall period of 2 weeks. Previous research has shown that recall periods longer than 3 months are subject to potential errors of recall bias (Severens et al., 2000). However, there is also potential overestimation when extrapolating work absence data collected over shorter periods of time, making a trade-off necessary, unless it is feasible to collect detailed data over a longer period of time (Braakman-Jansen et al., 2012).

Most of the research investigating agreement between resource utilisation data sources has focused on comparing self-reports with the healthcare provider or employee healthcare databases. One such study in the UK (Ferrie et al., 2005), found relatively good agreement between incidence and duration of self-reported work absence data over 12 months when compared with employee recorded sickness absence. In contrast, others (van Poppel et al., 2002; Grovle et al., 2012) reported poor agreement between self-reports and sickness absence database records for periods as low as 3 months. Although relatively good accuracy of data on the duration of sickness absence has been shown in employment sickness absence registers, practical challenges in accessing these data exist particularly within a clinical trial setting where employee data would be required from many different organisations.

The feasibility study presented in this thesis investigating the role of sickness certification records in assessing work absence in comparison to self-report was based on sickness certification data extracted alongside a large clinical trial and cohort study within primary care, with a diverse set of patients captured in the certification registry. Nevertheless, there were inherent problems with the sickness certification and consultation data. The main limitation of this investigation was the lack of information on the duration of sickness certificates and therefore reliable conclusions cannot be made on the differences in the magnitude of work absence between the two data sources. Further examination of the manually prepared sickness certification records, indicated that the most important source of discrepancy appeared to be the failure to accurately document certificates issued and the associated length of the sickness certificate. As a result, it was also not possible to use the sickness certification records to identify individual sickness episodes linked to self-reported time off work. In addition, one can question the quality and completeness of the database or the data extracted as the additional data had to be manually extracted from consultation medical notes.

However, care was taken to review all the relevant certification, consultation data sets and consultation free text for back pain related diagnoses and dates of sickness certificates. A further limitation was that, where included, it was not possible to identify whether an individual would spend the entire certified period away from work or might potentially return to work before the certified period expired.

These limitations mainly stem from the nature of the sickness certification system in the UK. Currently patients can self-certify sickness for an absence period of 7 days or less and would only be legally required to obtain a doctors fit for work note, formerly a sickness certificate, after an illness of more than 7 days. Additionally, doctors who issue sickness certificates are not mandated to record the duration of the sickness certification or to obtain information as regards to when individuals return from a sickness certification period.

Overall, it is evident from the findings in this chapter that sickness certification data in the UK often lacks information on the duration of sickness absence. Based on these findings, it would appear that medical certification data is not tailored for health economic research as data on key aspects such as duration of absence and occupation are not always accurately documented. However, because this analysis focused on a specific population, it cannot be assumed to be representative of the entire UK population. A wider exploration with a larger set of practices might be warranted to further explain the relationships between self-reports and sickness certification records.

From the findings in this investigation, time off work self-reports are potentially a more accurate data source than are sickness absence records and offer the best possibility for assessing time off work in economic evaluation studies, unless the certification recording system systematically improves to capture all sickness certifications and their associated durations. On the other hand, self-reports are subject to possible recall bias and non-response issues, and when used there are possibilities of under-reporting absence episodes especially over longer periods of time leading to underestimation of productivity costs. Conversely, there are also possibilities of overestimating productivity loss as a result of poor recall and inclusion of sickness absence episodes falling outside the relevant study period. A further complication is that sickness absence self-reports normally do not include data on the duration of specific sickness episodes unless specifically requested from the onset. For economic evaluation studies estimating productivity costs based on the friction cost approach and using self-reports, the tendency to ignore the capture of specific absence episodes should therefore be taken into account at the design stage to the study.

In the absence of duration data on episodes of work absence, a pragmatic approach was taken in this thesis assuming that the self-reported sickness absence comprised a single sickness absence episode. This sort of assumption may be more accurate in chronic diseases like back pain associated with both episodic and long periods of work absence, than for some other conditions.

6.9 Conclusions

This chapter has provided data about friction period estimates in the UK from different data sources and employment category levels. The findings showed vacancy durations, hence friction periods, differ between employment category levels and that this pattern is relatively similar across different data sources. A greater length of the average friction period was observed for higher occupational level jobs. This reflects that the friction period is sensitive to occupational level categories and the associated productivity losses may vary when more detailed friction periods are used. Obviously more research (a wider survey covering various occupational categories across different sectors) could be conducted to see if similar results are found, but the empirical work presented in this chapter will contributes to the debate on using the friction cost approach and knowledge as to how researchers can use more detailed friction periods when valuing productivity costs. An unanticipated conclusion is that the average friction period estimated has not changed considerably over recent years, despite the change in economic conditions.

The results also showed significant discrepancies between self-report records and sickness certification records. It also appears that the current medical certification registry does not accurately capture the duration of sickness certifications and therefore cannot be used to identify individual absence episodes, which are necessary in valuing productivity costs when using the friction cost approach. Medical certification records therefore do not provide a more practical solution to common challenges with self-reports in economic evaluation studies. Additional research (covering a more diverse set of GP practices) could be carried out to provide more insight into the comparability of self-reports with sickness certification data, but it is hoped that the results here generate an increased interest in reviewing GP certification records and a way forward for their use in economic evaluation research. The findings in this chapter highlight the difficulties in obtaining accurate information on the length of a sickness period in current economic evaluation research. Greater involvement of economists in the design of data collection instruments could provide more detailed and accurate data on friction periods. Chapter Eight assesses the potential impact of using more detailed friction periods within an economic evaluation carried out from a societal perspective. The next chapter reports empirical work assessing the construct validity and responsiveness of the single-item presenteeism question among low back pain sufferers.

CHAPTER SEVEN CONSTRUCT VALIDITY AND RESPONSIVENESS OF THE SINGLE-ITEM PRESENTEEISM QUESTION IN PATIENTS WITH LOWER BACK PAIN

7.1 Introduction

The overall aim of this chapter is to evaluate alternative methods used in the measurement of lost productivity with particular emphasis on reduced productivity at work. The inclusion of a single-item standalone question and a set of work and demographic questions in the STarT Back Trial and IMPaCT Back Cohort facilitated an exploratory investigation of the construct validity of the single-item presenteeism question. In this chapter, a rationale for assessing the validity of the single-item presenteeism question among low back pain sufferers is provided. This is followed by a brief overview of literature on psychometric properties of reduced productivity measures. Next, the methods and results of the validity and responsiveness of the single-item presenteeism measure are explored and presented. Finally, the discussion focuses on the implications of these findings and their application for work in the subsequent chapter.

7.2 Rationale

The literature review in Chapter Four of this thesis shows that the concept of presenteeism has received relatively little attention in comparison to absence from work, particularly within economic evaluation studies. The evidence from the review in Chapter Four shows that presenteeism costs comprise 28% to 85% of total costs of disease, and were often greater than absenteeism costs. It is therefore important that presenteeism is measured appropriately within economic evaluation studies. However, the methodological challenges in the measurement and valuation of presenteeism have meant that these costs are treated with caution within economic evaluation research (Brooks et al., 2010)

A range of tools have been developed to measure presenteeism as identified in Chapter Four of the thesis, and their psychometric properties of feasibility, reliability, and validity have been reported in previous reviews (Lofland et al., 2004; Beaton et al., 2009; Prasad et al., 2004; Mattke et al., 2007; Tang et al., 2009). The majority of such tools use multi-item presenteeism questions which have the drawback of being associated with difficulties in interpretation, increased responder burden, and the potential risk of asking irrelevant questions for certain occupations (Bowling, 2005). Conversely, the advantages of tools incorporating single-item questions, namely simplicity, ease of interpretation and reduced burden, come at the expense of obtaining more detailed information and potentially reducing sensitivity to change (Bowling, 2005).

This chapter investigates the validity of a single-item presenteeism question (SIPQ) as a standalone measure outside standard presenteeism questionnaires obtaining information over a period of 30 days. The SIPQ is a numeric rating scale asking respondents to rate their job performance over a specified period of time on a scale of 0 to 10. It was adapted from the standardised World Health Organisation Health and Work Performance (HPQ) questionnaire (Kessler et al., 2003). The HPQ single question asks respondents to rate their overall performance on days worked during the past 4 weeks. The wording from this question was adapted for a low back pain population group and to ask about a 30 day period

(Figure 7.1). A similar question is incorporated in the Work Productivity and Activity Impairment (WPAI) standardised questionnaire with a shorter recall period of seven days (Reilly et al., 1993).

Overall, few presenteeism measures have been evaluated within the context of back pain (Lerner et al., 2001; Beaton and Kennedy, 2005; Hagberg et al., 2002). The HPQ tool has been evaluated in chronic diseases such as arthritis, depression, chronic headaches and asthma, but not among back pain patients (Kessler et al., 2004). Similarly, the WPAI measure has been validated in a variety of diseases (Reilly et al., 2003; Reilly et al., 2004; Wahlqvist et al., 2007; Reilly et al., 2008; Reilly et al., 2010) but not yet among back pain patients. Such a validation is important because back pain is a common disorder affecting the working age population (Wynne-Jones, 2008), and is a leading reason for working individuals who consult their general practitioner in countries such as the United Kingdom (UK) (Foster et al., 2010). Practical methods of assessing the costs of reduced productivity, particularly when an economic evaluation is from a societal perspective, are needed in determining the full economic burden of back pain (Loeppke et al., 2007).

This chapter therefore reports:

1) An assessment of the validity of using a single-item presenteeism question in measuring at work reduced productivity due to back pain related symptoms with a 30 day recall period in two UK studies

2) An assessment of the responsiveness of the single-item presenteeism question for which there is currently no evidence in the context of back pain.

7.3 Psychometric testing of outcome measures

Psychometric testing has been widely used to measure subjective concepts such as health perception and attitudes, and has been used in measurement constructions, and the testing of established health status measures such as the SF-36 health survey, and EQ-5D (Harrison et al., 2010; Boonen et al., 2007). A range of criteria is used in assessing the performance of measures within the psychometric literature including reliability, validity and responsiveness (Brazier et al., 1999). Reliability is used to assess both random and systematic error inherent in any measurement (Streiner and Norman, 2006).

Responsiveness is defined as an instrument's ability to detect a meaningful or clinically important difference when changes occur over time (Lindeboom et al., 2005). The validity of an instrument assesses if an instrument is able to measure what it is supposed to measure within that particular situation (Streiner and Norman, 2006). There are different types of validity including content validity, criterion validity and construct validity. *Content validity* refers to the instrument's ability to measure all related characteristics of what is being measured (Brazier et al., 1999). *Criterion validity* assesses the association of a measure with some other alternative measure, usually a gold-standard measure within the area of study (Streiner and Norman, 2006). Alternatively, *construct validity* assesses whether the relationship between the scores of a measure and other measures or concepts being investigated correlates as hypothesized (Streiner and Norman, 2006). As there is currently no gold standard for assessing presenteeism, construct validity rather than criterion validity was assessed in this particular investigation.

7.4 Methods

7.4.1 Study population and design

Data were obtained from the Subgrouping for Targeted Treatment (STarT) Back Trial (Hill et al., 2011) and the IMplementation study to improve Patient Care through Targeted treatment for back pain (IMPaCT) Back Study (Foster et al., 2010) as described in Chapter Five. Briefly, The STarT Back Trial (SBT) was a primary care, multicentre, pragmatic randomised controlled trial investigating subgrouping for targeted treatment versus current best care (non-targeted) physiotherapy. The IMPaCT Back study (IBS) was a quality improvement cohort study investigating a new care system of subgrouping and targeting treatment for low back pain patients in routine primary care services.

7.4.2 Patient reported outcome measures

Patients completed self-administered questionnaires at baseline and at defined subsequent time-points throughout both studies. These questionnaires are summarised in Table 7.1 below and this summary is followed by a detailed description of the questionnaires in the next section.

Table 7:1 Summary of patient outcomes and instruments used in the SBT and IBS Studies

Construct	Outcome measure/Instrument	Summary description of instrument
Disability	Roland Morris disability questionnaire (RMDQ)	A 24 item questionnaire scoring 0 or 1. Higher scores represent high disability levels
Generic HRQOL	EuroQol (EQ-5D)	A multi-attribute generic health utility index with Mobility, self-care, usual activities, pain/discomfort, anxiety/depression. Higher scores represent better HR-QOL
Generic Health Status measure	SF-12	A shorter version of the SF-36 with a subset of questions from the SF-36 measure. It covers both physical (PC-12) and mental (MC -12) summary aspects.
Pain Severity	Numeric rating scale	An 11 point scale anchored by 'no pain' (0) to 'pain as bad as could be' (10). (past 2 weeks, usual pain, and present time)
Psychological health – Pain fear	Tampa Scale for Kinesiophobia(TSK)	A 17-item scale scored on a 4-point Likert scale anchored by 1 "strongly disagree" to 4 "strongly agree. Higher scores represent greater Kinesiophobia levels
Generic psychological health - Anxiety and Depression	The Hospital Anxiety and Depression questionnaire (HAD)	A 14 item measure of generalised anxiety (7 items) and depression (7 items). Higher scores represent higher anxiety and depression problems
Pain catastrophising	Pain catastrophising scale (PCS)	A 13- item measure each with a 5 point scale ranging 'Not at all' to 'all the time'. Higher scores represent greater Catastrophising
Change in health status	Global change numeric scale	A global question with a 6 point scoring scale anchored by 'completely recovered' to 'much worse' (6)

All items included in the START Back and IMPaCT Back databases except for the education component not collected in the IMPaCT Back cohort

7.4.2.1 The single-item presenteeism question

A single-item question asked respondents to estimate their reduced productivity at work due to back pain over the past 30 days. The question, administered at baseline and follow-up, comprised a single question presented on an 11-point Likert-type scale anchored by 'Not at all' (0) and 'The worst you can imagine' (10). Higher scores indicate a greater impact of back pain on work performance. The measure was used to assess the impact of back pain on work during the past 30 days. The question is shown in Figure 7.1.

“To what extent has back pain affected your **performance at work** over the past 30 days? Please rate this on a 0-10 scale, where 0 is ‘Not at all’ and 10 is ‘The worst you can imagine’.” *(Please cross one box)*

Not at all										The worst you can imagine
0	1	2	3	4	5	6	7	8	9	10

Figure 7:1 The Single-Item Presenteeism Question wording

7.4.2.2 Roland- Morris Disability Questionnaire (RMDQ)

Disability was measured using the RMDQ (Roland and Morris, 1983). The RMDQ is a self-administered 24-item questionnaire with yes/no items specifically related to physical functions. A RMDQ summary score is obtained by adding the items checked. The final score ranges between 0 (no disability) to 24 (severe disability). The psychometric properties of the RMDQ have been established in low back pain patients (Roland and Morris, 1983).

7.4.2.3 EuroQol (EQ-5D)

Health status was assessed using the EQ-5D. The EQ-5D is a self-completed generic health status measure with 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and with index values based on value sets elicited from a sample of the UK adult population (Dolan et al., 1996). Each dimension has 3 levels of severity. Scores in the UK value set range from -0.59 (worst) to 1.0 (perfect health) (Dolan, 1997). Higher scores indicate better HR-QOL outcomes. The established and validated 3 level version was used here. The EQ-5D has been found to be valid and responsive in low back pain patients (Obradovic et al., 2013).

7.4.2.4 SF-12

Functional well-being and the general well-being of patients was measured using the SF-12 (version 2).The SF-12 (v2) is a shorter version of the generic SF-36 health profile instrument (Ware et al., 1996). The SF-12 generates two summary measures: the physical (PCS-12) and mental (MCS-12) component summaries from the 12 questions. SF-12 scores range between 0 and 100, with higher scores indicating better quality of life (HR-QOL). The SF-12 has been found to be valid and responsive in low back pain sufferers (Obradovic et al., 2013).

7.4.2.5 Numerical rating scale (NRS) for back pain activity

The back pain is an 11 point numerical rating scale is a scale that allows individuals to rate their pain from 0 (no pain) to 10 (Pain as bad as could be) (Childs et al., 2005). Questions are focused on the intensity of least painful back pain in the last 2 weeks and the intensity of back pain at the present time using the NRS.

7.4.2.6 Tampa Scale for Kinesiophobia (TSK)

The TSK is used to assess fear of movement (Kinesiophobia) in pain sufferers (Hudes, 2011). The instrument is a 17-item self-completion questionnaire with each item scored on a 4-point Likert scale with scores ranging from 1 "strongly disagree" to 4 "strongly agree" (Feleus et al., 2007). The Overall scores range from 17 to 68, with higher scores indicating increasing degrees of Kinesiophobia. The TSK has been shown to be valid in low back pain sufferers (Roelofs et al., 2004).

7.4.2.7 Anxiety and Depression Scale (HADS)

The HADS was used to assess anxiety and depression in back pain patients. The HADS is a 14-item self-report measure used to assess anxiety (7 items) and depression (7 items) (Snaith, 2003). Each item is scored on a 4-point Likert scale with maximum subscale scores of 21 indicating the highest levels of depression and anxiety.

7.4.2.8 The Pain Catastrophising Scale (PCS)

Catastrophising involves worrying about negative outcomes from a situation (Turner et al., 2001). The patients' use of catastrophising strategies to cope with pain was assessed using the PCS which contains 13 questions on thoughts or feelings. Each question is asked on a 5 point scale from 0 ('not at all') to 4 ('all the time') and scores are summed to provide a total score (Sullivan et al., 1995). Scores range from 0 (lowest) to 52 (highest level of pain catastrophising) (Snaith, 2003).

7.4.2.9 Global change question

A global retrospective assessment question of treatment effect was used to assess the patients' perception of the overall change in their back pain condition. A 6-point Likert scale anchored by 'completely recovered' and 'much worse' was used and collapsed into a dichotomous variable with two categories 'improved' (completely recovered to better) and 'not improved' (no change to much worse). The variable was used as an external anchor for evaluation of responsiveness as in previous validation studies (Beaton and Kennedy, 2005; Linde et al., 2008).

7.4.3 Analysis

Alpha was set at .05 unless otherwise stated. All statistical analyses detailed below were carried out using STATA software version 12 and SPSS software version 19. *A priori* hypotheses were established and presented after each construct. Analysis was restricted to the employed population for whom data on presenteeism were collected as part of the data sets. A paired (samples) t-test based on matched individuals was used to test for the difference in reduced productivity between responders employed at baseline and follow-up.

7.4.4 Validity

Construct validity of the SIPQ was assessed by providing evidence of convergent and divergent validity. Convergent and divergent validity were tested by exploring whether the presenteeism tool was correlated with measures hypothesized to be related (convergent) or unrelated (divergent). Evidence of convergent validity was established by a significant and strong correlation between measures based on stated prior hypothesis (Marra et al., 2005), while divergent validity was determined by an insignificant and weak correlation between measures (Streiner and Norman, 2006).

Correlation analysis was conducted between the presenteeism measure, and standard reference constructs included in both studies for pain (NRS), disability (RMDQ), fear avoidance (TSK), anxiety and depression (HADS), Catastrophising (PCS) and quality of life (SF-12, EQ-5D). The measure was also correlated against non-clinical variables including age, education, gender and patient risk group. Correlation coefficients were calculated using Pearson's correlations for normally distributed data and Spearman's rank correlations for non-normal data. Correlations were classified as very strong ($|r| > 0.8$), strong ($0.6 < |r| < 0.8$), moderate ($0.4 < |r| < 0.6$), weak ($0.2 < |r| < 0.4$), or absent ($0.0 < |r| \leq 0.2$) (Aggarwal et al., 2009). Correlations were performed for interval scores in evaluating for cross-sectional validity.

Previous validation studies have shown productivity loss to be related to disease severity and quality of life (Prasad et al., 2004; Reilly et al., 2010). A number of *a priori* hypotheses of expected relationships were therefore developed and tested. It was hypothesized that the SIPQ would be highly correlated with: 1) pain intensity and severity 2) disability 3) disease specific measures and 4) quality of life. SIPQ was further hypothesised to be moderately correlated with fear and avoidance (measured by the TSK) and depression/anxiety (measured by the HADS). Such associations would reflect convergent validity of the SIPQ. In contrast, previous research has demonstrated that age, gender and education have relatively little impact on sickness presenteeism (Aronsson et al., 2000). A further hypothesis was that there would be an insignificant correlation between the SIPQ and these variables. Evidence in this context would reflect divergent validity of the SIPQ.

7.4.4.1 Presenteeism scores and EQ-5D

Lower levels of quality of life in back pain patients have been associated with efficiency loss and absenteeism (Lamers et al., 2005). Correlation analysis was therefore used to explore the relationship between presenteeism and EQ-5D at baseline and follow-up. The SIPQ was expected to be highly correlated with the EQ-5D.

7.4.4.2 Presenteeism scores and general health scores

Investigations were also carried out to compare the dimensions within the SF-12 such as physical limitations, psychological health problems at work, and general health satisfaction. Since pain in the lower-back generally affects physical health to a greater extent than mental health (Riddle et al., 2001), presenteeism was expected to be highly correlated with the physical summary component (PCS).

7.4.4.3 Presenteeism scores and disability

Previous studies have associated back pain with chronic illness which leads to disability and impairment at work (Croft, 1998; Mallen et al., 2007; Lin et al., 2011). The SIPQ was therefore hypothesised to be highly correlated with the disease specific RMDQ. This measure was also expected to be more highly correlated with the RMDQ than with the generic EQ-5D measure as disease specific measures have been known to have overall greater responsiveness than generic instruments (Taylor et al., 1999).

7.4.4.4 Presenteeism scores and pain characteristics of pain intensity and severity

Pain intensity (or pain severity) has been reported as a predictive factor for poor prognosis among musculoskeletal patients (Mallen et al., 2007). These characteristics have been associated with self-reported reduced productivity (Hagberg et al., 2002; Mannion et al., 2009). The SIPQ was therefore hypothesised to be highly correlated with pain intensity and it was also hypothesised that the severity of back pain would be associated with significantly high levels of reduced productivity.

7.4.4.5 Presenteeism scores and psychological health

Evidence of the influence of psychological factors, such as fear avoidance and anxiety, on back pain related reduced productivity has been conflicting. Some have found no impact of depression or anxiety on reduced productivity (Hagberg et al., 2002) while others found that fear avoidance beliefs about work and physical activities were significantly associated with both work absence and reduced productivity (Mannion et al., 2009). It was hypothesized that the SIPQ would be moderately correlated with fear and avoidance (measured by the TSK) and depression/anxiety (measured by the HADS).

7.4.4.6 Presenteeism and other factors – age, gender, education

Previous research has demonstrated that age, gender and education have relatively little impact on sickness presenteeism (Aronsson et al., 2000). It was therefore hypothesised that there would be an insignificant correlation between the SIPQ and these variables.

Discriminant validity of the presenteeism measure was tested using the known-group validity method by assessing the ability of the SIPQ to discriminate between pre-defined groups that were expected to differ (Streiner and Norman, 2006). Patients were categorised into subgroups of low risk, medium risk and high risk based on the presence of risk factors for chronicity, in order to identify patients at risk of work absence due to back pain sickness for targeted treatment. Using pre-defined groups from the validated STarT Back tool (Hill et al., 2008), it was hypothesised that the SIPQ would discriminate among known groups by showing greater presenteeism in patients thought to have more severe back pain (i.e. the 'high risk' group and 'medium risk' group, compared to the 'low risk') ($P < .05$). Sensitivity to the known differences was assessed using analysis of variance (ANOVA) with adjustments for multiple testing based on the Bonferroni correction method (Bland and Altman, 1995). Furthermore, using t-tests, known-group validity was also assessed by comparing participants who had improved in their back pain condition with those who had not improved according to a global response question and a change in RMDQ score described in the responsiveness analysis detailed below. *A priori* hypothesis was that the SIPQ would discriminate among patients classified into these known-groups.

7.4.5 Responsiveness and sensitivity to change

Responsiveness of the SIPQ was also assessed to investigate the ability of the tool to detect changes over time. A responsive measure is able to detect minimal important changes in patients (either improved or worsened) over time (Revicki et al., 2006). A minimal meaningful improvement is the smallest difference that is perceived as a significant change in the health condition of patients (Streiner and Norman, 2006). For this analysis, comparisons between the SIPQ and other instruments were performed for two patient groups: those who perceived their health as having improved between baseline and follow-up, and those who considered their health as having not improved. The change in health was established from two indicators with the following items for assessing changes in health: (i) patient perceived change in back pain condition over time (patient rated global score question); and (ii) change in percentage RMDQ score (clinical endpoint).

The external patient global rating of change (GRC) question in health status was used to map change scores to changes in clinical measures (Revicki et al., 2006): a RMDQ > 30% change is considered an appropriate threshold for assessing patient improvement in health among back pain patients (Jordan et al., 2006). Responses from the GRC question were given on a 6-point scale ranging from 1 “completely recovered” to 6 “much worse”. Respondents who scored 1, 2 or 3 were classified as having improved, and those scoring 4, 5 or 6 were considered not to have had any improvement. Similarly, respondents with a RMDQ > 30% change were classified as having improved, and those individuals with < 30 % change as not having had any improvement.

In quantifying and examining responsiveness, two analytical approaches were used:

I. correlation analysis
II. statistical

In the correlation approach, correlation analysis was conducted between the presenteeism measures and the two external indicators for assessing change as in the construct validity analysis above. It was hypothesised that at least moderate correlations would indicate responsiveness of the measure. Once groups of low back pain sufferers were identified as either improved or non-improved based on the external indicators, statistical analysis was then used to further examine responsiveness.

In the statistical analysis approach (Hays and Revicki, 2005), statistical significance of the differences in the mean presenteeism change scores between ‘improved’ and ‘non-improved’ groups was established using the student *t*-test statistic. Distribution-based methods were applied based on statistical properties of the presenteeism measure. These methods included: (i) effect size (SES) (mean change divided by the standard deviation of the baseline score) (Kazis et al., 1989); and (ii) standardised response mean (SRM) (mean change divided by the standard deviation of the change score) (Cohen, 1988). Anchor based testing included the Guyatt responsiveness index (GRI) (mean change in ‘improved’ subjects divided by the standard deviation of change in ‘non-improved’ subjects) (Guyatt et al., 1987; Yost and Eton, 2005). Responsiveness of the measure was categorized as small when less than 0.5, moderate when between 0.5 and 0.8, and large when greater than 0.8 (Cohen, 1988). *A priori* hypotheses were that large ESs and SRMs for the SIPQ would be observed for patients who had improved and that small ESs and SRMs would be observed for patients who had not improved, as previous studies showed lower values for subgroups of patients who had not improved in self-reported measures (Beaton et al., 2010).

Longitudinal construct validity assesses the extent to which change over time correlates with other indicators of change. Responsiveness was further measured through correlation analyses of change scores of the SIPQ with change scores of other constructs (as described in the construct validity assessment detailed above); only patients employed both at baseline and follow-up were included. *A priori* hypotheses were that low correlations ($.30 < r < .50$) would be observed between the change scores of the SIPQ and change scores of the pain, disability scales and health status measures.

7.5 Results

7.5.1 Participant Characteristics

The back pain patient samples are described for employed patients in Table 7.2. The SBT sample included 851 patients, of whom 62% (n = 524) were employed. 275 (53%) of these patients reported taking time off work in the past 12 months. Only 285 responders reported being employed both at baseline and follow-up. The IBS included 922 patients, of whom 60% (551) were employed at baseline; 243 (44%) of these patients reported taking time off in the past 12 months (Table 7.2). Overall, 281 responders reported being employed both at baseline and follow-up (the IBS was characterised with a lower response rate than the SBT).

Table 7:2 Characteristics of patients employed at baseline assessment.

	STarT Back (N[a]=524)			IMPaCT Back (N[a]=551)		
Measurements	**n[a]**	**Number /Mean**	**Percentage / SD**	**n[a]**	**Number /Mean**	**Percentage) / SD**
Absenteeism	524	275	53%	547	243	44%
Gender (Female)	524	292	56%	551	301	55%
Age	524	43.69	11.34	551	46.80	11.20
RMDQ	524	8.79	5.40	551	7.93	5.70
EQ-5D	517	0.60	0.27	546	0.65	0.29
SF-12 PCS	522	39	10.08	546	41.74	10.57
SF-12 MCS	522	48.69	10.79	546	49.93	10.49
Pain intensity – Present time	522	6.34	2.32	549	4.26	2.76
Pain intensity- Past 2 weeks	521	4.48	2.49	550	5.72	2.72
Average pain severity	518	4.88	1.97	548	4.99	2.54
SIPQ score (reduced productivity at work)	520	4.82	3.00	530	4.47	3.01

a represents participant who self-reported being employed at baseline; RMDQ Roland Morris disability questionnaire, SD Standard deviation, PCS Physical component summary scale, MCS Mental component summary scale, STarT: Subgrouping for Targeted Treatment, IMPaCT: Implementation study to improve Patient Care through Targeted treatment

Baseline characteristics were based on a sample of patients employed at baseline comprising 524 of 851 (SBT dataset) and 551 of 922 (IBS dataset).

7.5.2 Distribution of responses across the 11 points of the SIPQ

Overall, scores for the SIPQ were distributed across the whole scale in both studies as shown in Figure 7.2, suggesting a capacity for broad discrimination between different intensities of presenteeism. At follow-up, several participants reported presenteeism scores distributed between the numbers 0 (not pain at all) and 5. The SBT sample included 520 and 292 patients (for baseline and follow-up respectively) who responded to the SIPQ, of whom 60% and 89% (for baseline and follow-up) reported numerical scores distributed between numbers, 0 and 5. A similar trend was observed in the IBS dataset which included 530 and 279 patients (at baseline and follow-up respectively) who responded to the SIPQ, of whom 61% and 86% (for baseline and follow-up) reported a numerical score between the numbers 0 and 5.

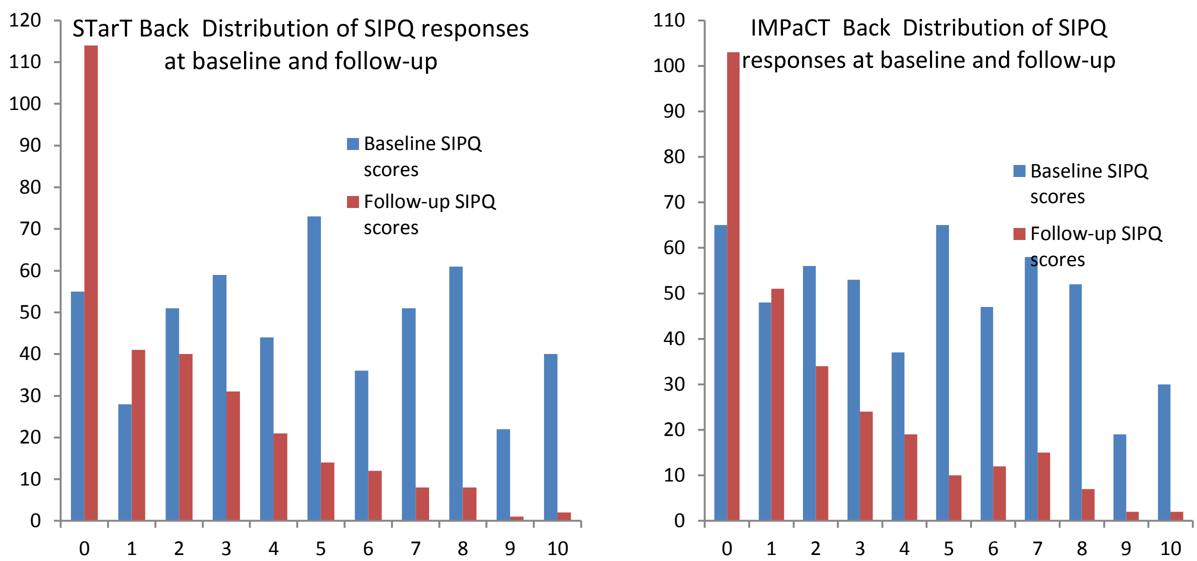


Figure 7:2 Overall distributions of SIPQ scores across the 11 points of the SIPQ

7.5.3 Reduced Productivity – Presenteeism

The difference in mean reduced productivity between patients employed at baseline and follow-up was calculated alongside the 95% confidence intervals (CI). This resulted in a mean difference of 2.46 (CI 2.09 to 2.82: *P* <0.001) and 2.25 (CI 1.87 to 2.61: *P* <0.001) for SBT and IBS respectively (Table 7.3). In both studies, the differences between the patients employed at baseline and follow-up were statistically significant (*P*<0.001) (Table 7.3).

Table 7:3 Reduced productivity for patients with back pain at baseline and Follow-up

STUDY	Measure	n[a]	Baseline [mean (95% CI)]	Follow-up [mean (95% CI]	Mean difference [Mean (SD)] [CI]	p-Value*
STarT Back	Presenteeism (employed patients baseline and follow-up 12 months)	283	4.43 (4.09, 4.78)	1.98 (1.70, 2.25)	2.46 (3.12) (1.87-2.61]	*P*< 0.001
IMPaCT Back	Presenteeism (employed patients baseline and follow-up 6 months)	262	4.38 (4.02, 4.74)	2.12 (1.83, 2.43)	2.25(2.62) (2.09-2.82)	*P* < 0.001

[a] values for reduced productivity changes calculated for data available at both baseline and follow-up

7.5.4 Construct Validity

Presenteeism was positively correlated with pain, psychological health (fear avoidance, anxiety depression, and catastrophising), disability, risk group categorization and time off work; and negatively correlated with quality of life (Table 7.4). Furthermore, the Spearman rank correlation between presenteeism and EQ-5D dimensions, risk group and time off work due to back pain showed a significant correlation. There was strong evidence of an association between reduced productivity and quality of life, disability, pain intensity and pain severity attributes as hypothesised. In addition, weak to moderate associations were observed between presenteeism and the general psychological measures (Table 7.4) and the anxiety and self-care dimensions of the EQ-5D. The highest correlations were observed between the SIPQ and physical summary component of the SF-12, the 'usual activities', and the 'pain' components of the EQ-5D ($0.29 \leq |r| \leq 0.66$) and in the disease specific measures related to disability ($0.55 \leq |r| \leq 0.70$) and pain ($0.38 \leq |r| \leq 0.77$) (Table 7.4). There was evidence of a weak association between the SIPQ and the mental health component ($0.21 \leq |r| \leq 0.29; P < 0.001$) (See Table 7.4).

Table 7:4 Construct validity: correlations between the SIPQ and other constructs at baseline and follow-up

	STarT Back Trial		IMPaCT Back Study	
Measurement (Pearson's Correlation)	**Baseline (N=520)**	**12- Month Follow-up (N=285)**	**Baseline (N=530)**	**6- Follow-up (N=279)**
Pain Dimensions				
Average pain severity	0.54 (*P*<0.001)	0.77 (*P*<0.001)	0.53 (*P*<0.001)	0.70 (*P*<0.001)
Intensity of back experience at present time	0.44 (*P*<0.001)	0.76 (*P*<0.001)	0.44 (*P*<0.001)	0.65 (*P*<0.001)
Psychological health				
Anxiety (HAD)	0.23 (*P*<0.001)	0.39 (*P*<0.001)	0.34 (*P*<0.001)	0.42 (*P*<0.001)
Depression (HAD)	0.35 (*P*<0.001)	0.47 (*P*<0.001)	0.45 (*P*<0.001)	0.53 (*P*<0.001)
Catastrophising (PCS)	0.35 (*P*<0.001)	0.56 (*P*<0.001)	N/A	N/A
Pain related fears (TSK)	0.34 (*P*<0.001)	0.40 (*P*<0.001)	0.35 (*P*<0.001)	0.42 (*P*<0.001)
Disability				
RMDQ score disability	0.53 (*P*<0.001)	0.70 (*P*<0.001)	0.54 (*P*<0.001)	0.70 (*P*<0.001)
SF-12 Dimensions				
physical component summary (SF-12 v2)	-0.56 (*P*<0.001)	-0.66 (*P*<0.001)	-0.55 (*P*<0.001)	-0.58 (*P*<0.001)
Mental component summary (SF-12 v2)	-0.21 (*P*<0.001)	-0.28 (*P*<0.001)	-0.29 (*P*<0.001)	-0.29 (*P*<0.001)
General health	-0.54 (*P*<0.001)	-0.62 (*P*<0.001)	-0.57 (*P*<0.001)	-0.59 (*P*<0.001)
Age	-0.01	0.03	-0.09	-0.03
Education	0.04	0.01	N/A	N/A
Gender	0.02	0.04	0.03	-0.01
Measurement (Spearman's Ranks Correlation)				
EQ-5D				
EQ-5D mobility	0.34 (*P*<0.001)	0.44 (*P*<0.001)	0.34 (*P*<0.001)	0.39 (*P*<0.001)
EQ-5D self-care	0.22 (*P*<0.001)	0.27 (*P*<0.001)	0.26 (*P*<0.001)	0.22 (*P*<0.001)
EQ-5D usual activities	0.36 (*P*<0.001)	0.50 (*P*<0.001)	0.43 (*P*<0.001)	0.51 (*P*<0.001)
EQ-5D pain	0.29 (*P*<0.001)	0.58 (*P*<0.001)	0.36 (*P*<0.001)	0.54 (*P*<0.001)
EQ-5D anxiety	0.14 (*P*<0.001)	0.27 (*P*<0.001)	0.22 (*P*<0.001)	0.32 (*P*<0.001)
EuroQol EQ-5D	-0.44 (*P*<0.001)	-0.60 (*P*<0.001)	-0.46 (*P*<0.001)	-0.62 (*P*<0.001)
Others				
Risk Group	0.49 (*P*<0.001)	0.41 (*P*<0.001)	0.47 (*P*<0.001)	0.49 (*P*<0.001)
Time off work	0.39 (*P*<0.001)	0.43 (*P*<0.001)	0.40 (*P*<0.001)	0.15 (*P*<0.05)

SIPQ Single-Item presenteeism question, RMDQ Roland Morris disability questionnaire, SD Standard deviation, HAD Hospital Anxiety and depression scale PCS Pain Catastrophising Scale, TSK Tampa Scale for Kinesiophobia, STarT: Subgrouping for Targeted Treatment, IMPaCT: Implementation study to improve Patient Care through Targeted treatment, N/A data not collected in the IMPaCT Back Study.

The single-item presenteeism scores correlated positively with education (*r* = 0.04 for SBT) and gender (*r* = 0.02, *r* = 0.03 baseline for SBT and IBS respectively) and negatively with age (*r* = -0.01, *r* = 0.03 for SBT and *r* = -0.09, *r* = -0.03 for IBS) (Table 7.4). These weak and insignificant relationships between the presenteeism tool and these variables (education, gender and age) showed divergent validity (Table 7.4).

The results of discriminant validity are summarised in Tables 7.5. At follow-up in both studies, mean scores of the SIPQ could differentiate between patient risk groups. The findings at baseline show that 'high risk' patients have statistically significant higher presenteeism than 'medium risk' and 'low risk patients' (*P* < 0.001) (Table 7.5). There was a significant difference between the medium and low risk groups and the high-risk and low risk groups after correcting for multiple testing. Similarly, mean scores of the SIPQ could differentiate between improved patients and patients who had not improved according to the global change question and RMDQ change score indicator. The findings showed that 'improved' patients have statistically significant higher mean presenteeism changes than 'non-improved' patients according to the global change to back pain (SBT: 3.25 versus 0.71, IBS: 2.85 versus 0.76; *p*<0.001) and the RMDQ score indicator (SBT: 3.16 versus 0.82; *p*<0.001, IBS: 3.22 versus 0.68) external indicators (Tables 7.8).

Table 7:5 Known-group validity of the SIPQ among risk groups at baseline and follow-up

	STarT Back Trial				IMPaCT Back Study			
Risk Group	**n**	**Baseline mean (SD)**	**n**	**Follow-up mean (SD)**	**n**	**Baseline mean (SD)**	**n**	**Follow-up mean (SD**
Low	161	2.66 (2.30)	92	1.38 (2.03)	222	2.83 (2.40)	113	1.31 (1.74)
Medium	239	5.44 (2.69)	128	2.27 (2.36)	202	5.28 (2.78)	114	2.30 (2.58)
High	120	6.49 (2.81)	78	2.39 (2.55)	106	6.34 (2.90)	52	3.42 (2.92)
Bonferroni Correction method	Diff(p-value)				Diff(p-value)			
Medium to Low risk	2.78 (*P=0.00*)		0.89 (P=0.02)		2.44 (*P=0.00*)		0.98 (*P=0.00*)	
High risk to low risk	3.83 (*P=0.00*)		1.00 (*P=0.01*)		3.50 (*P=0.00*)		2.11 (*P=0.00*)	

SIPQ Single-item presenteeism question, SD Standard deviation, Diff Difference, STarT: Subgrouping for Targeted Treatment, IMPaCT: Implementation study to improve Patient Care through Targeted treatment

Additionally, longitudinal construct validity examined using correlations between change scores in reduced productivity and change scores from selected constructs in both studies showed similar findings. The strongest association with presenteeism was observed with RMDQ scores and the SF-12 physical summary dimension (Table 7.6). However, low to moderate associations were observed between the SIPQ and the pain and HR-QOL measures.

Table 7:6 Longitudinal construct validity correlations between the SIPQ and pain, disability, psychological, general and quality of life measures

Measurement	STarT Back Trial (N^a=283)	IMPACT Back Study (N^a=262)
Pain Dimensions		
Intensity of back experience at present time (N=275;261)*	0.52 (*P* < 0.001)	0.34 (*P* < 0.001)
Psychological health		
Pain related fears (TSK) (N=277;259)*	0.27 (*P* < 0.001)	0.23 (*P* < 0.001)
HADS – Anxiety(N=279;261)*	0.22 (*P* < 0.001)	0.26 (*P* < 0.001)
HADS – Depression(N=279;261)*	0.36 (*P* < 0.001)	0.30 (*P* < 0.001)
Disability		
RMDQ –disability(N=283;262)*	0.54 (*P* < 0.001)	0.54 (*P* < 0.001)
EQ-5D Dimensions		
EuroQol EQ-5D(N=270;255)*	-0.36 (*P* < 0.001)	-0.31(*P* < 0.001)
SF-12 Dimensions		
physical component summary (SF-12 v2)(N=277;256)*	-0.50 (*P* < 0.001)	-0.51 (*P* < 0.001)
Mental component summary (SF-12 v2) (N=277;256)*	-0.15 (*P* < 0.001)	-0.08
General health (N=253)*	-0.46 (*P* < 0.001)	-0.09
Perceived change in Back pain		
Global change @12 months(N=278 , N=262)**	-0.38 (*P* < 0.001)	-0.30 (*P* < 0.001)
Change in RMDQ (RMDQ>30%)(N=283, N=242)**	0.55 (*P* < 0.001)	0.54 (*P* < 0.001)

** Spearman's ranks correlation; *Pearson's correlation coefficient; a Patients employed at baseline and follow-up with relevant data, (N) represents numbers of participants at base-line and follow-up in the STarT Back Trial and IMPACT Back Study respectively, SIPQ Single-Item presenteeism question, RMDQ Roland Morris disability questionnaire SD Standard deviation, HAD Hospital Anxiety and depression scale, PCS Pain Catastrophising Scale, TSK Tampa Scale for Kinesiophobia STarT: Subgrouping for Targeted Treatment, IMPaCT: Implementation study to improve Patient Care through Targeted treatment N represents numbers at follow-up in the SBT and IBS studies.
The responsiveness analysis is based on a subsample of patients employed at baseline and follow-up, 283 for the SBT and 262 for the IBS dataset.

7.5.5 Responsiveness

A total of 283 patients employed both at baseline and follow-up in the SBT were included in the responsiveness analysis based on the global change in back pain variable (See Table 7.6). A total of 188 (68%) of all the employed patients reported an improvement in their back pain condition (See Table 7.7). In the IBS, 262 of the employed patients both at baseline and at follow-up were included in the responsiveness analysis (See Table 7.6). Of these, 186 reported an improvement in their back pain condition based on the global assessment question (See Table 7.7). Change in SIPQ showed statistically significant associations with measures assessing improvement in back pain (Table 7.7). The majority of patients showed a significant improvement in back pain condition at follow-up as determined by both the global response change in the back pain question and the RMDQ > 30% change score. As anticipated, the improved patients group had a significant reduction in work impairment compared with the non-improved patients group (Table 7.7).

Effect size statistics in the overall SIPQ scores were large (SBT dataset – SES = 0.83, SRM=0.76, IBS dataset - SES = 0.76, SRM=0.74). Similarly, the effect size statistics from the subgroup of patients with improvement in back pain based on SIPQ scores were large (1.01 to 1.18) in both studies compared to those for patients reporting no improvement (SRMs: 0.24-0.28) – giving anchor-based GRI statistics of 1.12 and 0.94 (SBT and IBS respectively) (See Table 7.7). Patients classified as 'not improved' demonstrated small improvements in their presenteeism scores although the SRM values show these improvements were very small.

Table 7:7 Responsiveness of the SIPQ change scores by proxy measures (given by global change in condition of back pain variable and RMDQ>30% criterion)

	STarT Back Trial			**IMPaCT Back Study**		
Back pain change	**n**	**Mean change(SD)**	**SRM**	**n**	**Mean change(SD)**	**SRM**
Global change in condition						
Improved	188	3.25 (2.76)	1.18	186	2.85 (2.81)	1.01
Not improved	90	0.71 (2.91)	0.24	76	0.76 (3.03)	0.25
RMDQ Change						
Improved	198	3.16 (2.94)	1.07	152	3.22 (2.83)	1.14
Not improved	85	0.82 (2.95)	0.28	90	0.68 (2.83)	0.24

SIPQ Single-Item presenteeism question, RMDQ Roland Morris disability questionnaire, SD Standard deviation, SRM Standardised response mean, STarT: Subgrouping for Targeted Treatment, IMPaCT: Implementation study to improve Patient Care through Targeted treatment. The responsiveness analysis is based on a subsample of patients employed at baseline and follow-up, 283 for the SBT and 262 for the IBS dataset.

Similarly, the SIPQ was able to detect changes in back pain patients according to the transitional global change questionnaire that assessed transition of patient condition from “no change” to “better”, “much better” or “completely recovered” and from “no change” to “worse” or “much worse”. As anticipated, patients reporting their back pain condition as ‘completely recovered’ and ‘much better’ had a significant reduction in work impairment when compared with patients self-reporting as ‘worse’ and ‘much worse’ (Table 7.8).

Table 7:8 Mean change in presenteeism for the single-item question among back pain patients based on the change in back pain condition 6 point severity question.

Global change in severity pain question	**Presenteeism**			
	n	**SBT (Mean, SD)**	**n**	**IBS (Mean, SD)**
Completely Recovered	36	-4.36 (3.11)	38	-3.61 (2.78)
Much better	90	-3.53 (2.86)	92	-3.23 (2.73)
Better	62	-2.19 (2.53)	56	-1.71 (2.68)
No change	70	-1.28 (2.55)	46	-1.26 (2.61)
Worse	15	1.00 (2.24)	28	-0.04 (3.64)
Much worse	5	2.20 (3.49)	2	0.50 (2.12)

a change in back pain between baseline and follow-up in the 6 point back pain change scale measure

7.6 Discussion

This chapter has reported on the first study to examine the validity and responsiveness of the single-item presenteeism question as a stand-alone measure for estimating presenteeism among low back pain sufferers. There has been an increased number of presenteeism instruments developed in the literature. This, to a large extent, has been driven by the significant costs associated with reduced productivity, hence necessitating the need for practical tools assessing the impact of illness on productivity at the workplace. The review in Chapter Four of this thesis identified a diverse set of multi-dimensional instruments as well as standalone global single-item questions from costing literature. The validity of standalone questions used outside of standard presenteeism instruments in these costing studies, however, has not been reported among low back pain sufferers.

This study demonstrated the construct validity and responsiveness to change in worker productivity of the single-item response question (the SIPQ). The study substantiated most of the theorised constructs underpinning the SIPQ in relation to existing measures, suggesting that the presenteeism tool is measuring what it is expected to measure. When testing for convergent validity, the strongest relationships were observed between the presenteeism measure and pain, disability, and SF-12 physical component scales. Furthermore, divergent validity was established as the presenteeism question was as anticipated, not statistically significantly correlated with constructs of age, gender and education. Known-group validity of the single-item presenteeism tool was also established, as the tool was able to differentiate between presenteeism changes among subgroups of patients based on patient risk groups (low risk, medium risk, and high risk), and observed patient response to the intervention (improvement or not).

The responsiveness to change in worker productivity of the single-item response question was also confirmed. The SIPQ was found to be highly responsive, showing ability to discriminate between subgroups of patients and changes in their health conditions according to pre-defined anchors (patient rated global change and clinical end point: RMDQ > 30% indicators of change). Estimates of effect size indicators from the overall and pre-defined groups of patients were large showing high levels of responsiveness.

The results in this study generally agreed with previous validation studies that have found strong relationships between work productivity, disease severity, pain and quality of life in a range of conditions (Reilly et al., 2008; Giovannetti et al., 2009; Zhang et al., 2010a). On the other hand, the findings regarding the association between presenteeism and psychological measures were inconclusive. Chronic back pain is known to be commonly associated with the impaired ability to perform physical roles (Lin et al., 2011), but there is less clear evidence, generally, on the influence of psychological factors such as fear avoidance and anxiety. Some studies have found that the effect of depression or anxiety on reduced productivity could not be ascertained (Hagberg et al., 2002) and others report that fear avoidance beliefs about work and physical activity due to back pain are significantly associated with both work absence and reduced productivity (Mannion et al., 2009). As hypothesised, the results presented in this chapter show significant but weak associations between the SIPQ and psychological assessment measures. Further, the results in this investigation agreed with a previous study that investigated the validity of a single-item question against a multi-item work ability index questionnaire used to assess work ability among women on long-term sick leave (Ahlstrom et al., 2010). There was a strong association between both measures, and the predictive value of health-related quality of life and sick leave was similar in both tools.

In the study presented here, the strongest associations are observed between the SIPQ and the physical summary component of the SF-12; the ‘usual activities’, and the ‘pain’ components of the EQ-5D; and in disease specific measures related to disability and pain. Similar findings were observed in the longitudinal construct validity analysis. Although there is currently no gold standard for the assessment of statistically significant meaningful changes (Pengel et al., 2004), responsiveness of the measure was shown based on two proxy measures assessing improvement in back pain condition. These approaches have been used in prior psychometric studies (Beurskens et al., 1996; Beaton and Kennedy, 2005; Coelho et al., 2008). All statistical measures applied showed responsiveness of the SIPQ.

The study has both strengths and limitations. The major strength of the study is that the analysed data were from two large empirical datasets which contained other measures alongside which to validate the SIPQ, and thus enabled comparative analysis across the two studies. Furthermore, the data have been analysed using a number of well-established statistical methods for measuring validity and responsiveness. The analysis was not, however, able to cross-validate the SIPQ against multi-item presenteeism tools. This would be an avenue for further research and could be achieved by comparison of the SIPQ with other commonly used multi-item instruments such as the Work Limitations Questionnaire, the Work Productivity and Activity Impairment questionnaire and the Stanford Presenteeism Scale. Although some studies have compared across presenteeism measures (Ozminkowski et al., 2004; Tang et al., 2009; Beaton et al., 2010; Zhang et al., 2010b; Louise et al., 2012), none have, as yet, focussed on the comparison of these measures with a SIPQ. One study, however, has compared the Work Ability Index against a single-item question assessing work ability among women on long-term sick leave (Ahlstrom et al., 2010). Such studies would be important in providing greater insight into the comparability of productivity loss estimates from single and multi-item presenteeism tools. Moreover, the results presented here only focus on low back pain sufferers, and therefore cannot be generalised.

The significance of back pain related costs of absenteeism and presenteeism augments the importance of initiatives towards the further development of presenteeism measurement and valuation instruments. The SIPQ in particular potentially offers various benefits for those wishing to investigate presenteeism given the brevity, simplicity, and efficiency of the measure and its potential applicability to diverse occupations since it does not ask job-specific questions as in the majority of the current multi-item measures. Moreover, the SIPQ can be incorporated into clinical study questionnaires without imposing a significant burden on both the respondent and researcher. This is in contrast to many of the self-report multi-item presenteeism questions used for economic evaluation that are often less practical and quite lengthy.

This study has shown the validity and responsiveness to change of the single-item presenteeism question among back pain patients, which warrants further research in other disease areas and translation into practice. Nevertheless, there is still the need for further studies undertaking concurrent validation of the SIPQ alongside other commonly used multi-item instruments such as the Work Limitations Questionnaire and Stanford Presenteeism Scale. A previous study compared psychometric properties of four presenteeism measures in workers with shoulder or elbow disorders, reporting the existence of theoretical differences in what the tools measure (Tang et al., 2009). Further research is therefore needed to provide more insight into the comparability of productivity loss estimates from single and multi-item presenteeism tools.

7.7 Conclusion

The findings presented in this chapter show the validity and responsiveness of the single-item presenteeism question. Standardised validated presenteeism measures are necessary if the impact of disease on work productivity is to be accurately quantified and incorporated within economic analysis studies. While potentially less precise and limited in obtaining more detailed information than multi-item measurement scales, the SIPQ demonstrates good construct validity and responsiveness, and offers noticeable benefits within musculoskeletal research across diverse populations. It is also able to discriminate between specific patient groups, is responsive to change, and is feasible to use in the context of back pain. Construct validity and responsiveness are, however, best confirmed through evidence from various studies. The SIPQ should therefore be validated in other studies to investigate if the results obtained here will be replicated. Subsequently the SIPQ scores from both studies will be used to generate presenteeism costs. In the next chapter, these are included as part of the cost-effectiveness analyses from a societal perspective.

CHAPTER EIGHT : COST-UTILITY ANALYSES OF STRATIFIED MANAGEMENT CARE INTERVENTIONS TARGETED AT LBP PATIENTS: THE SBT AND IBS EVALUATIONS

8.1 Introduction

In Chapters Six and Seven, estimates of the length of a friction period relevant for the United Kingdom (UK), and an assessment of the construct validity and responsiveness of a single-item presenteeism question were presented and discussed. In this chapter, economic evaluations of the stratified care intervention conducted alongside the STarT Back Trial (SBT) (Hill et al., 2011) and IMPaCT Back Study (IBS) (Foster et al., 2010) are presented, with the aim of illustrating the application of the friction cost approach in economic evaluation studies. This includes how the friction cost approach can be used to value absenteeism and presenteeism, using multiple friction period estimates and adjusting costs for the effects of productivity loss on team productivity.

The burden of low back pain has been clearly established in Chapter Five. Stratified primary care management, which employs screening tools to allocate patients to risk-defined groups, has been proposed as a means of ensuring optimal treatment of LBP (Hay et al., 2008; Hill et al., 2011). The cost-effectiveness of such interventions from a societal perspective however remains unclear. Information on societal costs is useful to support broader decision making and to prevent potentially inappropriate policy decisions resulting from narrower healthcare perspectives.

Few studies have attempted to investigate the extent to which inclusion and exclusion of productivity costs affects the true total societal costs in this area. A cost-utility analysis was previously performed alongside the STarT Back clinical trial reported by Hill et al., (2011) and further analysis was performed within risk-defined subgroup analysis and reported by Whitehurst et al., (2012). The studies showed that, from a healthcare perspective, the stratified management approach was a highly cost-effective intervention among LBP patients in general, and across all risk-defined subgroups (Hill et al., 2011; Whitehurst et al., 2012). The cost-utility analysis for the IMPaCT Back study has not yet been reported. However, the limitation of these economic analysis studies is that they were performed from the conventional UK healthcare perspective, although there are strong arguments in favour of adopting a broader societal perspective (Drummond et al., 2005; Drummond and Rutten, 2008).

The economic analyses studies in this chapter were therefore performed using a societal perspective to investigate the impact of including productivity costs estimated using a FCA on economic evaluation results, and focusing on aspects of lost productivity such as absenteeism, presenteeism, and multiplier effects (related to effects of productivity loss on team productivity) in economic evaluation.

The aims of this chapter therefore are to:

(i) assess the impact of including productivity costs of absenteeism, presenteeism and multiplier effects in cost-effectiveness analyses (investigated through two cost-utility studies comparing the stratified management intervention with non-stratified current practice overall and within risk-defined subgroups);

(ii) assess the impact of using more detailed length of friction period values when valuing productivity costs using the friction cost approach on cost-effectiveness results (assessed through friction period estimates stratified by occupational levels obtained from four different UK data sources).

Productivity costs are generated based on various cost valuation models of the friction cost approach. The chapter also illustrates practical applications of the findings from empirical studies investigating detailed friction period estimates for the UK and the validated SIPQ tool described in Chapters Six and Seven within an economic evaluation setting. Subsections of this chapter describe the approaches used to generate productivity costs, and the methods, findings, and conclusions from the economic analysis studies.

8.2 Approaches to productivity assessment

8.2.1 Work-related outcomes

Data on absenteeism and presenteeism were collected among patients in paid employment at baseline and follow-up (6 months for IBS, and 12 months for SBT). Participants were asked about their work status (full time/part-time/not working) and their current/previous occupation. The occupation of each participant was then classified according to the SOC

2000 classifications and wages assigned, based on the Annual Survey of Hours and Earnings, 2008 as described in Chapter Six. Social security benefits such as sick pay, unemployment benefits and disability pay were excluded from the evaluation as these were considered to be transfer payments, rather than lost productivity as recommended for societal perspective evaluations (Brouwer et al., 2001; van den Hout, 2010).

Absenteeism was based on self-reported duration of back-pain related work absence, by asking respondents whether they had taken time off work as a result of their back pain problem and if so, the number of days off work during the study period. Presenteeism was assessed by asking participants in paid employment the extent to which back pain had affected their performance at work during the past month using the SIPQ validated in Chapter Seven of this thesis. In the base-case analysis, the impact of including the costs of lost productivity (i.e. absenteeism, presenteeism, and multiplier effects) was investigated in the economic analysis.

8.2.2 Valuation of productivity costs

In the base-case analysis, the valuation of lost productivity was carried out using the friction cost approach. Detailed methods for the models used in estimating productivity costs using the friction cost approach are described below. For comparison purposes, productivity costs were also recalculated using the more conventional approach of using the friction cost approach to value absenteeism and the human capital method to value presenteeism as part of the sensitivity analysis.

Wage-related multiplier estimates were used to account for potential differences in job characteristics of team work, substitution, and time sensitivity of output across different job classifications as a result of productivity loss. In order to do this, the occupational wage was adjusted using an occupational weighted multiplier. The multiplier values used were obtained from a published study conducted in the UK among early RA patients enrolled in the Early Rheumatoid Arthritis Network (ERAN) cohort (Zhang et al., 2012).

The multiplier is defined as, "*the cost of an absence as a proportion (often greater than one) of the absent worker's daily wage*", (Nicholson et al., 2006) (p 112). The study by Zhang et al., (2012) generated multipliers for selected professions in the UK that were used for these economic evaluations. The occupational classes from the IBS and SBT were mapped as closely as possible to those of Zhang et al., (2012) in order to obtain multipliers relevant for the evaluations in this thesis. The values ranged from 1.0 (elementary occupations) to 4.68 (managers) in absenteeism and from 1.0 to 2.78 for presenteeism (See Table 8.1).

Table 8:1 Average multiplier estimates per job category for Absenteeism and presenteeism

Job Category	Absenteeism		Presenteeism	
	n	Multiplier estimate	N	Multiplier estimate
Manager	5	4.68	15	2.78
Professionals	7	1.97	8	1.40
Technicians	7	1.78	10	1.55
Clerk	6	1.27	7	1.25
Services and sales	5	1.07	12	1.08
Craft	1	1.63	2	1.31
Operators	2	2.00	2	2.13
Elementary occupations	1	1.00	1	1.00

Source: Adopted from Zhang et al., (2012)

8.2.2.1 Valuation of Absenteeism

The questions on work absence used in the study data sets involved asking respondents whether they had experienced absence from work during the study period (6 months for IBS and 12 months for SBT).

Absenteeism based on the Human capital method

For comparison purposes with the friction cost approach, absenteeism costs were first calculated using the human capital method by multiplying the work absence time due to back pain by the gross national wage-rate for each occupation stratified by age and gender and excluding adjustments for multiplier effects based on the model below:

Equation 8.1

$$PC = t * w * h$$

- Here, *PC* represents absenteeism related costs valued using the human capital method.
- *t* represents self-reported time off work (in days).
- *w* represents the gross national wage-rate per hour stratified by age and gender
- *h* represents the number of hours worked per day.

Absenteeism using the friction cost approach including multiplier effects

In the base-case analysis, productivity costs were valued, based on the friction cost approach for different cost scenarios, which are described in the next section. Additionally, productivity costs were adjusted for the impact of work absence on team productivity (multiplier effects). The cost scenarios generated in this chapter are based on the friction period data obtained from DLA Piper UK (DLA), the Chartered Institute of Personnel and Development (CIPD), the Office of National Statistics (ONS) and the University of Birmingham (UoB), as described and reported in Chapter Six of this thesis. The results are therefore presented for five cost scenarios generated by the friction cost approach: i) scenario 1 applies a single friction period ii) scenario 2 applies a two occupational level friction period from the DLA dataset iii) scenarios three, four and five apply five occupational level friction periods from the CIPD, ONS and UoB datasets.

In cost scenario 1, productivity costs were estimated using the average friction period (found to be 2.7 months for the UK in Chapter Seven) while incorporating multiplier effects. The estimation of labour costs was based on age-adjusted and sex-adjusted wage-rates. Time off work due to back pain during the study period was multiplied by the labour wage-rate and an elasticity value of 0.8. The elasticity value was to take into account work compensations in the short-term as recommended by the Dutch costing manual (Hakkaart-van Roijen L et al., 2010), and based on the employee's length of absence. The length of absence was categorised into short-term absence (less than the friction period) and long-term absence (greater than the friction period). The elasticity value was then applied to the short-term absences. The friction costs for this period were therefore assumed to be 80% of wage costs.

The elasticity value is often used as an adjustment factor to account for compensation at the work place, for instance in situations where labour reserves exist within an organisation or where the individual makes up for lost work time on their return to work (Koopmanschap et al., 1995; Hanly et al., 2012). In the literature, and as demonstrated in Chapter Three of this thesis, a value of 0.8 has often been used to represent 80% productivity loss during the friction period. The productivity cost calculation is illustrated in Equation 8.2:

Equation 8.2

$$PC = t * w * h * l$$

- Here*,* *PC* represents absenteeism related costs valued using the friction cost approach.
- *t* represents the individual's self-reported time off work (days) (based on a uniform friction period of 2.7 months)
- *w* represents the gross national wage-rate per hour stratified by age and gender for an occupation adjusted for multiplier effects.
- *h* represents the number of hours worked per day.
- *l* represents the elasticity value

A variant of this cost model, comprising scenarios 2 to 5, involved using friction periods stratified by occupational classification from the four data sets as described in Chapter Six of this thesis.

Absenteeism – without accounting for multiplier effects

In the regular friction cost approach, absenteeism was estimated as commonly done using the same equation as above but without taking into account any multiplier effects. This was then used as part of the sensitivity analysis.

8.2.2.2 Valuation of Presenteeism

Reduced productivity was measured using the SIPQ. Respondents were asked to rate how much back pain had affected their overall work performance over a period of 30 days. The self-reported reduced productivity score was then translated into productivity loss by multiplying the self-reported presenteeism score by the actual number of days worked for each individual according to Equation 8.3 below:

Equation 8.3

$$R = d * (r / 10)$$

- Here, *R* represents productivity loss in relation to presenteeism over the past 30 days
- *r* represents the self-estimated level of back pain related reduced productivity during the past 30 days obtained at follow-up.
- *d* represents the actual days at work for each individual - estimated as overall days worked less absence from work days due to sickness.

Productivity losses resulting from reduced productivity at work were then converted into monetary estimates using the friction cost approach and the human capital method.

Presenteeism valuation: Using the friction cost approach including multiplier effects

In the friction cost approach, incorporated in scenarios 1 to 5, presenteeism related productivity loss was calculated by multiplying reduced productivity loss (in days) by the gross daily wage-rate and the elasticity value of 0.8 adjusted for potential internal compensations and the values extrapolated over the study periods. The details are shown in equation 8.4:

Equation 8.4

$$C_h = R_h * w * l$$

- Here, C_h represents presenteeism costs due to back pain over the past 30 days.
- R_h represents back pain related productivity loss while working (including number of days and hours per day)
- w represents the gross national wage-rate per day stratified by age and gender for an occupation adjusted for multiplier effects.
- l represents the elasticity value

Presenteeism valuation: Using the human capital method and excluding multiplier effects

In the conventional human capital method, approach 2, reduced productivity was calculated using the human capital method and excluding multiplier effects as commonly done by multiplying the estimated productivity loss in days by the age-sex dependent average wage (Equation 8.5). This was used as part of the sensitivity analysis.

Equation 8.5

$$C_h = R_h * w$$

- Here, C_h represents presenteeism costs due to back pain over the past 30 days.
- R represents back pain related productivity loss while working.
- w represents the UK national wage rate.

8.3 Methods 1: STarT Back Study

8.3.1 Study design

A cost-effectiveness analysis was performed alongside a pragmatic multi-centre randomised controlled trial carried out over a period of 12 months, including 851 participants (568 assigned to the stratified primary care management intervention and 283 to the non-stratified current best practice). Details of the trial were given in Chapter Five of this thesis. The economic analysis was undertaken from a broader societal perspective, including both healthcare and productivity costs.

8.3.2 Outcomes for economic analysis

Health outcome and work-related data were collected through self-reported questionnaires. Outcome data were collected at baseline, 4 months and 12 months. Work-related data including time off work (absenteeism) and presenteeism were collected at baseline and 12 months.

8.3.3 Health outcomes

The pre-specified outcomes for the economic analysis were change in physical factors for chronicity as measured by the Roland Morris Disability Questionnaire (RMDQ) and change in overall health status on the EQ-5D measure (Dolan, 1997). The EQ-5D responses were then used to generate Quality Adjusted Life Years (QALYs) using area-under-the-curve analysis across the study duration (Manca et al., 2005).

8.3.4 Healthcare resource use and costs

Resource use information was extracted from the study questionnaires and included consultations, other healthcare professional consultations, prescriptions, hospital based procedures and out-of-pocket expenditures. A telephone call at twelve months was used to obtain additional data for patients not responding to the questionnaire. Costs were included by multiplying resource use data with United Kingdom (UK) unit cost estimates, expressed in 2008/2009 average prices. This base year was used for comparability with the initial cost-effectiveness analyses. All costs within the healthcare sector were included regardless of whether they were borne by the government or privately by patients. Details of the relevant resource use and unit cost data are reported in Appendix C.1

8.4 Methods 2: IMPaCT Back Study

8.4.1 Study design

The IBS economic evaluation was performed alongside a prospective, population-based, quality improvement before and after study over a period of 6 months. The study included 922 participants, comprising 368 in phase 1 strategy (before intervention) and 554 in phase 3 strategy (after stratified intervention). Details of the study were given in Chapter Five. Similar to the STarT Back trial methods described above, the economic analysis conducted here was performed from a societal perspective.

8.4.2 Outcomes for economic analysis

Health outcome data were collected at baseline, 2 months and 6 months using self-reported questionnaires. Data regarding employment status, work absence (absenteeism) and presenteeism were collected at baseline and 6 months. Health resource use data were collected using a six months self-reported questionnaire.

8.4.3 Health outcomes

Similar health outcomes to the SBT design were collected over a shorter period (6 months). The EQ-5D responses were then used to generate Quality Adjusted Life Years (QALYs) (Manca et al., 2005).

8.4.4 Healthcare resource use and costs

Resource use information was extracted from a combination of medical record reviews and the six month self-report questionnaire, and included hospital stays, outpatient appointments, hospital visits, medication, over the counter treatments, health visits with the NHS and private practice. Unit costs assigned to these resources were obtained from published sources reflecting UK national averages. A telephone call at 6 months was used to obtain additional data for non-responders. For patients who provided permission for medical records to be reviewed, records were anonymously retrieved to gain further insight into self-reported absence reports reported in the 6 month questionnaire. Details of the relevant IBS resource use and unit cost data are reported in Appendix C.2

8.5 Analytic methods

Analysis was conducted based on the intention-to-treat principle. Details of other aspects including handling of missing data, costing and uncertainty analysis are included in the following sections.

8.5.1 Missing data

The multiple imputation method was used to impute missing data. The technique is used to replace each missing variable with a set number of possible values (Briggs et al., 2003). In both studies, five estimates for each missing variable were generated using a multivariate normal model (Whitehurst et al., 2012). The overall mean of the five estimated values was then used as the imputed observation included in the analysis. Imputation was carried out on both the costs and outcomes.

8.5.2 Exploring variations in cost and work outcomes between groups

Costs and outcomes were calculated for overall and risk-group analysis, and defined as means and standard deviations (SD) for costs and effects per patient per year. Bootstrap methods were applied to generate difference in mean costs and QALYs between the treatment groups (Barber and Thompson, 2000). Confidence intervals for costs were generated using the bias-corrected and accelerated bootstrapping method (1000 replications) (Barber and Thompson, 2000).

Cost valuation scenarios were presented for disaggregated costs of: i) healthcare costs with both absenteeism and presenteeism adjusted for multiplier effects ii) healthcare costs with both absenteeism and presenteeism excluding multiplier effects. Statistical analysis was performed using STATA 12 and SPSS 19.

8.5.3 Cost-Utility analysis

Incremental cost-effectiveness ratios (ICERs) were generated as the additional costs (difference in costs) divided by additional benefits (difference in QALYs gained) of the study intervention within each risk-defined subgroup and for the overall patient group.

Costs relating to absenteeism, presenteeism and multiplier effects of productivity loss on team work were included in the calculations alongside direct healthcare costs. In order to generate appropriate incremental QALY values, between-group imbalances were accounted for using linear regression analysis to adjust for baseline EQ-5D, RMDQ scores, age, gender and duration of pain at baseline. Given the time horizon of the cohort study (6 months) and randomised trial (12 months), no discounting was performed.

To account for uncertainty, a bootstrap approach based on 5000 replicates was used to generate the data used to produce cost effectiveness acceptability curves (CEACs) at different willingness-to-pay values. A comparison of costs and effects was illustrated graphically in the form of cost-effectiveness planes (CE plane) (van Hout et al., 1994; Briggs and Fenn, 1998). The horizontal axis showed the difference in the costs between the treatment arms (in this case stratified management minus usual practice), while the vertical axis showed the difference in effects.

The decision on whether the ICER obtained is acceptable often depends on the maximum amount that the decision maker is willing to pay for an additional QALY. The probability that the stratified care intervention for low back pain intervention is cost-effective at different values of the maximum acceptable ratio to society (willingness to pay values) was plotted as a CEAC (Fenwick and Byford, 2005). ICERs were generated from a societal perspective based on friction cost approach valuation models for different willingness-to-pay values. To assess the impact of using more detailed friction periods, CEACs were also generated for friction periods varying by occupation levels. In the subgroup analyses, the probability curves for the three risk groups were plotted together to allow direct comparison of the risk groups.

8.5.4 Sensitivity Analysis

A number of sensitivity analyses were performed to assess the impact of assumptions made during the analyses on the overall friction related costs. The first sensitivity analysis included using the conventional approach of incorporating productivity costs, i.e., using the friction cost approach for valuing absenteeism and the human capital method for valuing presenteeism and excluding multiplier effects. This aimed to assess possible differences between the two approaches. In the second sensitivity analysis, the elasticity factor was adjusted from 0.8 in the base case analysis to 1 in the sensitivity analysis. No published estimate of an elasticity factor currently exists for the United Kingdom. Therefore, in the absence of a published elasticity value, an assumption was made to use a value of 1 as done in a previous UK study (Hanly et al., 2012). In the final analysis, combined multiplier values from the UK study (Zhang et al., 2012) and a previous US study (Pauly et al., 2008) were applied in order to assess the sensitivity of these values on the overall impact.

8.6 Results 1: The STarT Back Trial

8.6.1 Study population

Following the multiple imputation procedure on the SBT study dataset, a total of 851 participants in the study were included in the base-case analysis; 568 in the intervention group and 283 in the control group. 567 (67%) of these patients provided cost utilisation data at 12 months, 386 in the intervention group and 181 in the control group. The resource use data are provided for the overall sample and for each risk-defined subgroup in Tables 8.2 to 8.3.

8.6.2 Work-related outcomes

The mean work-related resource per patient shows a consistent pattern of lower productivity loss in patients treated with stratified management (Table 8.2). Approximately 23% of the patients in the intervention group and 34% in the usual practice group reported having time off work at the 12 month follow-up period. The overall absence from work for the intervention group was 4.43 days compared to 12.18 days in the control group, a difference of 7.75 days (95% CI -14.2 to -1.31; p = 0.01). For the case of presenteeism, the stratified management intervention was associated with a productivity loss of 46.5 days compared to 53.2 days in the control group, a difference of 6.75 days (95% C: -20.4 to 6.89 ; p = 0.33).

In the subgroup analysis, a significant difference in lost days was observed in the medium-risk-defined subgroup (4.13 days versus 18.44 days for stratified and usual practice respectively) but not in the high-risk (9.85 days versus 10.57 days) and low-risk groups (0.37 days versus 3.0 days).

Table 8:2 Comparison of overall work-related outcomes for participants in paid employment between the two study arms

Resource/Cost Component	Cost(£)				Mean difference (95% CI)
	Stratified care (n = 200)		Usual care (n = 98)		
Overall Analysis	n/mean	%/SD	n/mean	%/SD	
Work-related outcomes – HCM					
Time off work due to low back pain: (n, %)	46	23%	33	34%	
Absenteeism days	4.4	21.2	12.2	35.1	-7.75 (-14.2 to -1.3)
Back-pain related absenteeism cost	447	2126.2	1113	3424.4	-69 (-1362.1 to 25.2)
Presenteeism days	46.5	54.4	53.2	59.7	-6.75 (-20.4 to 6.89)
Back-pain related presenteeism cost	3524	4480.2	4274	5222.5	-750 (-1924.93 to 424.60)
Work-related outcomes – FCA					
Back-pain related absenteeism cost					
Scenario 1	461	1760.4	1179	3598.4	-717 (-1334.4,-100.6)
Scenario 2	457	1733.8	1275	4119.0	-818 (-1493.4,-141.8)
Scenario 3	407	1460.5	1171	3664.6	-763 (-1356.4,-170.2)
Scenario 4	461	1760.4	1272	4101.3	-811 (-1486.0,-134.3)
Scenario 5	433	1586.2	1082	3156.6	-649 (-1194.8,-103.2)
Back-pain related presenteeism cost	4411	5510.1	5618	6896.4	-991 (-2344.9, 363.6)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

*The estimation of productivity loss focussed on the subsample of respondents in paid employment at 12 months for SBT (298/567).

8.6.3 Productivity Costs: The human capital method

The findings in Table 8.2 show that the overall mean costs based on the human capital method were lower in the stratified management group (intervention) than in the current practice group (control). The overall mean absenteeism costs were £447 in the stratified management group and £1113 in the current practice group, a difference of £669. For presenteeism, the corresponding productivity loss was £3524 and £4274 for the intervention and control groups respectively, a difference of £750 (See Table 8.2).

8.6.4 Productivity costs: Methodological applications of the friction cost approach

In the overall analyses, the results in Table 8.2 indicate that adjusting productivity costs for multiplier effects leads to estimates of productivity costs that are higher than the more traditional human capital method. The observed mean costs in the stratified care intervention were between £407 and £461 depending on the occupational friction period level used compared with over £1000 in the current practice groups (Table 8.2).

In the subgroup analysis, large and significant cost differences between stratified care and usual care were observed in the medium-risk group compared to other risk-groups. Moreover, mean productivity losses were higher in the medium-risk and low-risk categories for the stratified management care intervention. However, in the high-risk category, the control was associated with lower mean productivity costs compared to the stratified management care intervention (Table 8.3). As shown in Table 8.3, the results of the overall analysis were replicated in the subgroup analysis. In addition, there were no significant differences between the friction cost scenarios at the different occupational level categories.

Table 8:3 Comparison of risk-group work-related outcomes for participants in paid employment between the two study groups

Resource/Cost Component	Cost(£)				Mean difference (95% CI)
	Stratified care (n=65)		Usual care (n=29)		
Low-risk	n/mean	%/SD	mean	SD	
Time off work due to low back pain: (n, %)	8	12.3%	5	17.2%	
Back-pain related absenteeism cost					
Scenario 1	47	160.1	199	836.3	-152 (-368.5,63.9)
Scenario 2	47	160.1	199	836.3	-152 (-368.5,63.9)
Scenario 3	47	160.1	219	945.0	-172 (-414.9,69.7)
Scenario 4	47	160.1	199	836.3	-152 (-368.5,63.9)
Scenario 5	47	160.1	199	836.3	-152 (-368.5,63.9)
Back-pain related presenteeism cost	3123	4235.7	5270	8367.1	-2146 (-4785.6,492.0)
Medium-risk	(n = 81)		(n = 48)		
Time off work due to low back pain (n, %)	22	27.1%	20	41.7%	
Back-pain related absenteeism cost					
Scenario 1	422	1741.4	1870	4725.9	-1448 (-2604.1,-292.6)
Scenario 2	401	1607.4	2104	5660.3	-1703 (-3034.8,-371.5)
Scenario 3	336	1166.4	1878	5002.4	-1541 (-2695.2,-388.7)
Scenario 4	406	1647.1	2098	5637.6	-1691 (-3022.3, -361.7)
Scenario 5	369	1380.0	1709	4267.3	-1340 (-2361.4,-319.3)
Back-pain related presenteeism cost	4845	5329.8	5902	6322.5	-105 (-3130.2,1055.7)
High-risk	(n = 54)		(n = 21)		
Time off work due to low back pain: (n, %)	16	30%	8	38%	
Back-pain related absenteeism cost					
Scenario 1	1061	2656.4	866	1558.9	194 (-1043.6,1433.0)
Scenario 2	1053	2628.4	866	1558.9	187 (-1039.5,1414.3)
Scenario 3	966	358.2	869	1570.5	96 (-1021.8,1214.0)
Scenario 4	1061	2656.3	866	1558.9	194 (-1043.6,1433.0)
Scenario 5	1010	2477.4	866	1558.9	144 (-1021.7,1308.9)
Back-pain related presenteeism cost	5369	886.2	5404	6339.7	-35 (-3603.2,3531.8)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

The estimation of productivity loss focussed on the subsample of respondents in paid employment at 12 months for SBT (298/567).

8.6.5 SBT Friction cost approach: Impact of absenteeism, presenteeism and multiplier effects.

The friction cost approach has often been associated with lower societal costs compared to the human capital method. In this analysis, an investigation was carried out into the potential impact on total societal costs of using the friction cost approach for valuing both absenteeism and presenteeism, while including effects of productivity loss on team productivity. This was compared to the conventional approach of using the friction cost approach for absenteeism and the human capital method for presenteeism. The results in Table 8.4 show that applying the friction cost approach for both absenteeism and presenteeism leads to higher societal costs than the conventional approach.

Table 8:4 Back pain mean (SD) societal costs per patient and year for patients reporting resource utilisation data

	Stratified care (n=568)		Usual care (n=283)		Difference (95% CI BC)
	Mean	**SD**	**Mean**	**SD**	
Absence days	4.6	21.2	12.18	35.1	
Health-care costs (£)	240	261.3	274	278.2	-34 (-76.87; 3.3)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1881	4152.3	2568	5899.5	-687 (-1373.4, -1.4)
Scenario 2	1879	4147.9	2601	6066.2	-722 (-1417.5,-26.6)
Scenario 3	1862	4105.8	2565	5926.6	-703 (-1386.8,-19.3)
Scenario 4	1881	4152.3	2600	6025.5	-719 (-1413.0,-26.3)
Scenario 5	1871	4124.8	2535	5765.3	-663 (-1339.5,12.4)
Total back pain related costs using FCA approach for absenteeism and HCM presenteeism					
Scenario 1	1613	3402.7	2090	4500.0	-477 (-1020.0,65.8)
Scenario 2	1612	3403.5	2104	4572.4	-491 (-1038.0,55.6)
Scenario 3	1599	3367.7	2087	4509.4	-488 (-1028.7,52.5)
Scenario 4	1613	3402.7	2107	4550.4	-494 (-1040.2, 51.4)
Scenario 5	1606	3382.3	2071	4437.2	-465 (-1003.1, 72.4)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

8.6.6 SBT Friction cost approach: Impact of absenteeism, presenteeism and multiplier effects in subgroup analysis.

As shown in Tables 8.5 to 8.7, the costs when applying the friction cost approach and incorporating multiplier effects were generally higher than costs generated using the more conventional friction cost approach among risk-group populations. Again, the main cost driver for this difference was the effect of including multiplier estimates. Moreover, there were no significant differences between the total societal costs for the occupational level friction period cost scenarios. Overall, similar results to the overall analysis were observed in the subgroup analysis (Tables 8.5 to 8.7).

Table 8:5 Back- pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: Low-risk

	Stratified care (n=148)		Usual care (n=73)		Difference (95% CI BC)
	Mean	**SD**	**Mean**	**SD**	
Absence days	0.37	1.2	3.0	11.9	
Health-care costs (£)	96.1	166.9	160.4	153.8	-64 (-110.2; -18.4)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1445	3188.6	2188	6006.6	-743 (-1961.7,475.2)
Scenario 2	1445	3188.6	2188	6006.6	-743 (-1961.7,475.2)
Scenario 3	1445	3188.6	2196	6050.9	-751 (-1975.4,472.0)
Scenario 4	1445	3188.6	2188	6006.6	-743 (-1961.7,475.2)
Scenario 5	1445	3188.6	2188	6006.6	-743 (-1961.7,475.2)
Total back pain related costs using FCA approach for absenteeism and HCM for presenteeism					
Scenario 1	1169	2477.1	1632	4303.1	-467 (-1763.4; 422.8)
Scenario 2	1169	2477.1	1632	4303.1	-467 (-1763.4; 422.8)
Scenario 3	1169	2477.1	1644	4342.9	-475 (-1691.1; 374.9)
Scenario 4	1169	2477.1	1636	4303.1	-467 (-1650.0 ; 442.7)
Scenario 5	1169	2477.1	1636	4303.1	-467 (-1792.5; 443.3)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

Table 8:6 Back- pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: Medium-risk

	Stratified care (n=263)		Usual care (n=131)		Difference (95% CI
	Mean	**SD**	**Mean**	**SD**	
Absence days	4.13	15.94	18.44	47.2	
Health-care costs (£)	235	229.6	288	283.8	-53 (-105.0;-.4)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1819	3977.8	3136	6451.3	-1317 (-2355.0 ,278.9)
Scenario 2	1812	3957.4	3221	6839.8	-1409 (-2480.7,-337.4)
Scenario 3	1792	3896.3	3138	6553.9	-1345 (-2384.3,-307.6)
Scenario 4	1814	3963.4	3219	6761.6	-1405 (-2470.3,-340.3)
Scenario 5	1802	3924.6	3077	6263.3	-1274 (-2289.0,-259.2)
Total back pain related costs using FCA approach for absenteeism and HCM for presenteeism					
Scenario 1	1595	3442.6	2579	4952.4	-983 (-1967.0;-142.7)
Scenario 2	1592	3437.7	2616	5105.9	-1024 (-2024.8; -84.8)
Scenario 3	1579	3393.9	2567	4951.4	-988 (-1914.9;-105.8)
Scenario 4	1595	3442.6	2616	5046.0	-1021 (-2132.3;-116.6)
Scenario 5	1587	3414.7	2537	4830.3	-950 (-2001.7;-68.8)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

Table 8:7 Back- pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: High-risk category

	Stratified care (n=157)		Usual care (n=79)		Difference (95% CI)
	Mean	**SD**	**Mean**	**SD**	
Absence days	9.85	35.4	10.57	18.2	
Health-care perspective (£)	384	304.6	358	323.5	26 (-53.9; 121.0)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	2404	5122.5	1956	4531.8	448 (-892.6,1788.9)
Scenario 2	2401	5117.8	1956	4531.8	445 (-894.0,1785.6)
Scenario 3	2373	5075.1	1956	4532.9	416 (-915.4,1748.4)
Scenario 4	2404	5122.5	1956	4531.8	448 (-892.5,1788.9)
Scenario 5	2387	5093.4	1956	4531.8	431 (-903.7,1766.8)
Total back pain related costs using FCA approach for absenteeism and HCM for presenteeism					
Scenario 1	2063	4000.7	1701	3810.4	362 (-829.8; 1309.5)
Scenario 2	2068	4011.0	1693	3773.2	375 (-736.0; 1356.3)
Scenario 3	2039	3965.3	1702	3811.2	338 (-776.7; 1347.1)
Scenario 4	2063	4000.7	1701	3810.4	362 (-808.6; 1313.7)
Scenario 5	2051	3979.0	1701	3810.5	349 (-802.7; 1236.3)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB.

8.6.7 Cost-utility analysis

From the healthcare perspective, the stratified management intervention was associated with greater mean health benefit at lower mean healthcare costs (0.039 additional QALYs and a mean healthcare cost difference of -£34.39). Similarly, to the healthcare perspective, the stratified management intervention was cost-effective from a societal perspective in the overall analysis and in risk-defined groups (Table 8.8). For the overall analysis, a dominant position was observed for stratified management care as the intervention was associated with increased health benefits achieved at a lower cost.

In the subgroup analysis, stratified care had the highest probability of being cost-effective among medium-risk patients for willingness-to-pay (WTP) thresholds of £20,000 and below. However, at WTP thresholds greater than £20,000, stratified care was highly cost-effective among high-risk patients. More specifically, the observed mean societal incremental cost-effectiveness ratio (ICER) in the high-risk group ranged from £7,298 to £7,860 per QALY gained in the various cost scenarios, well below the accepted £30,000 threshold (McCabe et al., 2008).

Among low-risk category patients, the intervention was marginally less effective (-0.001 QALYs) and less costly (range:-£751 to -£743). The observed ICER was over £700,000 saved per QALY lost for stratified care. Although the QALY difference was not statistically significant, these results suggest that the cost saving is worth making. For patients in the medium-risk group, a dominant position of stratified care was observed in all cost scenarios as there were observed health benefits achieved at lower costs
(range -£1,409 to -£1,274) (Table 8.8).

Table 8:8 Incremental societal costs and incremental QALYs: Point estimates for the SBT base-case analysis

	P1	P2	P3	P4	P5
Societal Care perspective: Overall					
Mean cost difference per patient (£)	-687	-722	-703	-719	-663
Mean QALY difference			0·039		
ICER: Societal Cost per QALY gained (£)	Dominant	Dominant	Dominant	Dominant	Dominant
Low-risk category					
Mean cost difference per patient (£)	-743	-743	-751	-743	-743
Mean QALY difference			-0.001		
ICER: Societal Cost per QALY gained (£)	743,000	743,000	743,000	743,000	743,000
Medium-risk category					
Mean cost difference per patient (£)	-1,317	-1,409	-1,345	-1,405	-1,274
Mean QALY difference			0.044		
ICER: Cost per QALY gained (£)	Dominant	Dominant	Dominant	Dominant	Dominant
High-risk category					
Mean cost difference per patient (£)	448	445	416	448	431
Mean QALY difference			0.057		
ICER: Cost per QALY gained (£)	7,860	7,807	7,298	7,860	7,561

P1, single-friction period, P2 – two level friction period DLA, P3 five-level friction period CIPD , P4 five-level friction period ONS, P5 five-level friction period UoB.

8.6.8 Cost-effectiveness planes – overall healthcare and societal cost perspective

The Cost-effectiveness planes showing the scatter plot of incremental costs versus incremental QALYs for the overall analyses from the healthcare and societal perspectives analyses are presented in Figures 8.1 and 8.2 respectively. The location of the majority of bootstrapped cost-utility pairs was used to identify the nature of the uncertainty in the incremental costs and QALY estimates

From a healthcare perspective, the dominance of the stratified management intervention was observed with 96% of cost-utility pairs located in the south-east quadrant. The societal perspective strengthened the case for the intervention with stratified management care remaining dominant (98% average of the cost-effect pairs located in the southeast quadrant) in all societal cost scenarios (Figure 8.1).

8.6.8.1 Cost-utility planes for the overall analysis

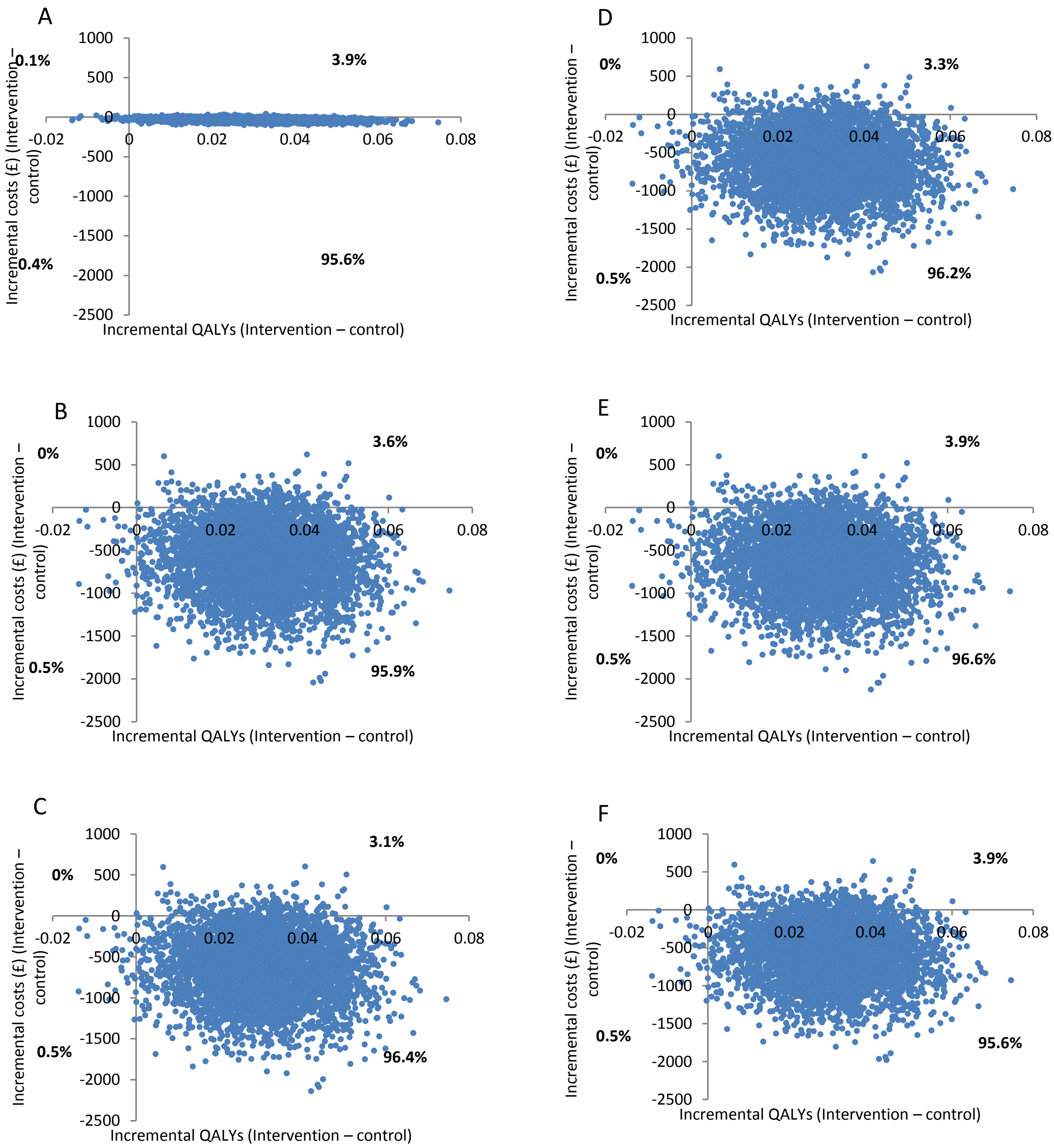


Figure 8:1 Cost-utility plane for the overall comparison of stratified management care compared to usual care for societal costs

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

8.6.9 Cost-effectiveness planes for the subgroup analysis

In the subgroup analysis, Figures 8.2 to 8.4 show the scatter plots for the incremental costs versus incremental QALY in the different risk groups. The results indicate much uncertainty in the bootstrapped data among the low-risk patient group as demonstrated by the distribution of the cost-effect pairs in all four quadrants of the CU planes (Figures 8.2 to 8.4).

8.6.8.2 Cost-utility analysis planes for the stratified analysis

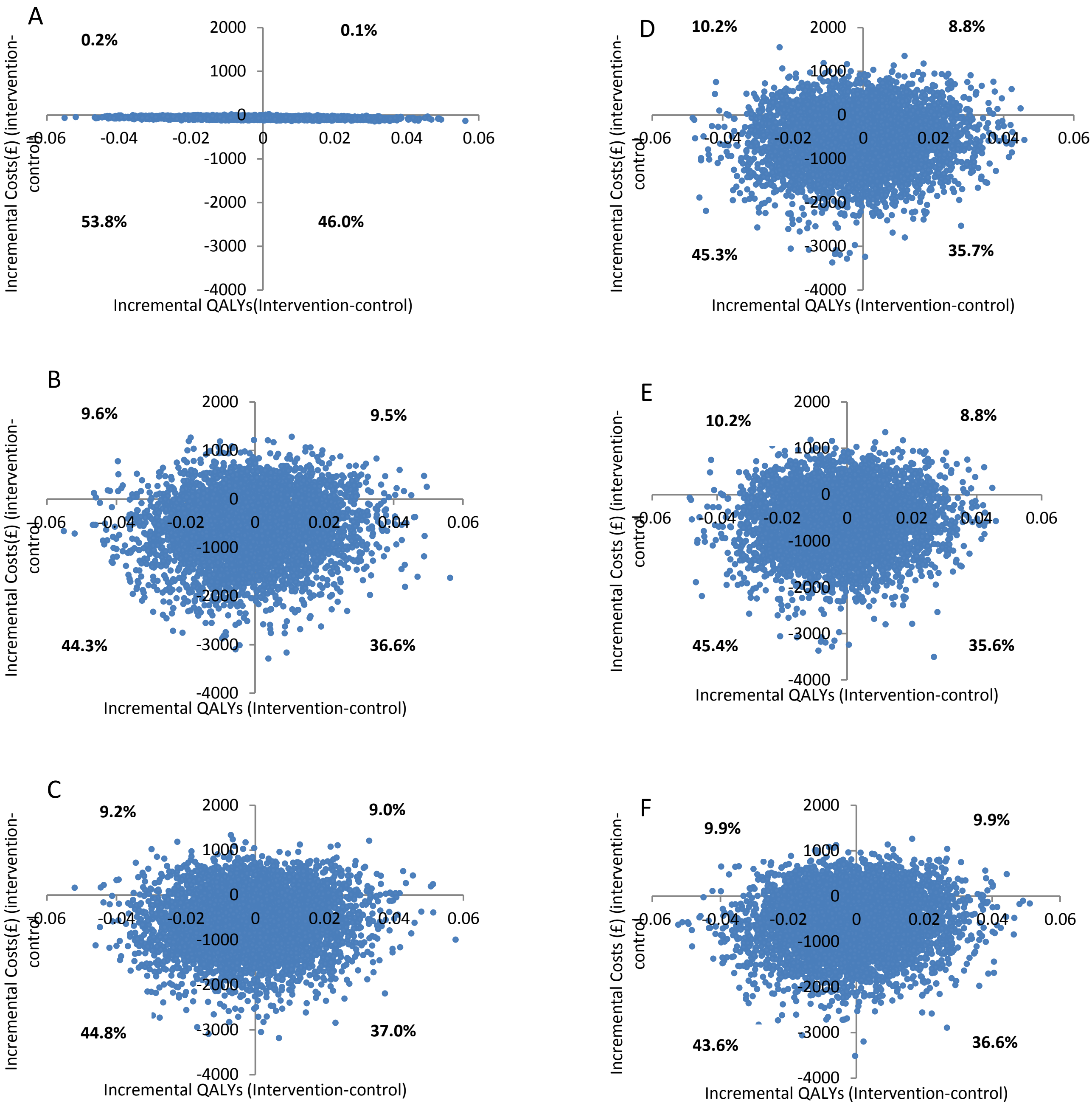


Figure 8:2 Cost-utility planes for the stratified analysis of stratified management care compared to usual care for a societal perspective – low-risk group

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

The findings in Figure 8.3 indicate that cost savings and increased benefits were observed for all scenarios among medium risk group patients (average 99% of bootstrapped CU pairs in SE quadrant). In other words, stratified care dominates usual care in 99% of the cases. Similar findings were demonstrated in the high-risk group with increased benefits and costs in all cost scenarios (average 76% of CU pairs in the NE quadrant).

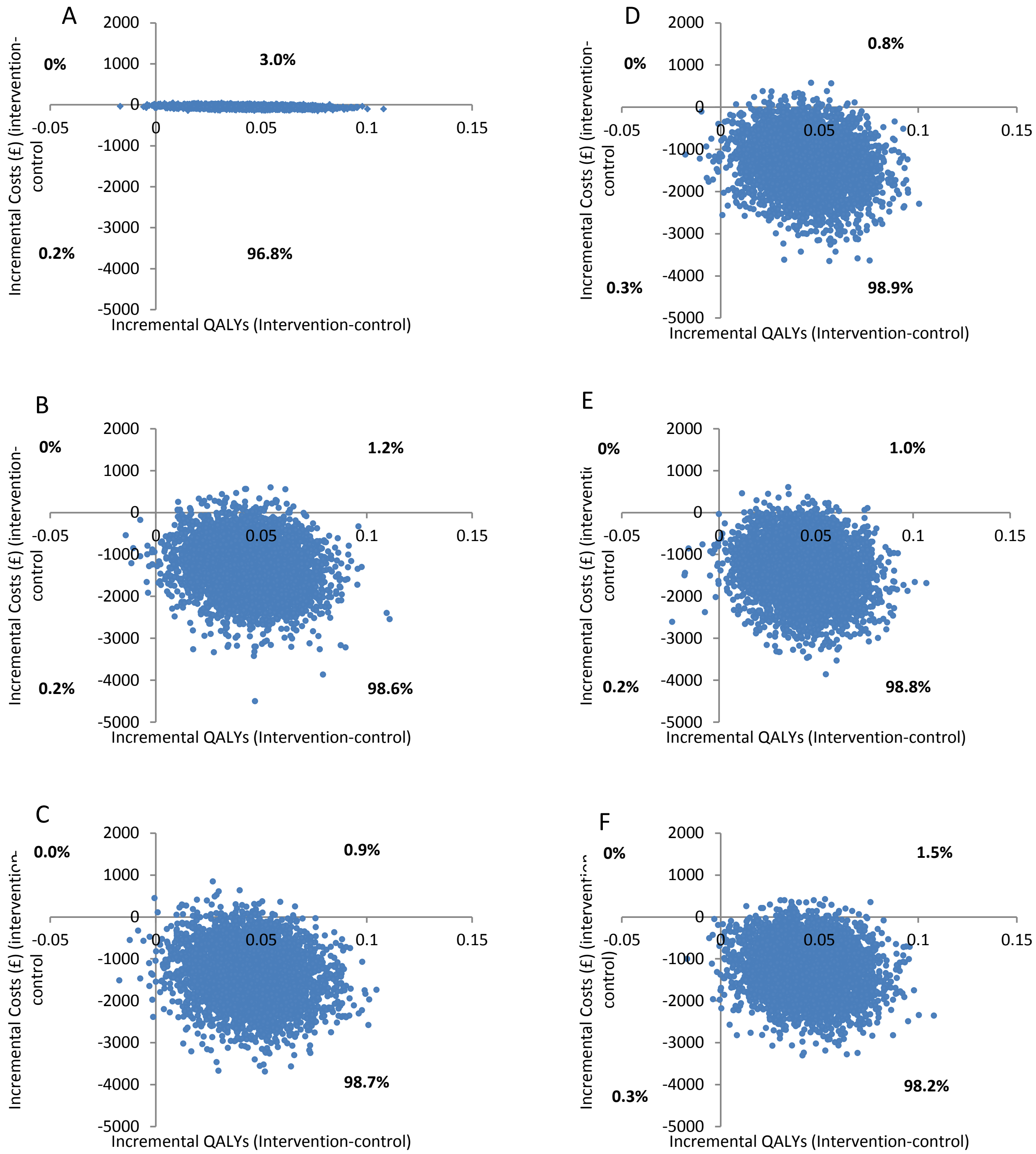


Figure 8:3 Cost-utility plane for the stratified analysis of stratified management care compared to usual care for a societal perspective: medium-risk category

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

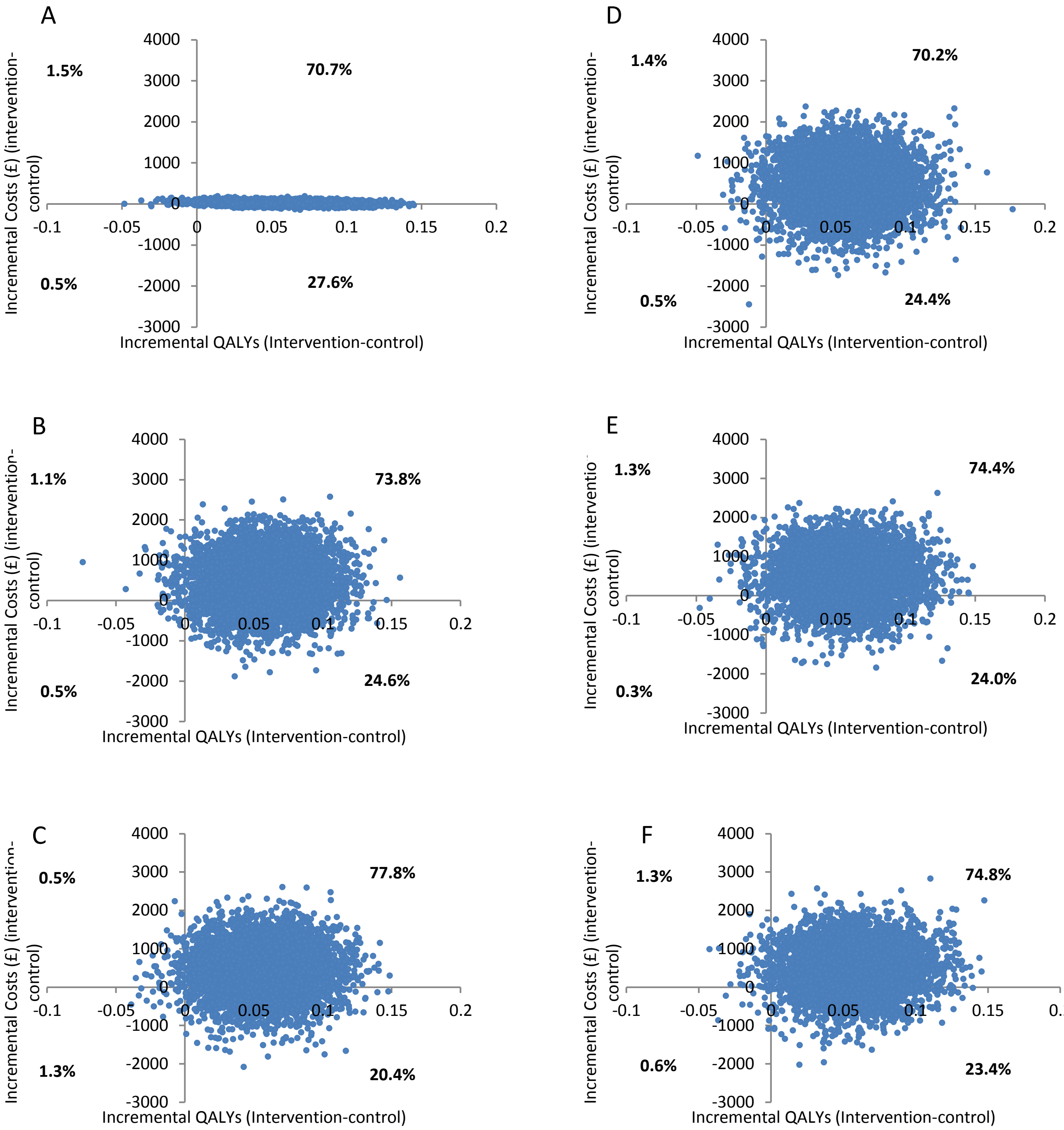


Figure 8:4 Cost-utility plane for the stratified analysis of stratified management care compared to usual care for a societal perspective: High-risk category

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal Cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

8.6.10 Healthcare perspective cost-effectiveness acceptability curves (CEACs)

The CEAC for the incremental cost-effectiveness ratios from a healthcare perspective is shown in Figure 8.5. The findings show that the stratified management intervention was associated with probabilities of being cost-effective exceeding 95% even with the willingness-to-pay value at £1000. This is well below the £30,000 per QALY UK threshold. The curve has an intercept of 96% reflecting the total proportion of points in the SW (0.4%) and SE (95.6%) quadrants. At high WTP values, the height of the curve approaches 99.4% reflecting the proportion of points in the NE (3.9%) and SE (95.6%) quadrants (See Figure 8.1).

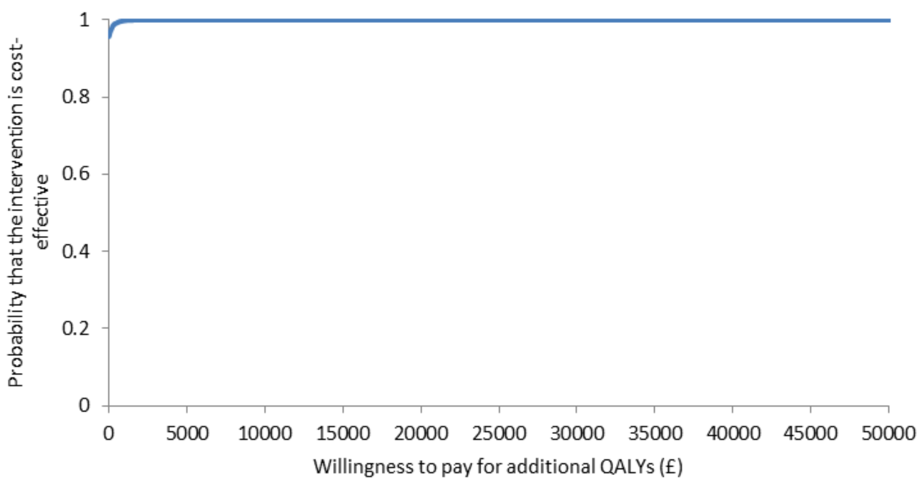


Figure 8:5 CEAC for the healthcare perspective

8.6.11 Societal cost perspective CEACs: overall analysis

The CEACs for the incremental cost-effectiveness ratios with friction period values disaggregated by occupation for the overall analysis are shown in Figure 8.6. The results showed very high probabilities of stratified care being cost-effective for all cost scenarios even with the willingness-to-pay threshold values as low as £5000. The total costs of the interventions were estimated according to cost scenarios with varying friction periods at two-level and five-level occupational classes.

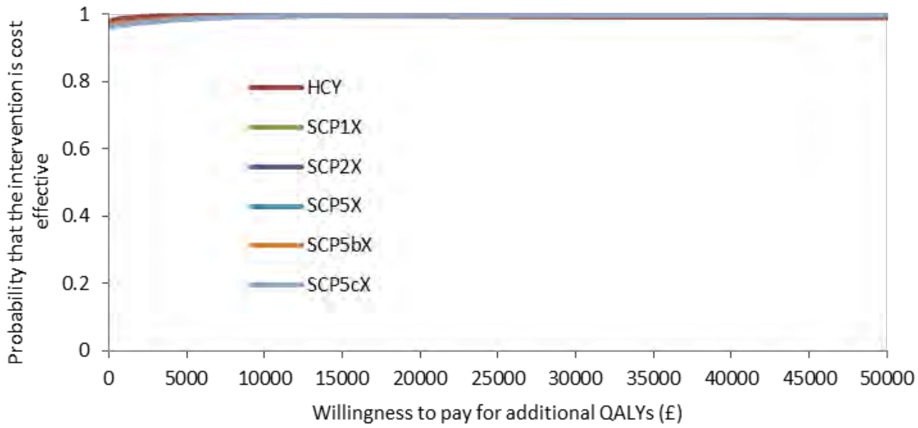


Figure 8:6 CEAC for the overall healthcare and societal perspective

*HCY human capital method, SCP1X Single friction period, SCP2X two-level period DLA, SCP5X five-level period CIPD, SCP5bX five-level period ONS, SCP5bX five-level period UoB.

8.6.12 Healthcare and societal cost perspective CEACs for the stratified analysis

In the healthcare perspective analysis, and at the £30,000 per QALY threshold, the probability of stratified care being cost-effective compared with usual care was highest among medium-risk and high-risk patients for WTP exceeding £10,000 QALYs. More specifically, the probabilities of stratified care being cost-effective at the £30,000 per QALY threshold in the low-risk, medium-risk and high-risk groups were on average across all cost-scenarios 0.52, 0.99, 0.96 (Figure 8.7). The probability of the intervention being cost-effective was highest (exceeding 0.9) for threshold values of £5,000 and below in all three risk groups (Figures 8.8 to 8.11).

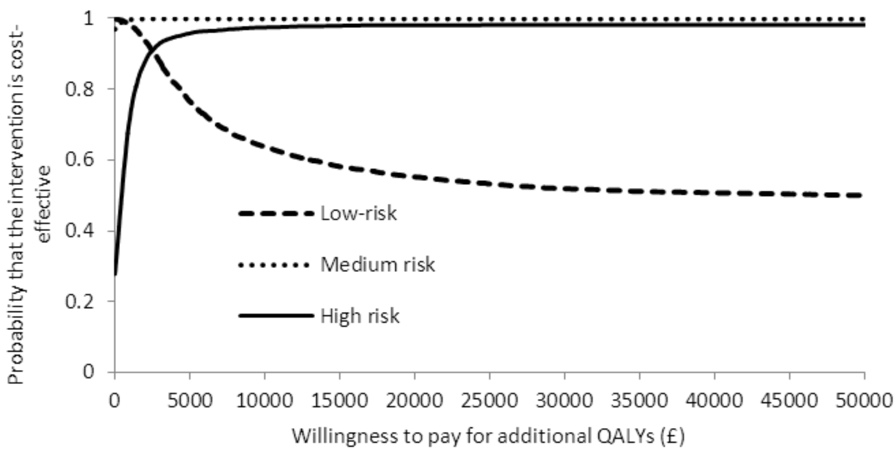


Figure 8:7 CEACs curves for the three risk group comparisons of stratified primary care management compared to current best practice from a healthcare perspective.

In addition, Figure 8.7 represents CEACs categorised by patient risk groups that have been drawn from a population which provided the overall CEAC reported in Figure 8.5. It should therefore be noted that the probability of cost-effectiveness of the intervention at any given WTP that has been reported is not an all-population weighted or unweighted average of the probabilities for the subgroups. This is because the CEAC represents a Bayesian probability that the intervention is cost-effective on the population as a whole and is not intended to show the proportion of patients for whom the intervention is expected to be cost-effective. There is therefore a greater degree of certainty about the cost-effectiveness of the intervention when considered as applying to the whole population than there is when it is applied to any particular subgroup. This is mainly because the all-population results are based on a larger sample size than any subgroup results.

From a societal perspective, the probabilities of stratified care being cost-effective at the £30,000 per QALY threshold in the low-risk, medium-risk and high-risk groups were on average 0.80, 0.99, 0.90 across all cost-scenarios. The medium-risk category was associated with very high probabilities for the intervention being cost-effective for all WTP thresholds. There was much uncertainty in the data regarding cost-effectiveness levels of the intervention at different WTP per QALY gain values after incorporating absenteeism, presenteeism and multiplier effects. For example, there was an increment in the probability of stratified management care being cost-effective among low-risk patients from 0.52 to 0.80 and significant reductions for high-risk patients at threshold values of £10,000 and below.

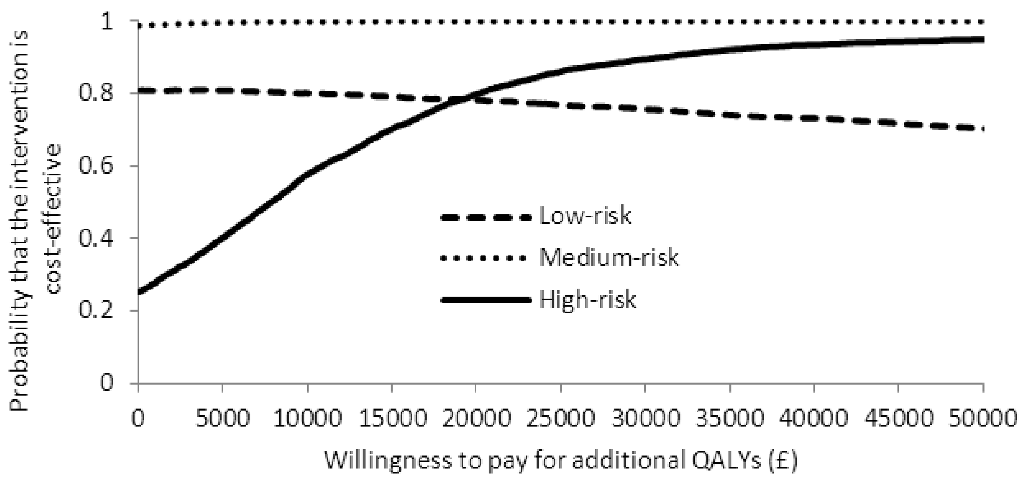


Figure 8:8 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice incorporating productivity costs of absenteeism, presenteeism and multiplier effects based on a single friction period.

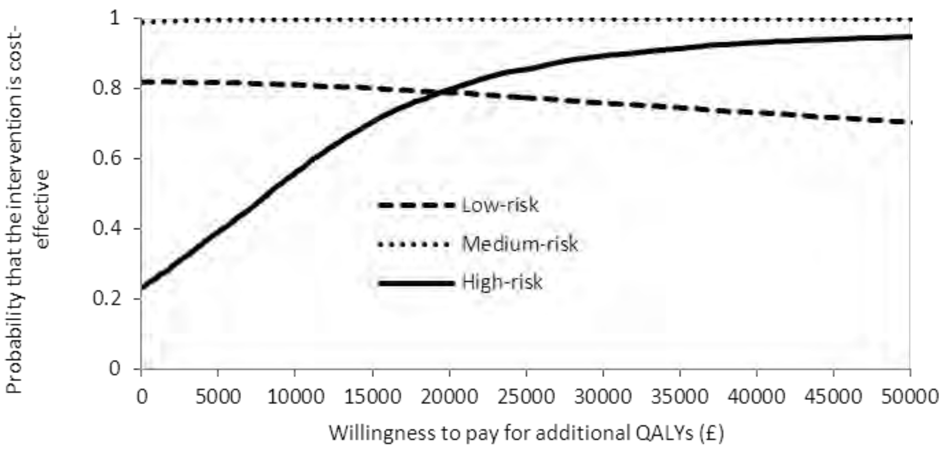


Figure 8:9 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice incorporating productivity costs of absenteeism, presenteeism and multiplier effects based on a two-level category friction period DLA.

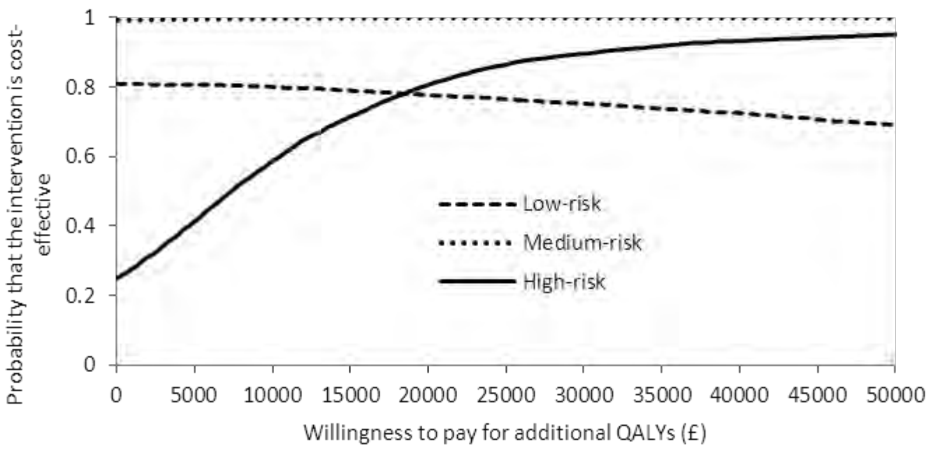


Figure 8:10 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice incorporating productivity costs of absenteeism, presenteeism and multiplier effects based on a five-level friction period CIPD

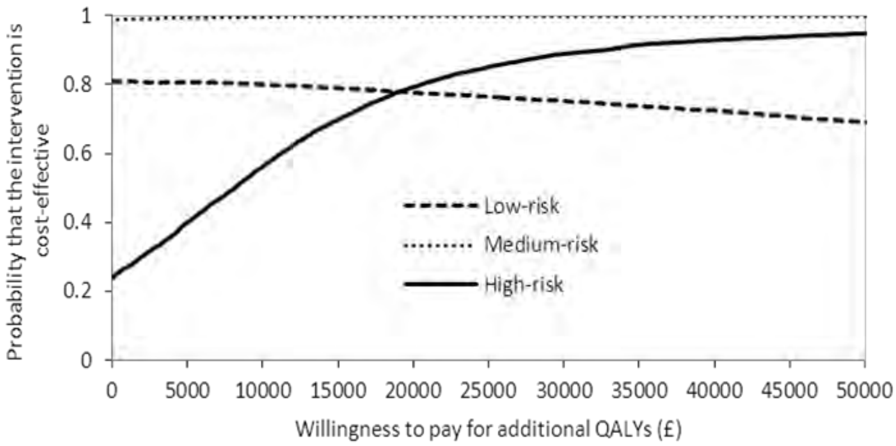


Figure 8:11 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice incorporating productivity costs of absenteeism, presenteeism and multiplier effects based on a five-level friction period ONS

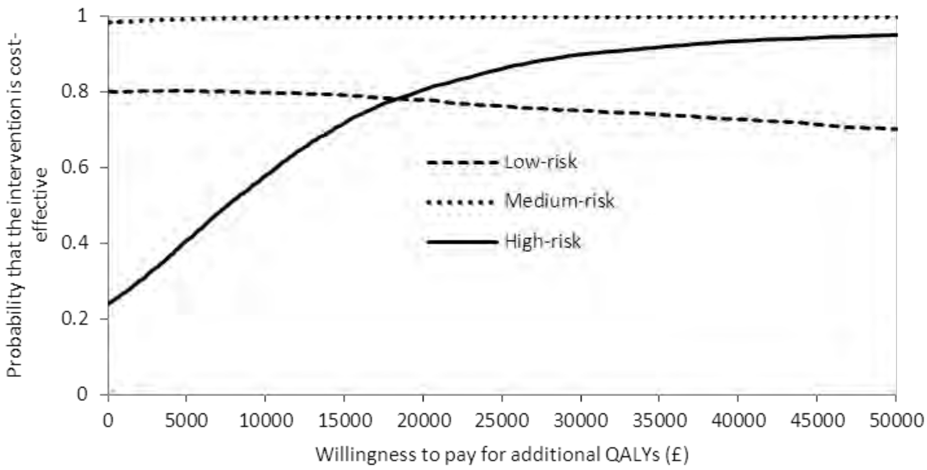


Figure 8:12 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice incorporating productivity costs of absenteeism, presenteeism and multiplier effects based on a five-level friction period UoB

8.6.13 Sensitivity Analysis

8.6.13.1 Regular approach: FCA for absenteeism and HCM for presenteeism excluding multiplier effects.

The fundamental interpretation of the cost-effectiveness analysis in both studies remained the same after applications of the more conventional societal perspective approach, which values presenteeism using the HCM and excludes multiplier effects in the overall analyses (Table 8.9). The intervention remained highly cost-effective at all WTP values investigated for all friction period values in the overall analysis. More specifically, stratified care remained highly cost-effective among medium-risk group patients across all WTP values investigated (Figures 8.13 and 8.14).

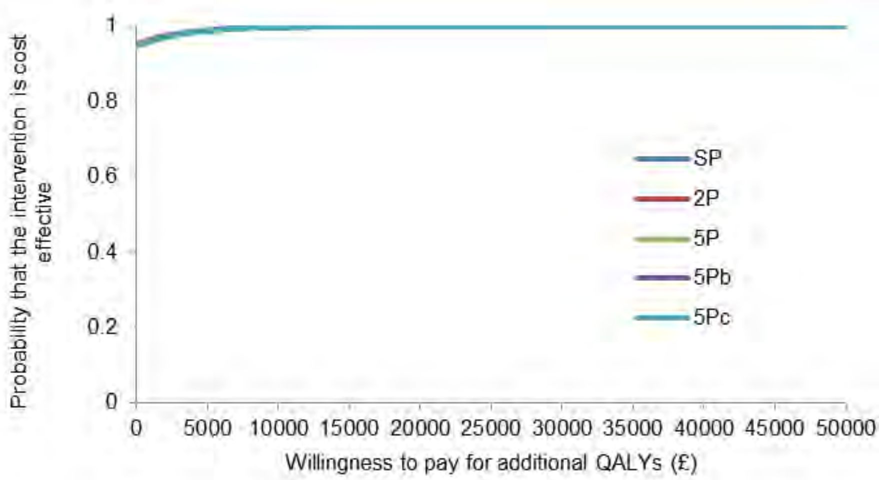


Figure 8:13 CEAC for the societal perspective excluding multiplier effects

* SP Single friction period, SP two-level period DLA, 5P five-level period CIPD, 5Pb five-level period ONS, 5Pc five-level period UoB.

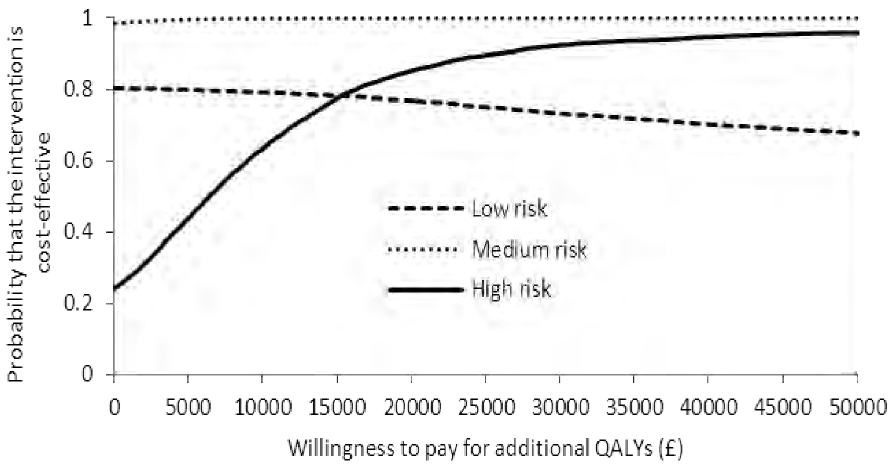


Figure 8:14 CEACs for the three risk group comparisons of stratified primary care management compared to current best practice from a societal perspective based on a single friction period using the FCA for absenteeism and HCA for presenteeism

8.6.14 Sensitivity analysis on cost-utility estimates

Similarly, the cost-effectiveness outcomes remained the same in the sensitivity analysis using the FCA for absenteeism and the HCM for presenteeism (SA1), the application of a higher elasticity factor of 1 assuming no short-term compensation mechanisms (SA2) and using adjusted multiplier values (SA3) (Table 8.9). The results showed that the dominant position for stratified care was not changed in the overall analysis or in the medium-risk patient category.

Table 8:9 Mean cost differences, mean QALY differences, incremental cost-effectiveness ratios from varying friction periods and corresponding three sensitivity analyses (SA1, SA2 and SA3)

	P1	P2	P3	P4	P5
Societal costs (£): Overall Analysis					
Main	-687	-722	-703	-719	-663
SA1	-803	-837	-809	-873	-779
SA2	-477	-491	-488	494	465
SA3	-355	-380	-372	-378	-341
Mean QALY difference	0·039				
ICER Main	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA1	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA2	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA3	Dominant	Dominant	Dominant	Dominant	Dominant
Low-risk category					
Main	-743	-743	-751	-743	-743
SA1	-912	-912	-905	-912	-912
SA2	-467	-465	-475	-467	-467
SA3	-269	-269	-279	-270	-270
Mean QALY difference	-0.001				
ICER Main	743,000	743,000	751,000	743,000	743,000
ICER SA1	912,000	912,000	905,000	912,000	912,000
ICER SA2	466,900	464,500	475,000	475,000	466,900
ICER SA3	269,000	269,000	279,000	270,000	270,000
Medium-risk category					
Main	-1317	-1409	-1345	-1405	-1274
SA1	-1543	-1616	-1552	-1695	-1481
SA2	-984	-1024	-988	-1021	-951
SA3	-1106	-1157	-1110	-1155	-1061
Mean QALY difference	0.044				
ICER Main	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA1	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA2	Dominant	Dominant	Dominant	Dominant	Dominant
High-risk category					
Main	448	445	416	448	431
SA1	535	533	522	535	519
SA2	362	375	338	362	349
SA3	822	818	773	822	796
Mean QALY difference	0.057				
ICER Main	7,860	7,807	7,298	7,860	7,561
ICER SA1	9,386	9,351	9,158	9,386	9,105
ICER SA2	6,346	6,574	5,928	6,349	6,128
ICER SA3	14,421	14,351	13,561	14,421	13,965

Main - Estimates from the suggested approach using the FCA for valuing absenteeism and presenteeism with adjusted multiplier values.

SA1 - Estimates from the conventional approach using the FCA for valuing absenteeism and the HCM for valuing presenteeism

SA2 - Using an elasticity factor value of 1assuming not short-term compensation costs

SA3 adjusting multiplier estimates in the suggested friction cost approach.

P1, single-friction period, P2 – two level friction period DLA, P3 five-level friction period CIPD , P4 five-level friction period ONS, P5 five-level friction period UoB.

8.7 Results 2: The IMPaCT Back study

8.7.1 Study population

A total of 922 participants were considered in the base-case analysis. The study included 368 individuals in the phase 1 strategy (current practice) and 554 individuals in the phase 3 strategy (stratified care). 539 (54%) patients were employed at 6 months, with 215 of 289 (74%) reportedly taking time off in the past 6 months. Table 8.10 shows the work-related outcomes and costs for the overall and stratified analysis according to different risk groups.

8.7.2 Work-related outcomes

Key outcomes for the study are presented in Table 8.10 below. The results for the average productivity costs per patient demonstrate lower productivity loss among patients in the phase 3 strategy compared to the phase 1 strategy. Additionally, work absence results were similar in both intervention phases (25% and 27%: phase 3 and phase 1). The overall absence from work for phase 3 was 4.2 days compared to 7.8 days in phase 1, a difference of 3.6 days (95% CI -8.0 to 0.8; p = 0.109).

In the case of presenteeism, the phase 3 intervention was associated with a productivity loss of 2 days compared to 2.2 days in phase 1, a difference of 0.2 days (95% C: -0.75 to 0.42 ; 0.574). Moreover, the phase 3 intervention was associated with lower absence days in the high-risk (11.0 days versus 5.0 days for phase 1 and phase 3 respectively) and in the medium-risk groups (15.4 days versus 9.0 days for phase 1 and phase 3 respectively). Relatively similar work absence levels were observed between the two interventions among the low-risk patients (1.5 days versus 1.2 days for phase 1 and phase 3 respectively).

Table 8:10 Comparison of overall work-related outcomes for participants in paid employment (work status, absenteeism and presenteeism loss) between the two study arms

Resource/Cost Component	Cost(£)				Mean difference 95% CI
	Phase 1 (n = 132)		Phase 3 (n = 158)		
Work-related outcomes - HCM	n/Mean	%/SD	n/Mean	%/SD	
Time off work due to low back pain (n, %):	35	27%	39	25%	
Absenteeism days	7.8	23.5	4.2	14.32	-3.6 (-8.02,0.82)
Back-pain related absenteeism cost	717	2280.4	343	1106.7	-375 (-789.2,40.2)
Presenteeism days	2.2	2.5	2	2.5	-0.2 (-0.75,0.418)
Back-pain related presenteeism cost	3233	3839.3	2209	2690.9	-1013 (-2600.3,572.6)
Work-related outcomes – FCA incorporating multiplier effects					
Back pain related absenteeism cost Scenario 1	991	3446.8	518	1697.6	-473 (-1102.1,155.6)
Scenario 2	1064	3885.0	516	1690.7	-548 (-1241.4,144.6)
Scenario 3	977	3502.3	490	1632.1	-487 (-1119.0,144.1)
Scenario 4	972	3307.4	514	1684.2	-457 (-1065.6,150.2)
Scenario 5	940	3130.2	507	1669.9	-433 (-1015.1,148.5)
Back pain related presenteeism cost	3721	4563.5	3013	4006.9	-709 (-2783.7, 1366.8)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB,

*Mean difference = Phase 3 mean cost estimate - Phase 1 mean cost estimate

*The estimation of productivity loss focussed on the subsample of respondents in paid employment at 6 months for IBS 132 of 229 (58%) in Phase 1, 158 of 310 (51%) in Phase 3.

8.7.3 Productivity costs: Human capital method related costs

As shown in Table 8.10, the results show that the mean absenteeism costs based on the human capital method were lower in the phase 3 strategy (stratified care) compared to the phase 1 strategy (usual practice) in the overall (£717 versus £343; phase 1 versus phase 3) and risk-defined subgroup analysis with the exception of the low-risk patients category (£223 versus £29; phase 1 versus phase 3) (Table 8.10 and 8.11). Mean presenteeism costs were lower in the phase 3 strategy than in the phase 1 strategy in the overall analysis and risk-defined subgroup analysis with the exception of the low-risk category.

8.7.4 Methodological applications of the friction cost approach

Resource use data on the total work-related costs in relation to the phase 3 and phase 1 strategies is presented in Table 8.10. Similar to the SBT results, the results are presented for five cost scenarios generated by the friction cost approach: i) scenario 1 applies a single friction period ii) scenario 2 applies a two occupational level friction period iii) scenarios three, four and five apply five occupational level friction periods from 3 different datasets. As shown in Table 8.10, estimating productivity costs based on occupation generated relatively similar productivity costs for the different scenarios with varying friction periods. The observed mean costs in the phase 3 strategy were between £490 and £518 compared with over £900 in the phase 1 strategy depending on the friction period occupational level used (Table 8.10). The corresponding presenteeism costs were £3721 and £3693 for the phase 1 and 3 strategy respectively.

Mean productivity costs for the phase 3 and phase 1 strategy for the subgroup analyses are presented in Table 8.11. The results from Table 8.11 indicate there were no observed differences in the mean productivity costs across the friction period cost scenarios in all risk groups.

Table 8:11 Comparison of work-related outcomes for participants in paid employment (work status, absenteeism and presenteeism loss) between the two groups

Resource/Cost Component	**Cost(£)**				**Mean difference** (95% CI)
	Phase 1		Phase 3		
Low-risk	**(n = 52)**		**(n = 62)**		
	n/Mean	%/SD	n/Mean	%/SD	
Time off work due to low back pain: (n, %)	5	10%	7	11%	
Back-pain related absenteeism cost					
Scenario 1	29	104.8	223	1043.1	193 (-100.12, 487.04)
Scenario 2	29	104.8	223	1043.1	193 (-100.12 ,487.04)
Scenario 3	29	104.8	223	1043.1	193 (-100.12 ,487.04)
Scenario 4	29	104.8	223	1043.1	193 (-100.12 ,487.04)
Scenario 5	29	104.8	223	1043.1	193 (-100.12, 487.04)
Back pain related presenteeism cost	2774	3789.2	4530	5549.8	1756 (-4883, 8395.7)
Medium-risk	**(n=56)**		**(n=67)**		
Time off work due to low back pain: n (%)	19	34%	19	29%	
Absenteeism					
Scenario 1	1477	3840.4	438	1182.9	-1038 (-2039.7,-372.0)
Scenario 2	1577	4315.9	433	1159.1	-1144 (-2255.8,-32.4)
Scenario 3	1469	3906.9	373	934.7	-1096 (-2094.7,-98.2)
Scenario 4	1471	3741.5	429	1136.6	-1042 (-2106.6,-68.4)
Scenario 5	1400	3481.7	412	1084.3	-987 (-1896.0,-78.5)
Back pain related presenteeism cost	2788	3291.5	1982	3347.3	-805 (-3130.2,1518.7)
High-risk	**(n=24)**		**(n=29)**		
Time off work due to low back pain: n (%)	11	46%	13	45%	
Absenteeism					
Scenario 1	1883	5297.3	1341	3117.0	-542 (-2955.7,1871.2)
Scenario 2	2045	6059.1	1341	3117.0	-704 (-3371.5,1963.5)
Scenario 3	1826	5396.0	1341	3117.0	-484 (-2930.0,1961.2)
Scenario 4	1790	4976.0	1341	3117.0	-449 (-2758.5,1859.9)
Scenario 5	1785	4795.0	1341	3117.0	-444 (-2695.7,1806.9)
Back pain related presenteeism cost	7001	7994.7	5677	6318.5	-1324 (-7522.2,4873.7)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB

*The estimation of productivity loss focussed on the subsample of respondents who reported being currently employment at 6 months for IBS (290/547).

8.7.5 IBS Friction cost approach: Impact of absenteeism, presenteeism and multiplier effects

In the IBS analysis, an investigation was carried out on the potential impact on total societal costs of using the friction cost approach for valuing productivity costs while incorporating adjustments for multiplier effects. These estimates were then compared with the conventional societal cost estimates.

As shown in Table 8.12, adjusting the conventional friction cost approach to productivity costs for multiplier effects leads to higher societal costs than the conventional friction cost approach and human capital generated estimates. In all cost scenarios, the main cost drivers were the effects of productivity loss on team productivity. Mean societal costs were lower in the phase 3 strategy than the phase 1 strategy.

Table 8:12 Back pain mean (SD) societal costs per patient and year for patients reporting resource utilisation data: Overall analysis

	Phase 1 (n=368)		Phase 3 (n=554)		Difference (95% CI BC)
	Mean	SD	Mean	SD	
Absence days*	7.9	23.5	4.3	14.5	0·49 (0·44, 0·55)
Health-care costs (£)	274	278.2	240	261.3	-34.4 (-76.87; 3.3)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	937	3074.4	583	2023.4	-354 (-683.4,-24.5)
Scenario 2	963	3287.3	582	2021.1	-380 (-723.4,-36.8)
Scenario 3	932	3010.0	576	2000.6	-356 (-686.2,-27.7)
Scenario 4	930	3008.2	582	2018.6	-348 (-673.1,-23.3)
Scenario 5	919	2932.2	580	2014.5	-339 (-658.8,-19.3)
Total back pain related costs using FCA approach for absenteeism and HCM for presenteeism					
Scenario 1	750	2130.4	534	1249.1	-285 (-503.4,-65.8)
Scenario 2	763	2225.3	464	1245.8	-298 (-523.6,-73.4)
Scenario 3	743	2126.2	458	1217.3	-285 (-501.6, 68.4)
Scenario 4	745	2096.6	464	1241.6	-281 (-497.4,-65.2)
Scenario 5	739	2064.4	462	1237.5	-276 (-490.6,-63.3)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB

*The analysis of work absence outcomes focused on the subsample of respondents in paid employment at 6-month follow-up; 132 of 229 (58%) in Phase 1, 158 of 310 (51%) in Phase 3.

8.7.6 IBS Friction cost approach: Impact of absenteeism, presenteeism and multiplier effects in subgroup analysis

The findings in the overall analysis were replicated in the subgroup analyses. Productivity costs incorporating multiplier effects generated higher costs than the more conventional approach that excludes multiplier effects in all risk groups. The differences in costs between the interventions were largely driven by the effects of absenteeism and presenteeism on team productivity (Table 8.13 to 8.15). The phase 1 strategy was associated with lower mean societal costs compared to the phase 3 strategy in the low-risk patient category (Table 8.13).

Table 8:13 Back pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: Low-risk

	Phase 1 n=136		Phase 3 N=214		Difference (95% CI BC)
	Mean	SD	Mean	SD	
Absence days*	0·52	2.29	0·85	3.40	-0.32 (-1.43,0.78)
Health-care costs (£)	141	210.1	138	171.3	2.90 (-39.3,45.3)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	243	821.4	306	1523.8	63 (-216.6,342.9)
Scenario 2	243	821.4	306	1523.8	63 (-216.6,342.9)
Scenario 3	243	821.4	306	1523.8	63 (-216.6,342.9)
Scenario 4	243	821.4	306	1523.8	63 (-216.6,342.9)
Scenario 5	243	821.4	306	1523.8	63 (-216.6,342.9)
Total back pain related costs using FCA approach for absenteeism and HCM for presenteeism					
Scenario 1	232	727.2	224	750.9	-8.1 (-153.4, 176.3)
Scenario 2	232	727.2	223	750.9	-8.1 (-139.5, 184.8)
Scenario 3	232	727.2	223	750.9	-8.1 (-144.4, 186.6)
Scenario 4	232	727.2	223	750.9	-8.1 (-147.6, 169.7)
Scenario 5	232	727.2	223	750.9	-8.1 (-125.7, 190.9)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB

*Work loss evaluation was based on a total subsample of 290 of 547 responders who reported being currently employed at 6 months follow-up

Table 8:14 Back pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: Medium-risk

	Phase 1 n=151		Phase 3 N=232		Difference (95% CI)
	Mean	SD	Mean	SD	
Absence days	6.11	20.0	2.3	12.8	-3.7 (-7.8, 0.43)
Health-care costs (£)	292	361.3	284.0	314.1	-7.79 (-73.4,56.9)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1107	3206.4	534	1371.2	-573 (-1041.8,-105.4)
Scenario 2	1144	3451.0	532	1363.1	-611 (-1107.4,-115.8)
Scenario 3	1105	3234.6	516	1287.9	-588 (-1054.3,-123.5)
Scenario 4	1105	3154.7	531	1354.1	-574 (-1035.3,-113.4)
Scenario 5	1079	3026.4	527	1339.2	-553 (-998.0, -107.3)
Total back pain related costs using FCA approach for absenteeism and HCM presenteeism					
Scenario 1	834	2064.1	462	1034.5	-372 (-727.4,-49.7)
Scenario 2	854	2180.6	461	1025.2	-393 (-816.0,-62.7)
Scenario 3	830	2067.5	446	939.3	-385 (-810.5,-61.2)
Scenario 4	834	2038.8	459	1012.7	-374 (-751.2,-63.2)
Scenario 5	819	1980.2	455	1000.7	-364 (-768.3, -61.9)

Scenario 1, single-friction period, Scenario 2 – two level friction period, Scenario 3 five-level friction period 1st data set , Scenario 4, five-level friction period 2nd dataset, Scenario 5 five-level friction period 3rd dataset

*Work loss evaluation was based on a total subsample of 290 of 547 responders who reported being currently employed at 6 months follow-up

Table 8:15 Back pain related mean (SD) societal costs per patient and year for patients reporting resource utilisation data: High-risk

	Phase 1 (n=81)		Phase 3 (n=108)		Difference (95% CI)
	Mean	SD	Mean	SD	
Absence days	7.89	26.3	4.47	12.44	-3.41 (-10.9,4.16)
Health-care perspective (£)	384	304.6	358	323.5	26 (-53.9; 121.0)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1784	4621.0	1237	3447.9	-547 (-1704.9,610.3)
Scenario 2	1832	4939.5	1237	3447.9	-595 (-1799.2,608.7)
Scenario 3	1767	4657.8	1237	3447.9	-530 (-1693.2,632.6)
Scenario 4	1757	4490.3	1237	3447.9	-520 (-1658.7,619.2)
Scenario 5	1756	4427.3	1237	3447.9	-519 (-1648.3,611.7)
Total back pain related costs using FCA approach for absenteeism and presenteeism					
Scenario 1	1464	3309.5	953	2067.2	-512 (-1393.8, 210.0)
Scenario 2	1488	3446.8	953	2067.2	-536 (-1472.3, 226.0)
Scenario 3	1442	3298.0	953	2067.2	-490 (-1503.9, 278.9)
Scenario 4	1446	3243.6	953	2067.2	-493 (- 1257.5,270.9)
Scenario 5	1445	3217.4	952	2067.2	-493 (-1252.7, 267.7)

Scenario 1, single-friction period, Scenario 2 – two level friction period DLA, Scenario 3 five-level friction period CIPD , Scenario 4 five-level friction period ONS, Scenario 5 five-level friction period UoB

*Work loss evaluation was based on a total subsample of 290 of 547 responders who reported being currently employed at 6 months follow-up

8.7.7 Estimation of the cost-utility analysis from a societal perspective

In the healthcare perspective, stratified care was associated with a small mean health benefit at a lower mean healthcare cost (0.003 additional QALYs and a reduced mean healthcare cost of -£33.54). Similarly, when considering the societal perspective for the overall analysis, a dominant position was observed for the phase 3 strategy, which was associated with a small increased health benefit achieved at a lower cost (0.003 additional QALYs and reduced mean societal costs ranging from -£380 to -£339).

The investigation was then carried out with a focus on subgroups of patients, comprising low-risk, medium-risk and high-risk patients. The results demonstrated a small but significant difference between the two groups among high-risk patients (0.0196 additional QALYs and reduced mean costs ranging from -£595 to -£519 (Tables 8.14 and 8.15). The high-risk cost-effectiveness planes show that on average 80% of the bootstrapped cost-utility pairs were situated in the southeast quadrant, indicating more effectiveness and fewer costs for the phase 3 strategy. Accordingly, stratified care dominates usual care, i.e. is more cost-effective regardless of the willingness-to- pay per QALY among high-risk patients. A dominant position negates the need to calculate an ICER.

However, there was considerable uncertainty around the incremental costs and effects in the low-risk category. A small health benefit of 0.00294 was observed at an increased average cost of £63 across all cost scenarios, resulting in an average incremental cost-per QALY gained of £21,429 when considering friction periods disaggregated by occupation. The results in this risk group demonstrate that the additional health benefits are worth the additional cost of the phase 3 intervention (Table 8.16).

Finally for the medium-risk group, stratified care was associated with lower mean societal costs ranging from -£611 to -£553 and a lower mean health benefit (0.00795 fewer QALYs). The observed incremental cost-effectiveness ratios ranging from £69,560 to £76,855 indicate that the additional cost required to achieve the small health benefit associated with the usual care group is too expensive (Table 8.16).

Table 8:16 Incremental societal costs and incremental QALYs: Point estimates incorporating absenteeism and presenteeism for the IBS base-case analysis

	P1	P2	P3	P4	P5
Overall Analysis					
Mean cost difference per patient (£)	-354	-380	-357	-349	-340
Mean QALY difference			0.003		
ICER: Cost per QALY gained	Dominant	Dominant	Dominant	Dominant	Dominant
Low-risk category					
Mean cost difference per patient (£)	63	63	63	63	63
Mean QALY difference			0.00294		
ICER: Cost per QALY gained	21,429	21,429	21,429	21,429	21,429
Medium-risk category					
Mean cost difference per patient (£)	-573	-611	-588	-574	-533
Mean QALY difference			-0.00795		
ICER: Cost per QALY gained	72,075	76,855	73,962	72,201	69,560
High-risk category					
Mean cost difference per patient (£)	-547	-595	-530	-520	-519
Mean QALY difference			0.01956		
ICER: Cost per QALY gained	Dominant	Dominant	Dominant	Dominant	Dominant

ICER: incremental cost-effectiveness ratio; QALY quality adjusted life year. P1 single-occupational category friction period, P2 two-level occupational category friction period, P3, P4, P5 five-level occupational category friction periods

8.7.8 Cost-utility planes – overall healthcare and societal cost perspective

Figure 8.15 shows the cost-effectiveness (CE) planes of incremental costs versus incremental QALYs for the overall analysis, based on 5000 bootstrap replicates. The planes illustrate the level of uncertainty in the mean incremental costs and QALY estimates. The findings demonstrate cost savings and increased health benefits in the overall analysis (the majority of replications (63%) located in the south-east quadrant).

In the societal cost analysis (Figure 8.15) the dominant position of stratified care (phase 3) was further strengthened with an average 67% of replications located in the southeast quadrant (showing greater benefits and reduced costs).

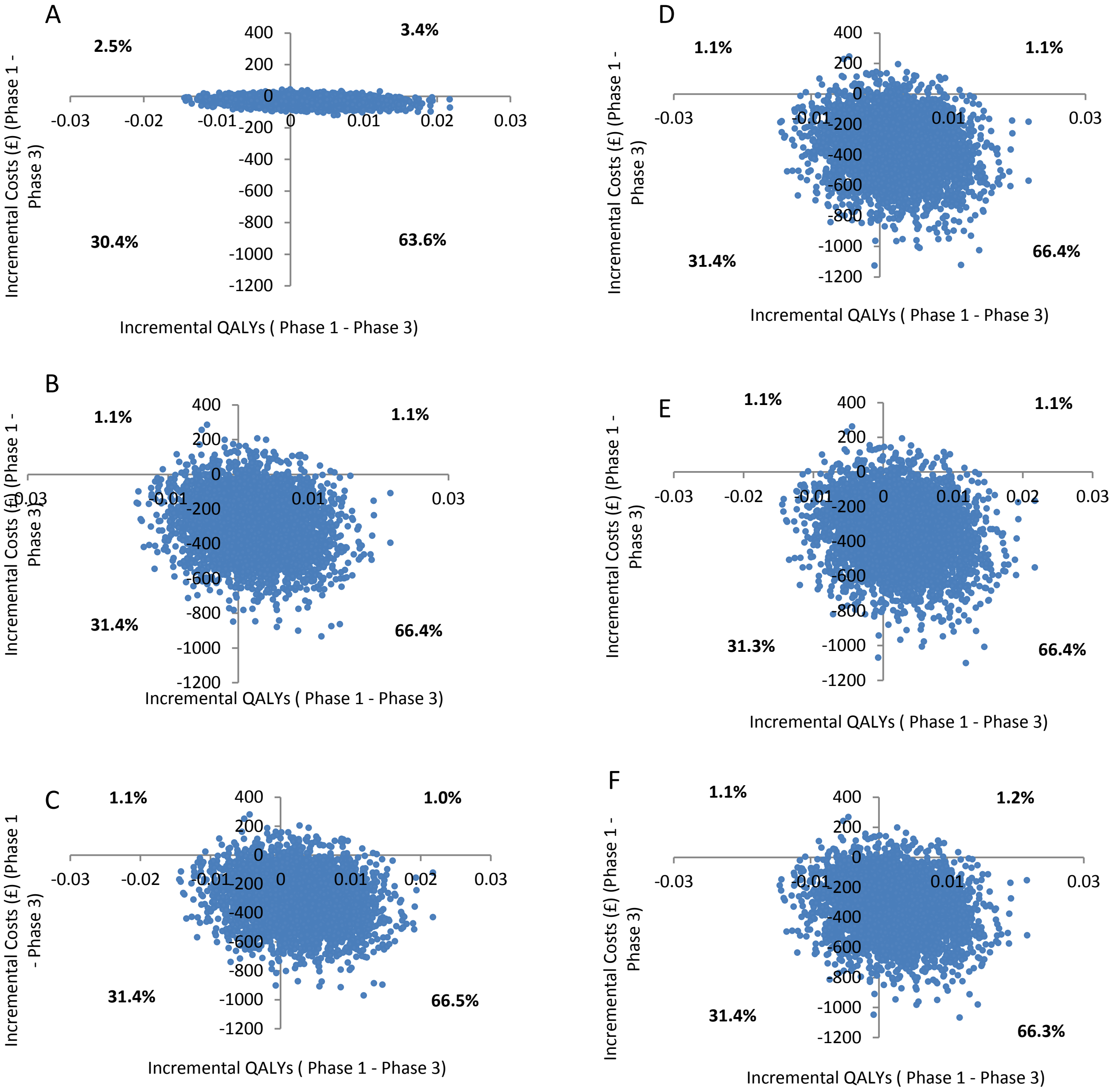


Figure 8:15 Cost-utility plane for the overall comparison of stratified care (phase 3) compared to usual care (phase 1) for the societal perspective

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

8.7.9 Cost-effectiveness planes for the subgroup analysis

Cost-effectiveness planes for the risk-group analysis showing the uncertainty around the cost-effectiveness ratios of the main QALY outcome are presented in Figure 8.16 to 8.18. There was much uncertainty in the bootstrapped data among the low-risk patients as seen by the distribution of the cost-effectiveness pairs in the four quadrants of the CU planes (Figures 8.16). In all cost scenarios, on average 48% of the replications were located in the NE quadrant.

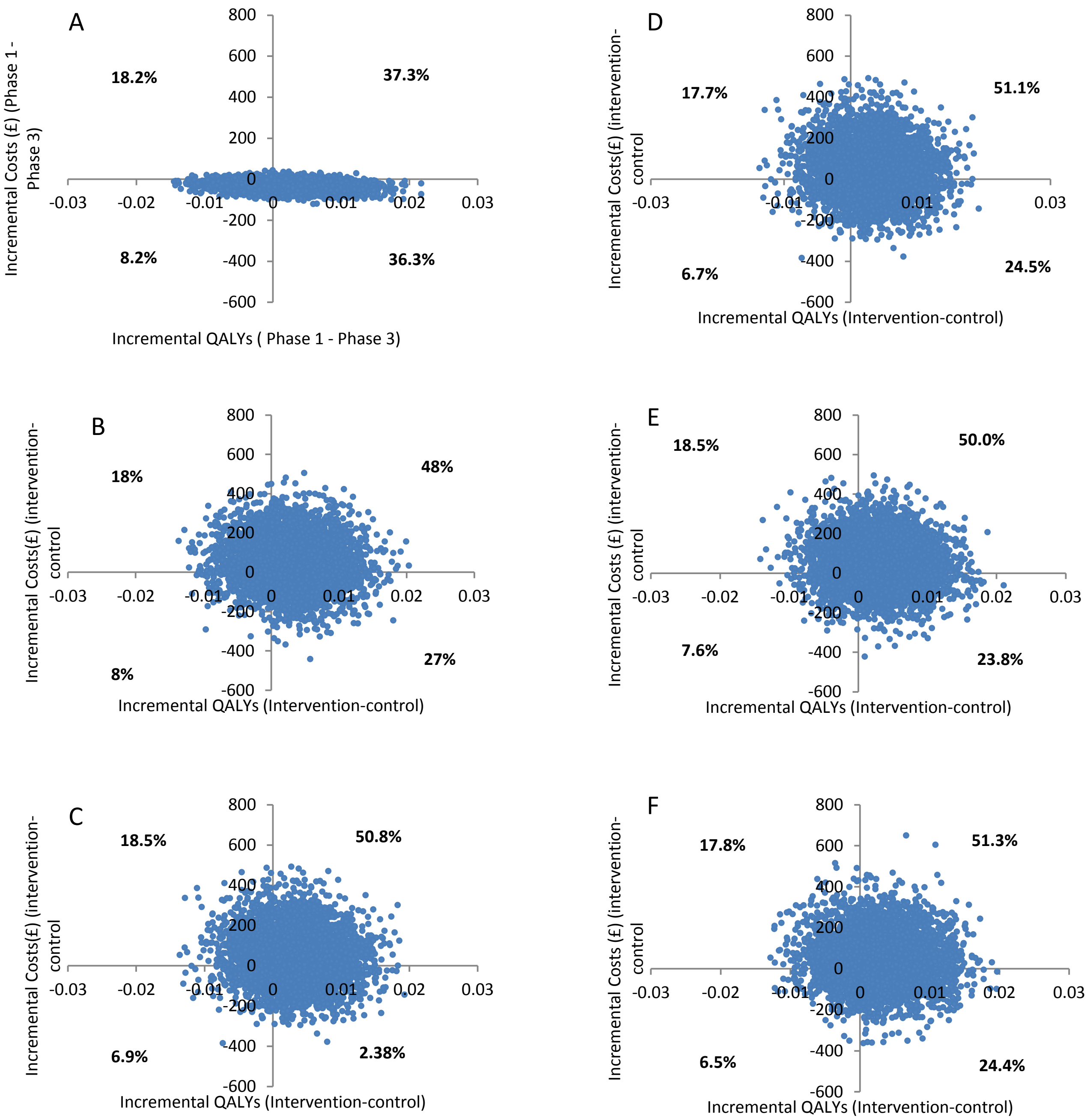


Figure 8:16 Cost-utility plane for the low-risk group analysis of stratified care (phase 3) compared to usual care (phase 1) from societal perspective

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB

The scatter plots for the medium-risk and high-risk patients are displayed in Figures 8.17 and 8.18. The intervention was associated with cost-savings, but no significant differences in health benefits between the two groups among medium-risk group patients. The majority (average 85% for all cost-scenarios) of the cost-utility pairs were located in the south-west quadrant (Figures 8.17). In contrast, the stratified care strategy was associated with cost-savings and increased effectiveness among the high-risk group, as the majority (average 80%) of cost-utility pairs were located within the south-east quadrant (Figure 8.18).

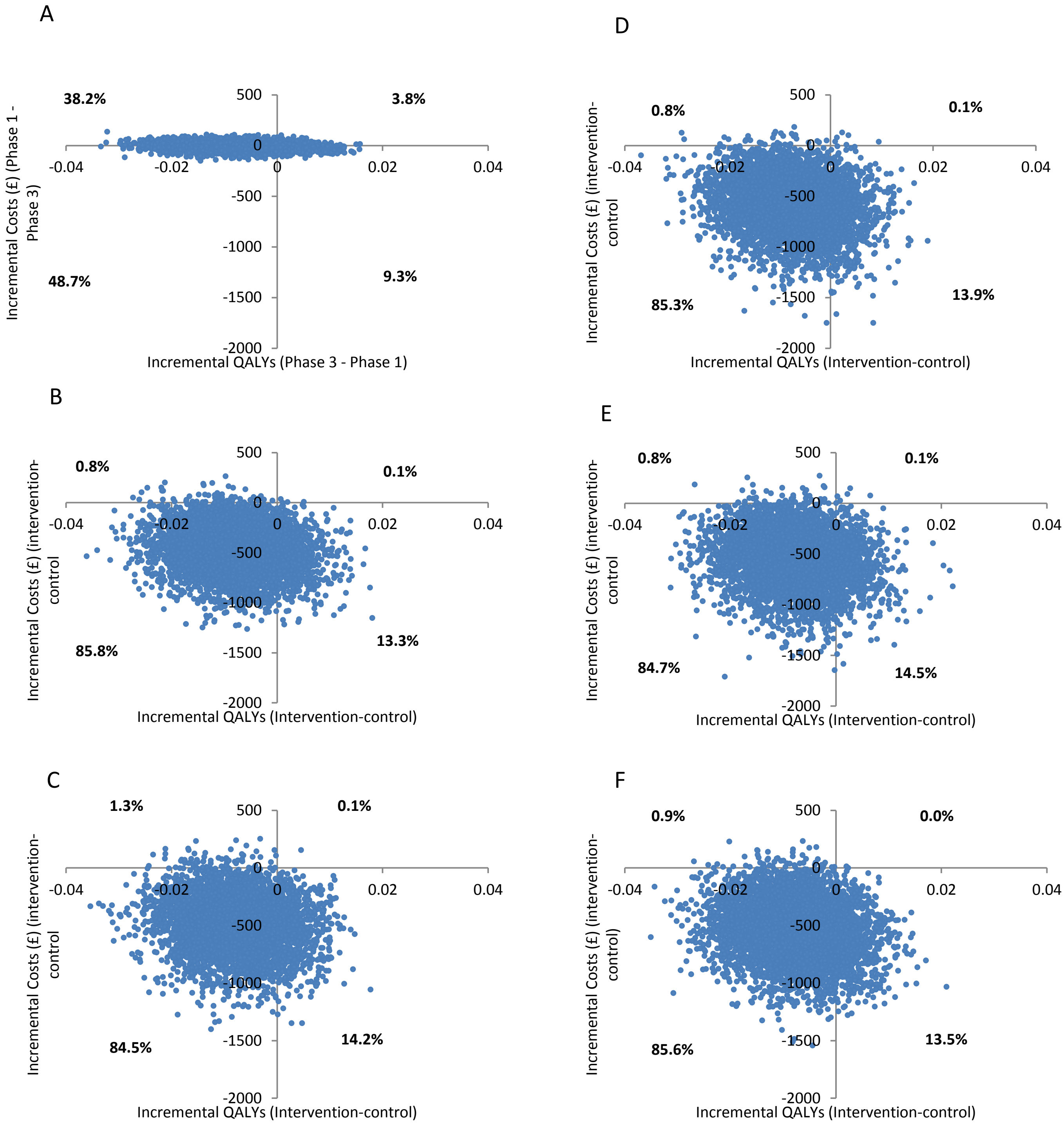


Figure 8:17 Cost-utility plane for the Medium-risk group analysis of stratified care (phase 3) compared to usual care (phase 1) from societal perspective

(A) Cost–utility plane comparing the stratified management approach (‘intervention’) to current best practice (‘control’) healthcare perspective. (B) Societal cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB.

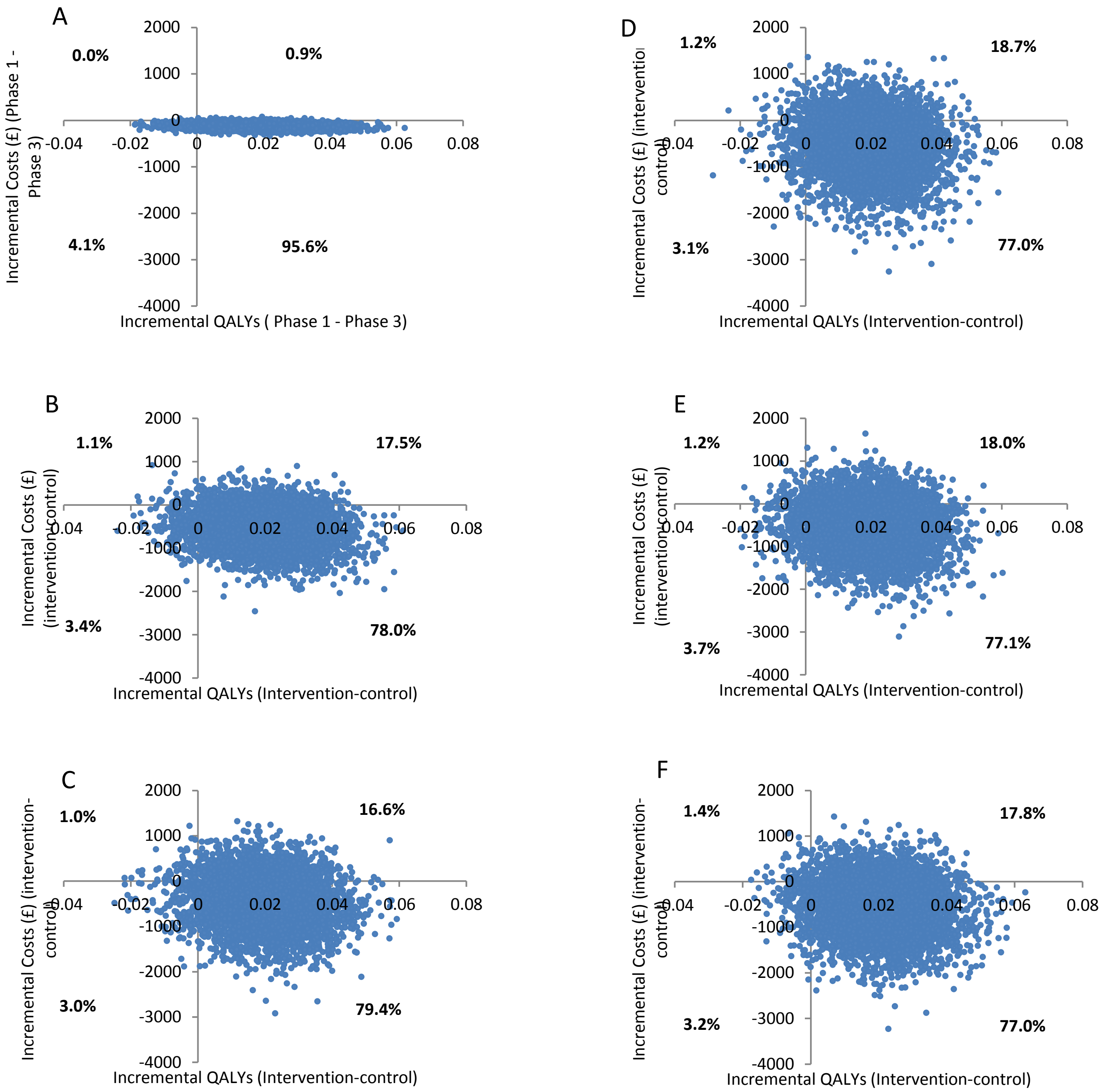


Figure 8:18 Cost-utility plane for the High-risk group analysis of stratified care (phase 3) compared to usual care (phase 1) from societal perspective

(A) Cost–utility plane comparing the stratified management approach ('intervention') to current best practice ('control') healthcare perspective. (B) Societal Cost perspective average friction period. (C) Societal cost perspective two-level friction period DLA. (D) Societal cost perspective five-level friction period CIPD. (E) Societal cost perspective five-level friction period ONS. (F) Societal cost perspective five-level friction period UoB

8.7.10 Healthcare and societal cost perspective cost-effectiveness acceptability curves (CEACs)

Figure 8.19 presents the cost effectiveness acceptability curve showing the probability of the phase 3 strategy being cost-effective compared with the phase 1 strategy for the overall analysis from a healthcare and societal perspective. The results show that the phase 3 strategy was associated with relatively high probabilities of being cost-effective for all willingness-to-pay thresholds investigated (up to £50,000 per QALY), but with the greatest probability of the intervention being cost-effective observed at low threshold values (below £10,000).

The CEACs for the societal perspective for the overall analysis are also presented in Figure 8.19 alongside the detailed friction period (varying friction periods at two and five-occupation levels) cost scenarios. As in the healthcare perspective, the phase 3 strategy was associated with a very high probability of being cost-effective (over 95% compared to 75% from the healthcare perspective) for the majority of willingness-to-pay thresholds investigated. The inclusion of productivity costs therefore further strengthened the case for the phase 3 strategy. Similar results were demonstrated in all cost scenarios applying the more detailed friction periods (Figure 8.19).

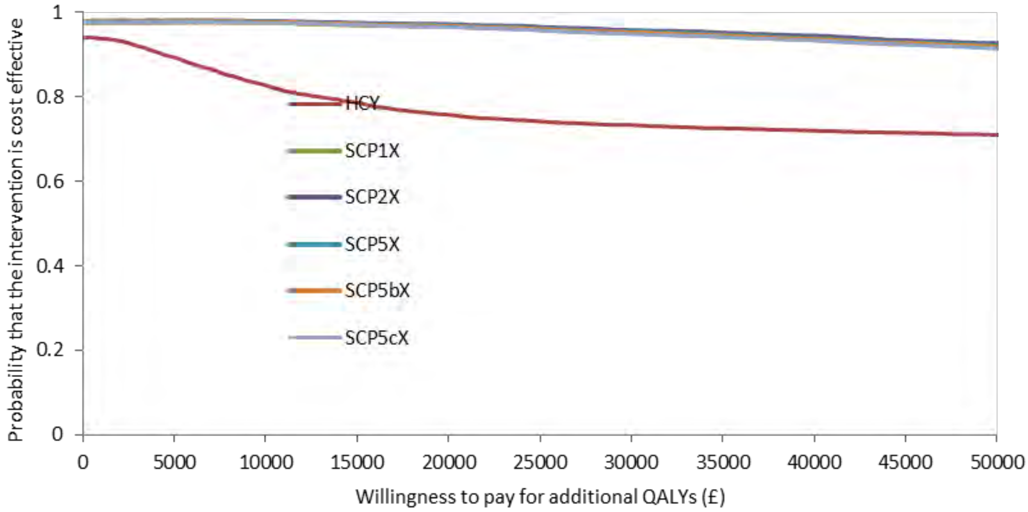


Figure 8:19 CEAC for the overall healthcare and societal perspective

HCY Healthcare perspective, SCP1X Single friction period, SCP2X two-level period DLA, SCP5X five-level period CIPD, SCP5bX five-level period ONS, SCP5bX five-level period UoB.

8.7.11 Cost-effectiveness acceptability curves (CEACs) for the subgroup analysis

The CEACs for the healthcare perspective and societal cost perspectives are presented in Figures 8.20 to 8.25 with all three risk group analyses presented on the same figure. In the analyses from the healthcare perspective, and at the £30,000 per QALY threshold, the phase 3 strategy was associated with very high probabilities (exceeding 95%) of being cost-effective among high-risk patients. For the medium-risk group, the plateau of around 20% for probability of stratified care (phase 3 strategy) being cost-effective at the £30,000 threshold reflects the negligible difference in benefits between the two groups. The probabilities that stratified care was cost-effective at the £30,000 per QALY threshold in low-risk, medium-risk and high-risk patients were on average across all cost-scenarios (i.e. scenarios varying friction periods at one, two and five-occupation levels) 0.70, 0.20, 0.98.

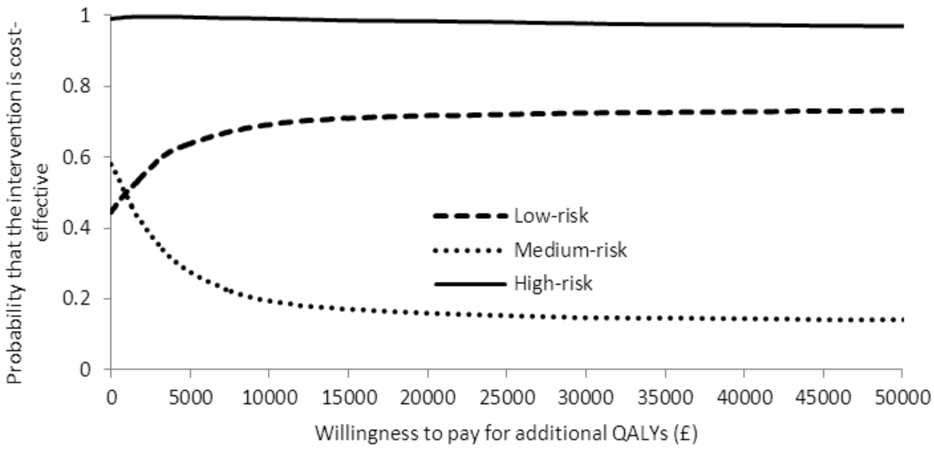


Figure 8:20 Cost-utility acceptability curves for the three risk groups analysis of stratified care (phase 3) compared to usual care (phase 1) (healthcare perspective).

8.7.11.2 Cost-effectiveness acceptability curves: Societal cost perspective

The inclusion of societal costs significantly altered the probabilities of stratified care being cost-effective in the subgroup analyses, particularly among medium-risk patients. Societal costs were calculated according to cost scenarios with varying friction periods at two-level and five-level occupational classes. The probabilities that stratified care was cost-effective at the £30,000 per QALY threshold in low-risk, medium-risk and high-risk patients were on average across all cost-scenarios 0.50, 0.80, 0.90. Slight but insignificant changes in the probability of stratified care being cost-effective were observed in the different cost scenarios. Notable, however, was the increment in the probability of stratified care being cost-effective among medium-risk patients from 0.20 to 0.80 at the £30,000 per QALY threshold. The improved cost-effectiveness of stratified care at low WTP could potentially be attributed to societal cost savings.

However, the stratified care strategy remained highly cost-effective among high-risk patients with probabilities exceeding 90% of being cost-effective at WTP threshold values greater than £5000 per QALY gained for all cost scenarios (Figures 8.20 to 8.25). In the low-risk category, the correspondingly observed probability plateau averaging 50% in all cost scenarios reflects the relatively small difference in QALY gains (0.00294 QALYs).

Overall, there were no significant alterations in the probability of the phase 3 strategy being cost-effective within the different cost scenarios as a result of using stratified occupational level friction periods. Consequently, the same policy conclusions are observed irrespective of the level of friction period used.

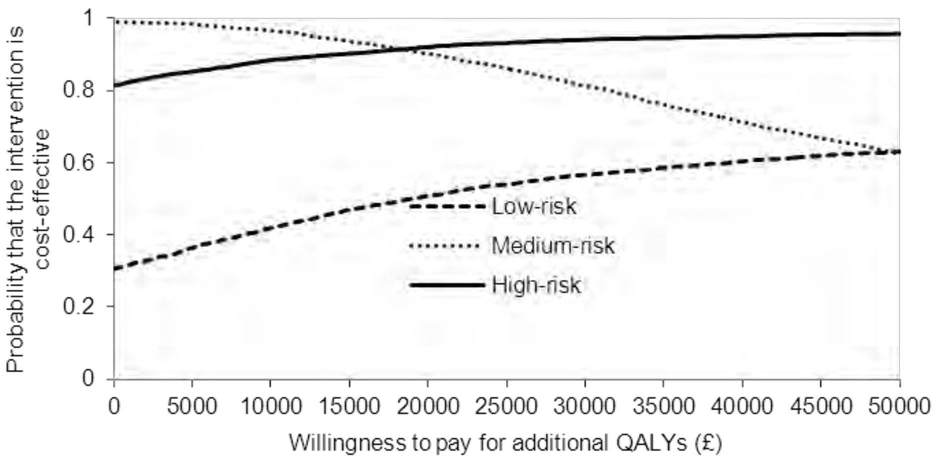


Figure 8:21 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (single occupation level friction period).

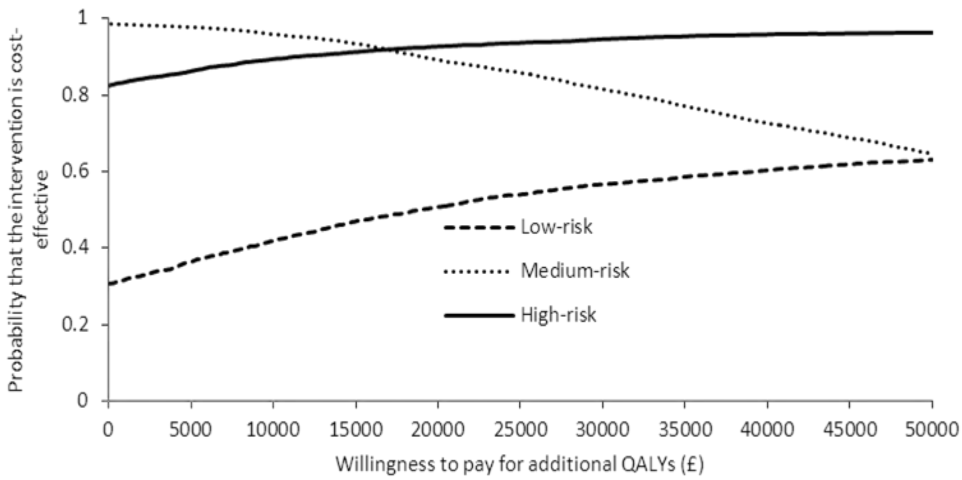


Figure 8:22 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (two level occupation level friction period DLA)

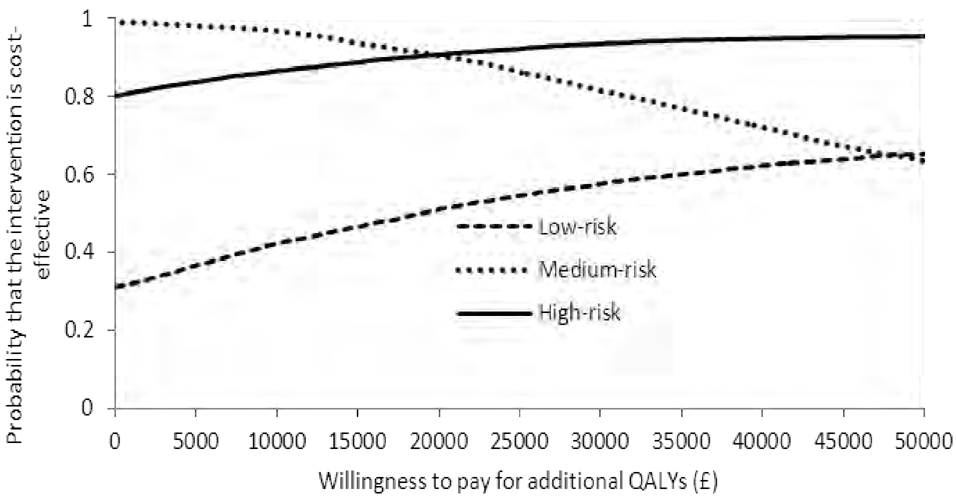


Figure 8:23 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (five occupation level friction period CIPD).

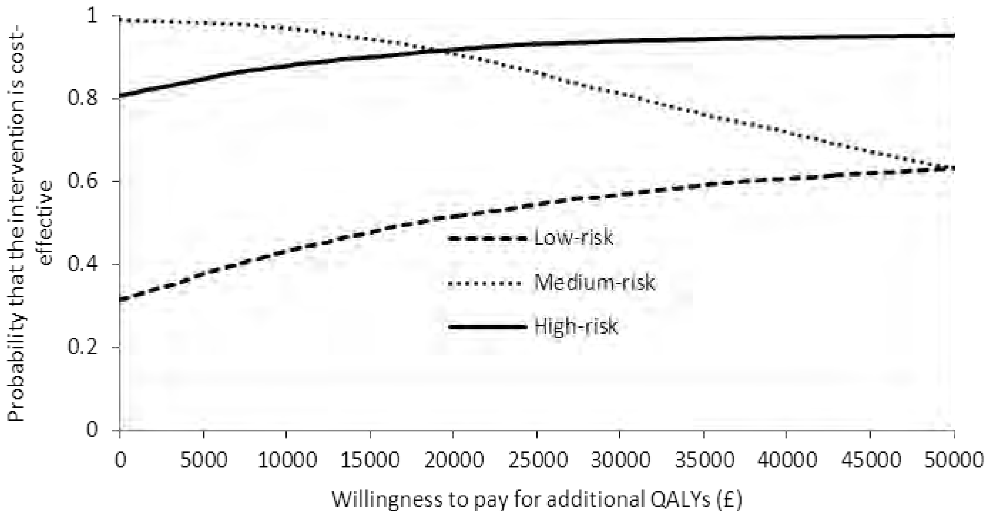


Figure 8:24 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (five occupation level friction period ONS)

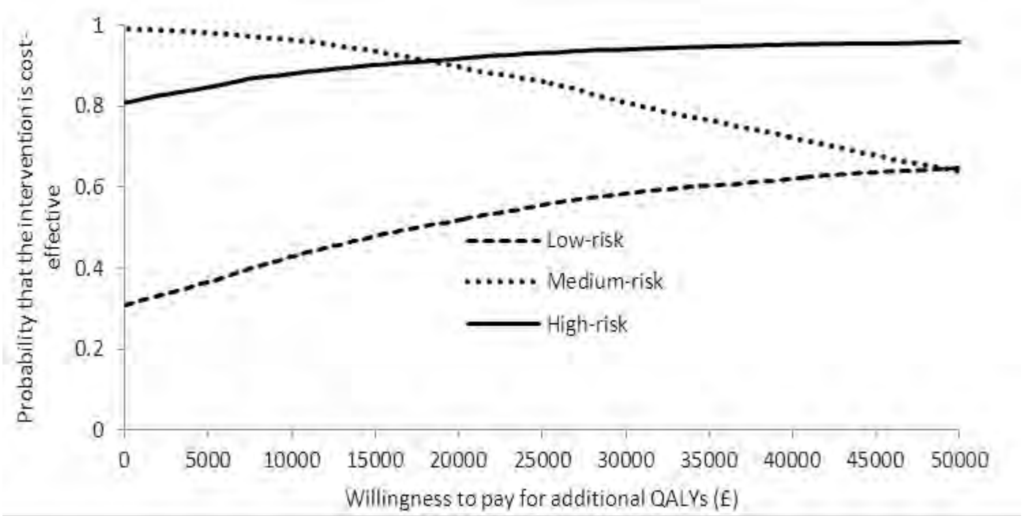


Figure 8:25 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (five occupation level friction period UoB)

8.7.12 Sensitivity Analysis

8.7.12.1 Regular approach: FCA for absenteeism and HCM for presenteeism excluding multiplier effects.

Overall, similar policy implications to the base-case were replicated in the sensitivity analyses. The CEAC for the incremental cost-utility ratios for a societal perspective with disaggregated friction period values for the overall analysis is shown in Figure 8.26. From a societal perspective, stratified care (phase 3) was associated with a very high probability of being cost-effective (over 90%) for the majority of willingness-to-pay thresholds investigated. At the £30,000 per QALY threshold, stratified care was highly cost-effective with probabilities exceeding 90%.

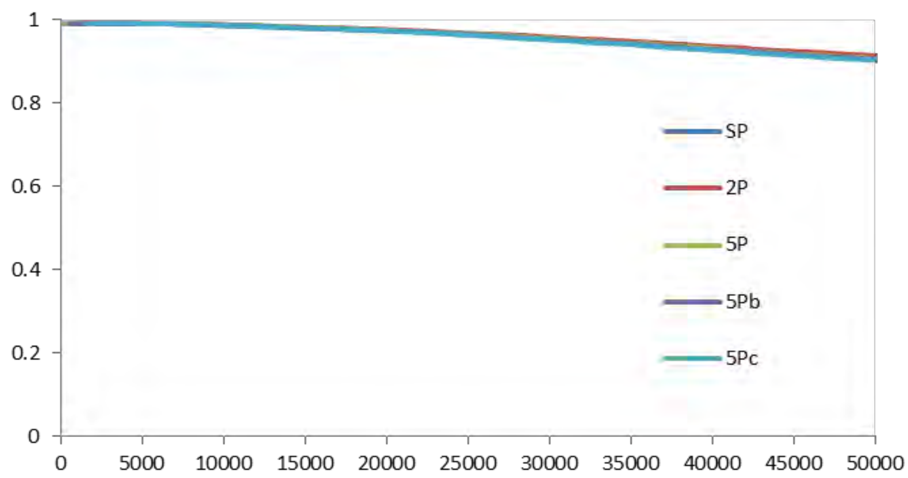


Figure 8:26 CEAC for the overall societal perspective sensitivity analysis

Sensitivity analysis results from the subgroup analysis are shown in Figure 8.27 with results presented only for the single-friction period. The findings were generally in the same direction as the base-case analysis. However, relatively lower probabilities for the phase 3 strategy being cost-effective were observed in the medium-risk group, and higher probabilities observed in the low-risk patient group, when the conventional friction cost approach was applied in all three risk-groups.

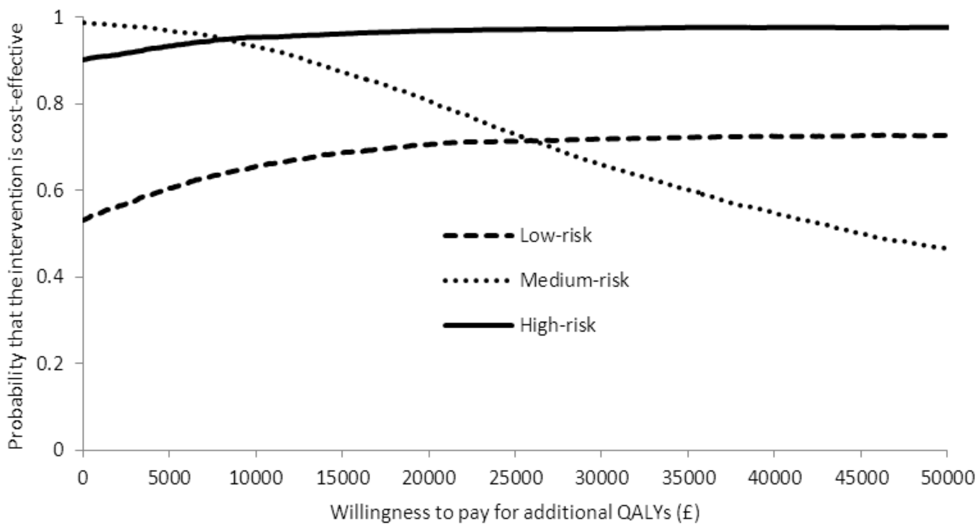


Figure 8:27 Cost-utility acceptability curves for the three risk group analysis of stratified care (phase 3) compared to usual care (phase 1) (a single occupation friction period)

8.7.13 Sensitivity analysis on cost-utility estimates

In the overall analysis, and for each risk-defined group, the findings were in the same direction as in the base-case analysis with similar interpretations of the cost-effectiveness outcomes replicated in all sensitivity analyses. As in the healthcare perspective, the phase 3 strategy remained cost-effective in the overall analysis and among patients in the high-risk group. The sensitivity analyses were based on calculating societal costs using the more conventional approach that values absenteeism using the FCA and presenteeism using the HCM (SA1), incorporating a higher elasticity factor of 1 assuming no short-term compensation mechanisms (SA2) and adjusting multiplier estimates in the suggested friction cost approach (SA3) (Table 8.17).

Table 8:17 Mean cost differences, mean QALY differences, incremental cost-effectiveness ratios from varying friction periods and corresponding three sensitivity analyses (SA1, SA2 and SA3)

	P1	P2	P3	P4	P5
Societal costs: Overall Analysis					
Main	-354	-380	-356	-348	-339
SA1	-383	-409	-378	-415	-415
SA2	-285	-298	-285	-281	-277
SA3	-366	-385	-366	-362	-357
Mean QALY difference			0.003		
ICER Main	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA1	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA2	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA3	Dominant	Dominant	Dominant	Dominant	Dominant
Low-risk category					
Main	63	63	63	63	63
SA1	96	96	96	96	96
SA2	-8	-8	-8	-8	-8
SA3	24	24	24	24	24
Mean QALY difference			0.00294		
ICER Main	21,429	21,429	21,429	21,429	21,429
ICER SA1	32,653	32,653	32,653	32,653	32,653
ICER SA2	-2742	-2742	-2742	-2742	-2742
ICER SA3	8,163	8,163	8,163	8,163	8,163
Medium-risk category					
Main	-573	-611	-588	-574	-553
SA1	-662	-700	-659	-715	-717
SA2	-372	-393	-385	-374	-364
SA3	-521	-549	-536	-523	-512
Mean QALY difference			-0.00795		
ICER Main	72,075	76,855	73,962	72,201	69,560
ICER SA1	83,270	88,050	82,893	90,189	90,189
ICER SA2	46,793	49,434	48,428	47,044	45,786
ICER SA3	65,535	69,057	67,421	65,786	64,403
High-risk category					
Main	-547	-595	-530	-541	-519
SA1	-534	-582	-517	-577	-577
SA2	-512	-536	-490	-493	-493
SA3	-656	-689	-631	-633	-632
Mean QALY difference			0.01956		
ICER Main	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA1	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA2	Dominant	Dominant	Dominant	Dominant	Dominant
ICER SA3	Dominant	Dominant	Dominant	Dominant	Dominant

Main estimates from the suggested approach using the FCA for valuing absenteeism and presenteeism with adjusted multiplier values, SA1 estimates from the conventional approach using the FCA for valuing absenteeism and the HCM for valuing presenteeism; SA2 using an elasticity factor value of 1assuming not short-term compensation costs; SA3 adjusting multiplier estimates in the suggested friction cost approach.

8.8 Discussion

In this chapter, two economic evaluation case studies were presented illustrating a practical application of the friction cost approach in a UK setting. Cost-effectiveness was estimated in a clinical trial and in routine primary care practice. The analysis used friction period estimates from Chapter Six and a single-item presenteeism question validated in Chapter Seven to measure presenteeism.

The current studies were the first to examine the cost-effectiveness of stratified care in primary care while incorporating absenteeism, presenteeism and multiplier effects alongside friction period estimates disaggregated by occupational categories. The findings were consistent in a clinical trial and routine practice setting both from the healthcare and societal perspectives. The cost-effectiveness result in both cases was favourable to the stratified care intervention in low back pain sufferers when compared to standard UK cost-effectiveness thresholds (Appleby et al., 2007; McCabe et al., 2008). The results demonstrate that the inclusion of productivity costs strengthened the case for stratified care in the overall analysis for both studies.

An important finding was that there were observed changes in cost-effectiveness results among subgroup populations when productivity costs and multiplier effects were incorporated. For example, the probability that stratified care was cost-effective among low-risk patients improved from 52% to 80% when productivity costs were incorporated within the SBT analysis at the £30,000 cost-effectiveness threshold. Stratified care was associated with improved probabilities (from 20% to 78%) among medium-risk patients as a result of including productivity costs in the IBS analysis. The stratified care model appears appropriate for a general primary care setting. When considering subgroups, the model appears to be more appropriate for high-risk patients in routine practice.

The results in this chapter provide insights into the implications of using more detailed friction periods disaggregated by SOC occupational categories. The results indicate that the use of stratified friction periods in this population did not, in these particular cost-effectiveness evaluations, alter the overall outcome of the base-case analysis.

To date, there is no evidence from economic evaluations that have simultaneously incorporated presenteeism, multiplier effects, and stratified friction period estimates in the UK. A previous review showed a scarcity of studies estimating productivity costs using the friction cost approach in comparison to the total number of economic evaluations, particularly in the UK (Pritchard and Sculpher, 2000). Although Pritchard and Sculpher's (2000) findings may have been due to the date of their review in relation to the development of the friction cost approach, similar findings were found in the systematic review within Chapter Four of this thesis. This found few economic evaluations in the UK applying the friction cost approach, and no studies that have attempted to estimate reduced productivity loss using the friction cost approach in the UK or that have considered multiplier effects resulting from absenteeism and presenteeism in other countries (Nicholson et al., 2006; Zhang et al., 2012). Applications of the friction cost approach remain important as the method potentially generates more realistic productivity costs than the conventionally applied human capital method (Koopmanschap and Rutten, 1996).

The decision to include or exclude productivity costs has been shown to significantly impact incremental costs in depressive disorders (Krol et al., 2011) and in neck pain interventions (Lewis et al., 2007). Although their inclusion in the evaluations in this chapter did not alter the overall cost-effectiveness decision, back pain has been associated with significant absence and disability costs (Ostelo et al., 2008). The exclusion of back pain productivity related costs could therefore generate underestimates of the true value of interventions providing treatment in this area. Additional empirical research on productivity costs in economic evaluations is important in light of the significant proportion of productivity costs in relation to total costs. The findings in both studies showed that productivity costs on average comprised more than half of the societal costs for stratified care among low back pain sufferers. And, their inclusion in the economic evaluation studies altered the probability level of the stratified care model being cost-effective at acceptable willingness-to-pay threshold values particularly within the subgroup analyses.

In the analyses presented here, cost-effectiveness acceptability curves were used to assess the cost-effectiveness of stratified care for increasing monetary valuations per QALY gained and assessed based on the £30,000 per QALY threshold. Although many cost-effectiveness analyses of healthcare interventions using the QALY outcome in the UK are based on this threshold value (Appleby et al., 2007; McCabe et al., 2008), there is currently no clear guidance on an acceptable threshold for an additional QALY gained in the UK when considering a societal perspective. Cost-effectiveness studies performed from a societal perspective provide further support for the hypothesis that inclusion of productivity costs may necessitate application of different ICER thresholds from the traditional £30,000 per

QALY threshold (Eichler et al., 2004). Some questions remain unanswered from recent research towards adjusting the threshold based on the social value for a QALY, and on opportunity costs of foregone health (Donaldson et al., 2011; Claxton et al., 2013). In future investigations, evidence is needed on whether and how to estimate this threshold when considering a societal perspective.

Methodologically, the results presented in this investigation provide further insights into the impact of including seldom considered productivity costs of presenteeism and multiplier effects within an economic evaluation setting in applications of the friction cost approach. There has been a lot of discussion about whether and how to value indirect costs of lost productivity in relation to presenteeism (Brooks et al., 2010; Johns, 2010). Several issues remain unresolved at present on this topic. Some have argued that these costs should be excluded, at least, or used with considerable caution because of criticisms on existing conversion and monetary translation approaches (Brooks et al., 2010). Given that both presenteeism and absenteeism costs are highly associated with the clinical area of back pain, however, excluding these costs could potentially lead to the generation of inaccurate estimates. In this thesis, the costs of presenteeism were therefore calculated with applications of the friction cost approach for the first time in the UK by applying the elasticity value of 0.8 on the calculated presenteeism costs.

Additionally, costs resulting from negative impacts of absenteeism and presenteeism on productivity of co-workers particularly where team work is involved (multiplier effects) were incorporated in the economic analyses presented in this thesis using multiplier values derived from a UK national study (Zhang et al., 2012). Empirical evidence on multiplier effects and their impact on overall societal economic evaluation outcomes remain scarce in UK studies. This exploration therefore increased knowledge on how multiplier effects from secondary sources can be used and their impact on productivity costs and cost-effectiveness outcomes. There is room for further progress in this area by determining a wider set of multiplier values as well as work related compensation mechanisms for UK data.

Previous studies have demonstrated that multiplier effects significantly impact on overall productivity costs (Zhang et al., 2012; Krol et al., 2012). In the literature, productivity costs have been estimated to be 28% higher than the regular friction cost approach estimates when multiplier effects are incorporated (Krol et al., 2012). The findings here agreed with previous research showing alterations in total societal costs as a result of incorporating multiplier effects. The cost-effectiveness analyses incorporating multiplier effects did not alter policy implications compared with analysis conducted from the healthcare perspective. However, their inclusion alongside absenteeism and presenteeism costs seemed to generate much uncertainty in the data among subgroup populations. The validity of the findings provided here, however, may require further empirical investigation since the multiplier values applied were obtained from a UK cohort of rheumatoid arthritis patients with very small sample sizes. Overall, the results provided suggest that current estimates of productivity costs in economic evaluation studies could be underestimated, in that multiplier effects are often not included.

On the other hand, it has been suggested that productivity loss is often partially compensated during working hours by colleagues and/or sick employees (Severens et al., 1998; Pauly et al., 2002; Krol et al., 2012; Zhang et al., 2012). In such cases, it has been argued that productivity loss from an absent individual or an individual working with limitations could potentially be reduced (Krol et al., 2012). It was not possible to incorporate such effects of work compensation in the cost-effectiveness studies presented in this thesis because data on compensation during working hours and extra workers were not collected. Although not investigated here, the estimation of job-related compensation mechanisms and how to translate such costs into productivity loss, and consequently incorporate their effects on cost-effectiveness outcomes, deserves further attention.

As part of a sensitivity analysis, the impact of key assumptions on the overall societal costs was assessed. All sensitivity analyses carried out generated similar results to the base-case analyses. For instance, the elasticity factor is a key parameter normally used in valuing productivity costs. It is common practice in applying the friction cost approach to use an elasticity value of 0.8 according to the Dutch costing guidelines as was done here. Although used by various studies, mostly in the Netherlands, there have not been any other reported published estimates of this elasticity factor. It therefore remains unclear how appropriate this value currently is, particularly within the UK context. The impact of changing the elasticity value on cost-effectiveness results was investigated in the sensitivity analysis, seeing there is no reported value in the UK. An elasticity value of 1 was applied in the sensitivity analysis. The results showed that, in this case, the overall policy implications were not changed as a result of altering the elasticity value. Similarly, in order to assess the impact of different methodological approaches to productivity cost estimates, cost-effectiveness estimates were recalculated using a possibly more conventional approach of incorporating productivity costs, i.e., using the friction cost approach for valuing absenteeism and the human capital method for valuing presenteeism and excluding multiplier effects. The results from both studies showed approximately similar policy implications to the regular friction cost approach of valuing productivity costs, suggesting the traditional valuation approach generates comparable policy implications in this case study.

Strengths of the economic analysis

The analysis presented here has some key strengths. The economic analyses were conducted alongside a large RCT and replicated in routine practice. Hence all resource use, work-related data and health outcomes data were obtained prospectively during the study periods. This is an important strength, as reliable resource use data with corresponding outcome data may not readily be available retrospectively. Another strength of this study is that the multiple imputation technique was used in dealing with missing data for back pain patients who dropped out before the end of the study period. This involved imputation of five samples of cost and effects which allowed the use of full costs and effects, for patients who did not complete the study. Adopting a complete-case analysis could have underestimated the real costs involved in the stratified and usual care interventions and among subgroups, as severely affected back pain patients are more likely to drop out of the study than less affected patients.

Few studies have provided evidence on the impact of simultaneously including absenteeism, presenteeism and multiplier effects alongside healthcare costs in economic evaluations. This research therefore contributes to this evidence base. Moreover, these are the first studies to explore application of friction periods stratified by occupational levels. Various cost-valuation models were used to generate societal costs for varying occupational level friction periods. This analysis therefore provides estimates that go beyond the more conventional friction cost approach of applying a single friction period. The results are also presented by risk-based subgroup populations. These are all important strengths of this research.

Limitations of this analysis

A limitation of the evaluations presented and discussed here is that the productivity cost analyses were limited to few participants who provided work loss related data. Moreover, the results presented here only focus on economic evaluations of low back pain sufferers, and it is a question as to how generalisable the implications of including productivity costs would be to other interventions for other conditions. The methodological approaches developed here should be explored further in other disease conditions for their impact to be more clearly understood.

Furthermore, inherent challenges were observed in the current data collection tools for estimating work related productivity loss. The lack of episode specific work related data during economic evaluation studies posed challenges in the practical application of the friction cost approach. The collection of such data requires the identification of individual sickness episode occurrences and duration data that were not available here. This would require detailed capture of time off work data from patients throughout the study period. Nevertheless, although criticised for using complicated data, which are not always available at the national level (Luce, 1996; Liljas, 1998; Knies et al., 2010), the friction cost approach provides more realistic estimates than the traditionally used human capital method. Also, the application of the friction cost approach in valuing productivity costs is steadily growing in the literature as demonstrated by the review studies reported in Chapter Three.

The SIPQ used to assess presenteeism in this investigation was first validated among LBP patients and then applied to assess reduced productivity, although estimates from other instruments could potentially lead to different results. Furthermore this instrument does not allow for capture of multi-dimensional work aspects or aspects of teamwork and substitutability. However, the SIPQ can potentially be used alongside other questionnaires that are able to measure such aspects, and provides the essential basis for presenteeism valuation within an economic evaluation.

The feasibility of applying the friction cost approach when incorporating more detailed friction period estimates is a necessary consideration in economic analysis. However, improvements in the tools used to obtain data on specific illness episodes and their duration within trial based studies is a necessary requirement for fostering increased application of the friction cost approach. Furthermore, regular updates and standardisation of friction period data sources are important, particularly at the national population level, as they are likely to vary depending on the state of the economy. Ideally, the increased availability of these data will provide more accurate estimates of friction cost related productivity costs.

Importantly, it was not possible to adjust productivity loss for compensation mechanisms due to the lack of appropriate work characteristic compensation data. Nevertheless, the importance of these work compensation costs has been clearly established as earlier discussed. For this investigation, it was assumed that the application of a 0.8 elasticity factor within the cost equation possibly incorporates workers compensating for short-term absences during normal working hours. However, additional analyses incorporating both multiplier effects and detailed compensation mechanisms estimates would be very informative and desirable. Furthermore, the multiplier values used here were based on values from an external study, and the validity of the approaches used in generating these estimates relies on the population for that investigation (defined by Rheumatoid Arthritis) being representative for low back pain sufferers.

8.9 Conclusion

This chapter presented and discussed cost-effectiveness results of the stratified care model in a clinical trial and in routine care practice while taking into consideration absenteeism, presenteeism and multiplier effects. The analyses concluded that the stratified care model was highly cost-effective in both settings. An interesting finding was the considerable change in probability levels for the stratified care model being cost-effective among subgroup populations when multiplier effects and other work related outcomes such as presenteeism and absenteeism are incorporated in the cost-effectiveness analysis. Applications of stratified friction periods using the friction cost approach, however, did not alter the overall fundamental interpretation of cost-effectiveness outcomes. This finding provides further support for future research in economic evaluations in other interventions and disease areas. In the next Chapter, the overall discussion, conclusions and recommendations of the thesis are provided.

CHAPTER NINE GENERAL DISCUSSION AND CONCLUSIONS

9.1 Introduction

The overall objective of the thesis was to contribute to the debate on the valuation of productivity costs using the friction cost approach and the inclusion of absence from work (absenteeism), reduced productivity (presenteeism) and wage-related multiplier effects in economic evaluation. To achieve this objective, the thesis addressed four research areas that relate to the estimation and inclusion of productivity costs and multiplier effects resulting from effects of productivity loss on team productivity within economic evaluations, in the United Kingdom. The research questions emerged from systematic reviews of cost-of-illness and economic evaluation studies that applied the friction cost approach and reported presenteeism costs. The searches were conducted in a systematic and rigorous way to obtain a diverse set of studies across different clinical areas. Methodological and empirical studies were then presented and discussed in Chapters Six and Seven of this thesis. The first study in Chapter Six focussed on estimating the length of friction period estimates for the UK setting. The second included a comparison study of sickness certification records with self-reported data. In Chapter Seven, the construct validity and responsiveness of the single-item presenteeism question as a standalone measure among low back pain sufferers was investigated. The empirical work from both chapters was then applied in Chapter Eight to investigate the potential impact on cost-effectiveness results of including full productivity costs in relation to absenteeism, presenteeism and multiplier effects on cost-effectiveness using the friction cost approach.

The aim of this chapter is to pull together the methodological and empirical work contained in this thesis. The first section provides an overview of the findings from the study questions addressed and compares the findings here to existing literature. The next section then considers the implications of the findings for policy and future research. Finally, an overall conclusion from the research is drawn.

9.2 Issues addressed in this thesis

How common is the application of the friction cost approach in valuing productivity costs in UK cost-of-illness and economic evaluation studies?

While there has been much debate about the benefits associated with the friction cost approach in valuing lost productivity, little evidence exists on whether and how the friction cost approach has been applied within cost-of-illness and economic evaluation studies. Chapter Three found that few attempts have been made at valuing productivity costs using this approach in the UK, but there is a general lack of clarity on the parameters and methods used. The friction cost approach is commonly applied in studies performed in the Netherlands, where the approach originated (Koopmanschap and van Ineveld, 1992; Koopmanschap et al., 1995), and where there appears to be more readily available country specific parameters necessary for its application. Nevertheless, variations in key parameters such as the friction period were observed across the studies from the different countries.

One of the key issues emerging from this review was that evidence of friction periods remains sparse in the UK, and no attempts have been made to estimate friction periods using occupational status, which is more relevant to economic evaluation. The findings from this systematic review further supported the idea of estimating country specific parameters for the friction cost approach, if increased application of this approach is to be realised. No attempts have been made to incorporate multiplier effects resulting from absenteeism and presenteeism in existing economic evaluation studies. The findings of the review conducted here supported earlier work showing that the friction cost approach needs to be given further consideration in the UK (Pritchard and Sculpher, 2000). The gaps in application of this approach informed part of the work in this thesis.

How has presenteeism been measured and valued in cost-of-illness and economic evaluation research?

A systematic review of economic studies assessing productivity costs in practice was reported in Chapter Four of this thesis, showing variations in standard and non-standardised presenteeism questionnaires. The findings of the review were in agreement with previous research, particularly Schultz et al., (2009) showing limited applications of presenteeism costs compared with other work-place outcomes such as absenteeism. Most of the assessment of presenteeism costs in current practice was found to be in cost-of-illness studies, with very few economic evaluations assessing presenteeism. On the costs of presenteeism in relation to absenteeism and overall total costs of illness, the systematic review here found presenteeism comprised the greatest proportion of total illness costs, being greater than absenteeism costs. Chapter Four showed that the effects of presenteeism on team productivity were not taken into account. Further, economic evaluation studies which take into account presenteeism costs and their effects on teamwork need to be undertaken. Only one study from the Netherlands, reported generating presenteeism costs using the friction cost approach, and not one in the area in the UK. It is possible therefore that attempts to apply the friction cost approach in generating presenteeism costs may lead to lower productivity loss estimates than current estimates based on the human capital method. Further research is needed to add to this evidence base. An important finding was that a number of studies using a global presenteeism question were identified. Little was known, however, about the validity of using the global item presenteeism questions.

The results were significant in at least three major aspects: i) cost-effectiveness results excluding full presenteeism costs may possibly fall short of true societal costs; ii) further empirical research is needed to use the friction cost approach in providing more realistic presenteeism cost estimates; and iii) no attempts have been made to estimate presenteeism using the friction cost approach in the UK or to consider multiplier effects resulting from the impact of presenteeism on teams. Further studies exploring this issue were recommended and, in particular it was noted that further work was necessary to establish the validity of the standalone global presenteeism question. These issues formed part of the foci of this thesis.

What is the appropriate length of a friction period for the UK?

Chapter Six provided the first UK specific length of friction period values which can be used when applying the friction cost approach in a UK setting. The overall friction period values obtained were comparable to previous friction period estimates reported in the literature review described in Chapter Three (For example, (Maniadakis and Gray, 2000; Luengo-Fernandez et al., 2006; Rivero-Arias et al., 2010)) , but gave an empirically based UK specific value.

Further, these data were also used to generate friction period values stratified by SOC occupational level categories. No such estimates have been generated before in the UK. However, a previous study by Koopmanschap et al., (1995) estimated friction period values that were disaggregated according to educational levels in the Netherlands. The observed friction period estimates in this thesis varied according to different occupational level categories. The friction period was longer for higher occupational levels suggesting higher unemployment levels in lower occupational jobs, and a longer period needed for hiring higher occupational level employees. These findings contribute to better understanding of the implications of varying friction periods when applying the friction cost approach in economic evaluations. Further, empirical research was recommended to investigate the impact of using stratified friction period values to avoid an inaccurate estimation of productivity costs. A practical illustration of the impact of the estimates generated here within a cost-effectiveness framework is provided in Chapter Eight of the thesis.

Can sickness certification records be used as a proxy for self-reported work absence data?

Several methods are available for collecting productivity loss, the most popular in use have been self-reported questionnaires and administrative records (Pole et al., 2006; Evans, 2004). A strong relationship between self-reported sickness absence records and administrative database records has been reported in the literature (Ferrie et al., 2005; Severens et al., 2000). Administrative database records have limitations in data accessibility, as multiple providers have to be approached for each patient, while self-reported questionnaires are susceptible to recall bias. GP sickness certification records may provide a rich source of sickness absence data; however, the relationship between these records and self-reported records has not been firmly established.

The study carried out in Chapter Six of this thesis found poor comparability between self-reported sickness absence and GP sickness certification records. Furthermore, the findings of this study suggest that using sickness certification records to generate time-off work data is not a viable option in economic evaluation. The results from the current research, which considered periods of 6 months and 12 months, differed from a previous study that found good comparability between the two data sources over a period of two weeks (Wynne-Jones et al., 2008). The results were however in agreement with the study by Wynne-Jones et al., (2008) study that showed a lack of work absence duration data from GP certification records. As a result it was not possible to analyse the comparability of GP sickness certification duration data with self-reported work absence duration data. These results provide further support for continued improvements in self-report sickness absence questionnaires used to estimate work absence in economic evaluation studies. In their current state, sickness certification records might have a limited role in estimating time off work for economic evaluation research. However, more research on this topic is needed before the association between self-report absence records and GP certification records is more clearly understood. The identification of productivity loss is a necessary component in economic evaluations performed from a societal perspective. Until GP certification data improves, this study shows self-reported questionnaires offer the best possibility for deriving work absence data in economic evaluation studies. However, there is a need for improved methods of capturing duration of absence data relevant for the friction cost approach when using self-reported instruments.

Is the Single-Item Presenteeism Question (SIPQ) a valid and responsive measure among low back pain patients?

In Chapter Four, the review found that presenteeism costs have rarely been incorporated in economic evaluation studies. To some extent, this could be explained by the majority of existing measurement instruments that are sometimes lengthy and complex, with several disparate questions compared with a single-item global presenteeism question. As discussed in Chapter Four, commonly used multi-item measures such as the WLQ (Lerner et al., 2001), HPQ (Kessler et al., 2003), WPAI (Reilly et al., 1993), and SPS (Turpin et al., 2004) differ in the type of presenteeism questions asked and in their detail. The validity of these measures has previously been investigated and reported (Lofland et al., 2004; Prasad et al., 2004). Single-item global questions have practical advantages for researchers in terms of simplicity, cost and the possibility of generating presenteeism estimates in diverse occupations that could be used in cost-effectiveness analysis (Bowling, 2005). However, the validity and responsiveness of a single-item presenteeism question has not been fully investigated.

In Chapter Seven of this thesis, evidence is provided showing the construct validity and responsiveness of a single-item presenteeism question (SIPQ) among LBP patients from the two thesis datasets. As such, it appears that the SIPQ work loss outcomes are assessing constructs that are relevant to patients with low back pain. The validity and responsiveness of the single-item presenteeism question has not been fully established against commonly used multi-item presenteeism questionnaires. However, a previous research study investigated the agreement between a multi-item Work Activity Index (WAI) and single-item question on work ability among women on long-term sick leave, and found a strong association between the two measures for all participants (Ahlstrom et al., 2010). There is, therefore, room for further progress in comparing standard presenteeism questionnaires against non-standardised presenteeism global questions. The use of a single-item presenteeism question to estimate presenteeism costs could be preferable to ignoring these costs due to theoretical and practical limitations of existing multi-item measures.

How does the inclusion of productivity costs using the alternative valuation methods explored in this thesis impact on the results of the cost-effectiveness analyses?

Chapter Eight provides the first estimates of the cost-effectiveness of stratified care compared to usual care in LBP sufferers from a societal perspective using the friction cost approach, and incorporating presenteeism and multiplier aspects of presenteeism on team productivity. The economic analysis valued productivity costs using more-detailed friction periods stratified by SOC occupational level categories. The overview of the clinical area in Chapter Five of this thesis showed the significant burden associated with the back pain condition in comparison to other illnesses in the UK.

Perhaps contrary to expectations, the inclusion of productivity costs in the evaluations in this thesis did not alter the findings on the cost-effectiveness of these strategies when evaluated from a healthcare perspective (Hill et al., 2011; Whitehurst et al., 2012). The inclusion of detailed productivity costs strengthened the case for stratified care management, as the intervention remained highly cost-effective from both a societal and healthcare perspective. An interesting finding was that the incorporation of productivity costs and multiplier effects altered the probability of the stratified care model being cost-effective among subgroup populations of low back pain patients.

The use of more detailed friction periods stratified by SOC occupational level also did not alter the interpretation of the cost-effectiveness results. Similar results were observed in all sensitivity analyses performed. However, more research needs to be undertaken to investigate the impact of simultaneously including compensation mechanisms and multiplier effects alongside other productivity cost outcomes. The present studies were successful in illustrating methodological applications of the friction cost approach in the UK. This remains an important area for future research.

9.3 Comparison with other contributions

Pritchard and Sculpher (2000) reviewed the way absenteeism costs have been assessed in practice within economic evaluations. They found very few studies that have applied the friction cost approach in the UK and internationally, and recommended the need for explicit and detailed descriptions of the methods when this approach is used. Furthermore, their review revealed the need for additional empirical work around testing applications of presenteeism within economic evaluations.

The more up-to-date review conducted here, confirmed that applications of the friction cost approach remain limited in the United Kingdom. Chapter Six took further steps in generating detailed friction period estimates relevant for the UK and in Chapter Eight, an illustration of the practical application of the friction cost approach in economic evaluations was provided in a UK context, providing an explicit and detailed description of methods, as recommend by Pritchard and Sculpher (2000).

An overview of research studies that have attempted to investigate vacancy durations in the United Kingdom and in the Netherlands was also provided in Chapter Six of this thesis. Previously published research on vacancy durations, particularly the work of Van Ours and Ridder (1991), which modelled vacancy durations according to educational levels in the Netherlands were reviewed. Their work was extended by Koopmanschap and colleagues (1995) to incorporate data from a large number of Dutch firms. The authors reported their friction periods stratified by educational status. In the United Kingdom, four research studies were identified (Beaumont, 1978; Roper, 1988; Adams et al., 2002; Andrews et al., 2008) , but these do not provide vacancy duration information relevant for estimating a friction period that can be used for valuing productivity loss using the friction cost approach. Chapter Six, provides more recent estimates of detailed specific vacancy durations relevant for a UK context. The friction period estimates generated are comparable to previous estimates used in UK economic costing studies, for example, in Maniadakis and Gray (2000), and in Rivero-Arias et al (2010). However, the work in this thesis is extended to include stratified friction periods as recommended by Koopmanschap et al (1995).

Both Zhang and colleagues (2012) and Krol and colleagues (2012) have recently demonstrated the additional productivity costs associated with absenteeism and presenteeism multiplier effects and compensation mechanisms in rheumatoid arthritis patients from the United Kingdom and in smokers from the Netherlands respectively. The productivity costs incorporating multiplier effects in cost-effectiveness studies within this chapter corroborate their findings showing significant alterations in productivity costs when these costs are included. Further work is done by assessing the impact of including multiplier effects in productivity costs alongside absenteeism and presenteeism costs on the cost-effectiveness analyses reported in Chapter Eight of this thesis. This was done by adapting the multiplier values generated from the study by Zhang et al., (2012) and mapping them as closely as possible to the occupations in the study datasets for valuation purposes. It was however, not possible to assess the impact of simultaneously including multiplier effects and compensation mechanisms as work compensation data was not collected within the thesis datasets.

9.4 Strengths of the research

One of the major strengths of this thesis is that it assessed the cost-effectiveness from a societal approach of a stratified care model in a large randomised clinical trial and evaluated the same model in routine primary care practice, to illustrate the inclusion of productivity costs. Further, comprehensive sensitivity analyses were conducted alongside the two economic evaluation studies.

The empirical estimation of friction periods for the UK is the first to generate more detailed friction period estimates disaggregated by occupation, and will provide researchers with practical data for applying the friction cost approach. A comprehensive illustration of methodological applications of the friction cost approach in the UK is also provided.

Another strength is that the choice of empirical work was based on a comprehensive and rigorous review of literature on the assessment of productivity cost estimation methods in cost-of-illness and economic evaluations in current UK and international practice and an overview of the clinical area of low back pain. The lack of appropriate context specific data for applications of the FCA, highlighted from the review, has provided support for estimation of friction period estimates for the UK. Similarly, the review highlighted the need for validation of the global standalone presenteeism questions often used outside standard presenteeism questionnaires.

9.5 Limitations of the research

Although this thesis has addressed some important issues relating to the estimation of productivity costs and guidance on how to apply these alternative methods in economic evaluations, some limitations are recognised. Many of these provide important grounds for further research. Firstly, it is possible that the systematic reviews in Chapters Three and Four could have excluded some relevant studies in which the friction cost approach was used, as well as studies reporting presenteeism costs because of the search strategy employed. The decision to use only original economic costing studies and exclude model-based CEA can be limiting. However the aim was to assess how these approaches are used in primary costing studies and a comprehensive systematic literature and hand search was carried out to identify all relevant studies.

The research presented in Chapter Six focussed on estimation of a friction period for the UK and a comparability study of GP sickness certification records and self-report data. A limitation with the data on friction periods is that it was obtained from a variety of data sources with plausible varying degrees of validity, although relatively comparable estimates were generated. Nevertheless, a more comprehensive survey for the various SOC levels could provide more rigorous friction period estimates for the UK. Moreover more detailed explorations were based on broad SOC occupational classifications. Estimations of SOC level friction periods have not been investigated previously in the UK and it was not possible to make any comparisons with estimates based on other approaches.

The exploration of the validity and responsiveness of the SIPQ was limited to tests against health-related and socio-demographic related constructs. It would have been useful to have been able to obtain data that included a standard reduced productivity at work instrument for validation purposes. Nevertheless, previous validations have made similar comparisons (Wahlqvist et al., 2002; Reilly et al., 2008; Reilly et al., 2010).

A further limitation with the economic evaluation studies is practical. A lack of duration data led to the pragmatic decision to assume the number of days lost over the study period was representative of the total friction period between the measurement points. The impact of applying the friction cost approach using these estimates in these case studies therefore could be a limited representation of the true impact of this approach, especially in relation to absences with short absence durations. This suggests that economists using this approach need to collect more detailed data on absence duration, as well as total days of absence.

Further, the economic evaluation studies in this thesis did not include unpaid labour as part of the assessment of productivity costs. It could be argued that the exclusion of costs outside the labour market could also lead to an underestimation of productivity loss estimates and potentially lead to the selection of interventions discriminating against those not employed (including those caring for dependents, doing voluntary work, retired, students and the unemployed) in this clinical area. Unpaid labour costs are commonly measured using questionnaires and valued using the proxy good method or opportunity costs method (van den Berg et al., 2006). With the opportunity costs method, costs are valued at the rate at which individuals would be paid if they were in paid employment. In the proxy good method, costs are valued at rates of paid work closely related to the unpaid person's work (Knies et al., 2010). Although there is justification for attaching monetary valuations to non-market labour and informal labour costs, it was not possible to include these costs here since lost time related to unpaid labour was not collected in the study datasets. Further research should attempt also to explore the importance of these costs.

Furthermore, considering economic evaluations in the UK are currently based on a form of the extra-welfarist framework, that does not provide for inclusion of productivity costs, little attention has been paid to the generation of explicit guidance concerning when productivity costs should be considered by decision makers. Although this thesis has discussed and presented normative and theoretical support for considering a societal perspective, current recommendations from economic guidelines in the UK do not provide for inclusion of such costs in both the reference case and non-reference case analyses. The question of whether and how the current £20,000 to £30,000 per QALY threshold (Appleby et al., 2009) value should be adjusted in the case of a societal perspective is an important question, but was beyond the scope of this thesis.

9.6 Implications for future research

The results of the reviews in Chapter Three and Four indicate that current practice in economic evaluations within the UK could benefit from more empirical studies investigating applications of the friction cost approach. Several areas of research in productivity cost valuation using the friction cost approach in the UK remain unexplored.

Overall, further economic evaluation studies that value productivity costs using the friction cost approach, need to be undertaken. On a methodological note, when conducting economic evaluations, the choice of the valuation method should be based, where possible, on both the human capital method and friction cost approach for comparability purposes.

Another important finding is the significant cost associated with absenteeism and presenteeism related multiplier effects. Currently, these costs seem to be largely ignored in economic evaluations. More empirical research is needed to investigate the potential influence of multiplier effects and compensation mechanisms alongside common work related outcomes of absenteeism and presenteeism on overall cost-effectiveness results, using more case studies on other disease areas. This would provide more understanding and recommendation on best practice approaches of including these costs in economic evaluations.

More importantly, current estimates of short-term work compensation mechanisms within the UK are generally based on an elasticity value obtained from the Netherlands. The value of an appropriate elasticity value for UK and the existence of internal replacement mechanisms in firms remains unexplored. Data on unpaid labour were also excluded in the analyses presented here. It would be interesting to investigate how to measure unpaid productivity loss and incorporate costs of unpaid labour in economic analysis.

In order to facilitate more applications of the friction cost approach in the UK, further developments in measuring relevant work absence data are encouraged in this area. For example, empirical research on data collection instruments with the best approach for measuring absenteeism and presenteeism data relevant for the friction cost approach is needed. As a starting framework for future research on this, exploration could include comparisons of questions asking respondents to recall self-reported specific sickness episodes with selecting a range of absence periods from a questionnaire. Similarly, questionnaire approaches could be compared against the more comprehensive diary approaches. However, as a start, the findings from this thesis provided some general guidance as to what data collection instruments need to include for accurate application of the friction cost approach. These are listed in Box 9.1.

Box 9.1 Common items needed for the friction cost approach

- Individual spells of absenteeism and presenteeism
- Start and end dates of each spell
- Categorise absenteeism data into periods of absence shorter or longer than the friction period.
- Incorporate additional questions where relevant to obtain information on work characteristics and compensation mechanisms
- When calculating productivity costs, absences longer than the friction period should be considered as representative of the total productivity loss during that period.
- Absences shorter than the friction period should be considered as representative of productivity loss subject to an elasticity factor (often a value of 0.8 is applied)
- The friction period should incorporate time until vacancy, time to fill vacancy, and time till first work day

The SIPQ tool provides an easy, less burdensome and less costly option of generating presenteeism estimates, than the more commonly applied complex and lengthy questionnaires. However, further research is encouraged to enable more definitive recommendations of such an approach. A cross validation study against standardised instruments is recommended as this would provide comparable data on productivity loss and costs estimates from these tools, and provide more credibility to estimates resulting from single-item presenteeism questions.

9.7 Implications for policy

The comprehensive length of friction period estimates generated from this thesis are relevant to the economic evaluation framework in the UK as they provide practical data that can be used to apply the friction cost approach in a UK context. Further research should also be conducted to assess the relative impact of using these stratified friction period values on cost-effectiveness outcomes in other interventions and disease areas.

On an empirical note, the results showing the validity and responsiveness of the SIPQ indicate important results in relation to measuring presenteeism. The results suggest that using a single-item question on assessing presenteeism in primary care may be a good alternative to the commonly used multi-item instruments among low back pain sufferers. The benefits of single-item questions have been discussed by researchers such as Bowling (2005) and Ahlstrom et al., (2010), particularly when there is less time, and where ease of interpretation, reducing responder burden and cost-effectiveness are important.

The two cost-utility analyses from this research strengthen the case for stratified care analysis and sub grouping both from a healthcare and societal perspective as a contribution to policy making. Based on the findings here, stratified care should be recommended in a primary care setting and among high-risk patients within routine care. The two CEA studies showed that the stratified care model is highly cost-effective.

9.8 Conclusion

Research on application of the friction cost approach in the UK is evolving, compared with the human capital method. There is merit for further insight and research on this topic because productivity costs need to be calculated as realistically as possible in order to guide decision makers in broader healthcare decision making. This thesis, has contributed methodological and empirical knowledge on estimation of productivity costs, particularly using the friction cost approach. This was done by providing first estimates of stratified friction periods by occupation for the UK, and cost-effectiveness analyses incorporating the relatively new concepts of presenteeism and multiplier effects alongside healthcare costs and absenteeism. The findings can guide future application of the friction cost approach in the UK. A simple and valid single-item presenteeism question that can be used across diverse occupations to generate cost data in economic evaluations is presented in this thesis. Further, the comparability study in Chapter Six shows self-reported data provides more complete time off work data for economic evaluation than sickness certification data. However, self-reported data can be further modified to ensure they capture specific work absence duration episodes when using the friction cost approach.

In summary, the discussions provided in this thesis highlight a merit for further methodological and empirical research on the role of productivity costs within the United Kingdom and internationally. The findings in the thesis contribute to increased understanding about how the friction cost approach can be used to estimate productivity costs while stratifying segments of the labour market by occupation. In particular, it provides insight to how absenteeism, presenteeism and multiplier effects of team work can be incorporated into cost-effectiveness analysis. A simple way to measure and value presenteeism in comparison to the existing lengthy and complex multi-item instruments is also presented. Added to this is the length of friction period estimates that can be used when applying the friction cost approach in economic evaluation within a UK context. A lack of data can limit application of these approaches in a UK context, as was observed in the research here.

Overall, the research provided in this thesis contributes to methods for assessing productivity costs in economic evaluation, and illustrates the feasibility of using them in the UK. Given the significance of productivity costs, this thesis adopted an approach that further examined the effect of using more detailed friction periods (in line with the varying labour market situation) on overall cost-effectiveness outcomes, rather than a standard friction period as often reported in economic evaluation studies. Consequently, the thesis presents a more comprehensive and accurate estimation approach of cost-effectiveness outcomes of interventions in this clinical area, and therefore provides a starting point for further research and also for improving practice in the UK. Further testing of these approaches in other clinical areas is merited in the UK, considering the growing evidence and value for accurate estimates of productivity costs.

Appendices

Contents

Appendix A.1 Chapter Three systematic review search strategies

Database: Embase

Search term

#1	friction cost approach.mp.
#2	limit #1 to (english language)
#3	friction cost method.mp.
#4	limit #3 to (english language)
#5	friction period.mp.
#6	limit #5 to (english language)
#7	friction cost.mp.
#8	limit #7 to (english language)
#9	#2 or #4 or #6 or #8

Database(s): Ovid MEDLINE(R)

Search term

#1	friction cost approach.mp.
#2	limit #1 to (english language)
#3	friction cost method.mp.
#4	limit #3 to (english language)
#5	friction period.mp.
#6	limit #5 to (english language)
#7	friction cost.mp.
#8	limit #7 to (english language)
#9	#2 or #4 or #6 or #8

Database(s): PsycINFO

Search term

#1	friction cost approach.mp.
#2	limit #1 to (english language)
#3	friction cost method.mp.
#4	limit #3 to (english language)
#5	friction period.mp.
#6	limit #5 to (english language)
#7	friction cost.mp.
#8	limit #7 to (english language)
#9	#2 or #4 or #6 or #8

Database: NHSEED

(friction cost approach) OR (friction cost method) OR (friction period) OR (friction cost)

Appendix A.2 Chapter Three references for individual studies identified

1. Borghouts JAJ, Koes BW, Vondeling H, et al. (1999) Cost-of-illness of neck pain in The Netherlands in 1996. *Pain* 80: 629-636.
2. Bosmans JE, van Schaik DJ, Heymans MW, et al. (2007) Cost-effectiveness of interpersonal psychotherapy for elderly primary care patients with major depression. *Int J Technol Assess Healthcare* 23: 480-487.
3. Brouwers EP, de Bruijne MC, Terluin B, et al. (2007) Cost-effectiveness of an activating intervention by social workers for patients with minor mental disorders on sick leave: a randomized controlled trial. *Eur J Public Health* 17: 214-220.
4. Brunenberg DE, Steyn MJv, Sluimer JC, et al. (2005) Joint Recovery Programme versus Usual Care: An Economic Evaluation of a Clinical Pathway for Joint Replacement Surgery. *Medical Care* 43: 1018-1026.
5. De Bruijn C, Goossens M, de Bie R, et al. (2007) Cost-effectiveness of an education and activation program for patients with acute and subacute shoulder complaints compared to usual care. *Int J Technol Assess Healthcare* 23: 80-88.
6. Dirksen CD, Ament AJHA, Adang EMM, et al. (1998) Cost-Effectiveness of Open Versus Laparoscopic Repair for Primary Inguinal Hernia. *International Journal of Technology Assessment in Healthcare* 14: 472-483.
7. Fautrel B, Clarke AE, Guillemin F, et al. (2007) Costs of rheumatoid arthritis: New estimates from the human capital method and comparison to the willingness-to-pay method. *Medical Decision Making* 27: 138-150.
8. Gallefoss F and Bakke PS. (2002) Cost-benefit and cost-effectiveness analysis of self-management in patients with COPD--a 1-year follow-up randomized, controlled trial. *Respir Med* 96: 424-431.
9. Goossens ME, Rutten-Van Molken MP, Kole-Snijders AM, et al. (1998) Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. *Health Econ* 7: 39-51.
10. Hakkaart-van Roijen L, Hoeijenbos MB, Regeer EJ, et al. (2004) The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta Psychiatrica Scandinavica* 110: 383-392.
11. Hanly P, Timmons A, Walsh PM, et al. (2012) Breast and prostate cancer productivity costs: a comparison of the human capital method and the friction cost approach. *Value in Health* 15: 429-436.
12. Huscher D, Merkesdal S, Thiele K, et al. (2006) Cost of illness in rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and systemic lupus erythematosus in Germany. *Annals of the Rheumatic Diseases* 65: 1175-1183.

13. Hutubessy RC, van Tulder MW, Vondeling H, et al. (1999) Indirect costs of back pain in the Netherlands: a comparison of the human capital method with the friction cost approach. *Pain* 80: 201-207.
14. Jellema P, van der Roer N, van der Windt DA, et al. (2007) Low back pain in general practice: cost-effectiveness of a minimal psychosocial intervention versus usual care. *Eur Spine J* 16: 1812-1821.
15. Kaitelidou D, Ziroyanis PN, Maniadakis N, et al. (2005) Economic evaluation of hemodialysis: implications for technology assessment in Greece. *International Journal of Technology Assessment in Healthcare* 21: 40-46.
16. Korthals-de Bos IB, Hoving JL, van Tulder MW, et al. (2003) Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. *BMJ* 326: 911.
17. Lewis M, James M, Stokes E, et al. (2007) An economic evaluation of three physiotherapy treatments for non-specific neck disorders alongside a randomized trial. *Rheumatology (Oxford)* 46: 1701-1708.
18. Liem MS, Halsema JA, van der Graaf Y, et al. (1997) Cost-effectiveness of extraperitoneal laparoscopic inguinal hernia repair: a randomized comparison with conventional herniorrhaphy. Coala trial group. *Ann Surg* 226: 668-675; discussion 675-666.
19. Liu JLY, Maniadakis N, Gray A, et al. (2002) The economic burden of coronary heart disease in the UK. *Heart* 88: 597-603.
20. Luengo-Fernandez R, Leal J, Gray A, et al. (2006) Cost of cardiovascular diseases in the United Kingdom. *Heart* 92: 1384 - 1389.
21. Luijsterburg PA, Lamers LM, Verhagen AP, et al. (2007) Cost-effectiveness of physical therapy and general practitioner care for sciatica. *Spine (Phila Pa 1976)* 32: 1942-1948.
22. Maniadakis N and Gray A. (2000) The economic burden of back pain in the UK. *Pain* 84: 95-103.
23. McEachan R, Lawton R, Jackson C, et al. (2011) Testing a workplace physical activity intervention: a cluster randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity* 8: 29.
24. Mol BW, Hajenius PJ, Engelsbel S, et al. (1999) Treatment of tubal pregnancy in the netherlands: an economic comparison of systemic methotrexate administration and laparoscopic salpingostomy. *Am J Obstet Gynecol* 181: 945-951.
25. Neovius K, Rehnberg C, Rasmussen F, et al. (2012) Lifetime productivity losses associated with obesity status in early adulthood: a population-based study of Swedish men. *Appl Health Econ Health Policy* 10: 309-317.
26. Nikken JJ, Oei EH, Ginai AZ, et al. (2005) Acute peripheral joint injury: cost and effectiveness of low-field-strength MR imaging--results of randomized controlled trial. *Radiology* 236: 958-967.
27. Oliva J, Lobo F, Lopez-Bastida J, et al. (2005) Indirect costs of cervical and breast cancers in Spain. *Eur J Health Econom* 6: 309-313.
28. Poley MJ, Stolk EA, Langemeijer RA, et al. (2001) The cost-effectiveness of neonatal surgery and subsequent treatment for congenital anorectal malformations. *J Pediatr Surg* 36: 1471-1478.

29. Ponto KA, Merkesdal S, Hommel G, et al. (2013) Public health relevance of Graves' orbitopathy. *J Clin Endocrinol Metab* 98: 145-152.
30. Rivero-Arias O, Gray A and Wolstenholme J. (2010) Burden of disease and costs of aneurysmal subarachnoid haemorrhage (aSAH) in the United Kingdom. *Cost effectiveness and resource allocation* 8: 6.
31. Rutten-van Molken MP, van Nooten FE, Lindemann M, et al. (2007) A 1-year prospective cost-effectiveness analysis of roflumilast for the treatment of patients with severe chronic obstructive pulmonary disease. *Pharmacoeconomics* 25: 695-711.
32. Saka m, McGuire A and Wolfe C. (2009) Cost of stroke in the United Kingdom. *Age and Ageing* 38: 27-32.
33. Smit F, Willemse G, Koopmanschap M, et al. (2006) Cost-effectiveness of preventing depression in primary care patients: randomised trial. *Br J Psychiatry* 188: 330-336.
34. Soegaard R, Buenger CE, Christiansen T, et al. (2007) Circumferential fusion is dominant over posterolateral fusion in a long-term perspective - Cost-utility evaluation of a randomized controlled trial in severe, chronic low back pain. *Spine* 32: 2405-2414.
35. Stant AD, Ten Vergert EM, den Boer PC, et al. (2008) Cost-effectiveness of cognitive self-therapy in patients with depression and anxiety disorders. *Acta Psychiatr Scand* 117: 57-66.
36. Steenstra IA, Anema JR, van Tulder MW, et al. (2006) Economic evaluation of a multi-stage return to work program for workers on sick-leave due to low back pain. *J Occup Rehabil* 16: 557-578.
37. Steuten L, Vrijhoef B, Merode FV, et al. (2006) Evaluation of a regional disease management programme for patients with asthma or chronic obstructive pulmonary disease. *International Journal for Quality in Healthcare* 18: 429-436.
38. Steuten LMG, Bruijsten M and Vrijhoef HJM. (2007) Economic evaluation of a diabetes disease management programme with a central role for the diabetes nurse specialist. *European Diabetes Nursing* 4: 64-71.
39. van den Hout WB, Goekoop-Ruiterman YPM, Allaart CF, et al. (2009) Cost-Utility Analysis of Treatment Strategies in Patients With Recent-Onset Rheumatoid Arthritis. *Arthritis & Rheumatism-Arthritis Care & Research* 61: 291-299.
40. van der Roer N, van Tulder M, van Mechelen W, et al. (2008) Economic evaluation of an intensive group training protocol compared with usual care physiotherapy in patients with chronic low back pain. *Spine (Phila Pa 1976)* 33: 445-451.
41. van Eijsden MD, Gerhards SA, de Bie RA, et al. (2009) Cost-effectiveness of postural exercise therapy versus physiotherapy in computer screen-workers with early non-specific work-related upper limb disorders (WRULD); a randomized controlled trial. *Trials [Electronic Resource]* 10: 103.
42. van Enckevort PJ, Koopmanschap MA, Tenvergert EM, et al. (1997) Lifetime costs of lung transplantation: estimation of incremental costs. *Health Econ* 6: 479-489.
43. van Roijen LH, van Straten A, Al M, et al. (2006) Cost-utility of brief psychological treatment for depression and anxiety. *Br J Psychiatry* 188: 323-329.
44. Van Schayck CP, Kaper J, Wagena EJ, et al. (2009) The cost-effectiveness of antidepressants for smoking cessation in chronic obstructive pulmonary disease (COPD) patients. *Addiction* 104: 2110-2117.

45. Van Tubergen A, Boonen A, Landewe R, et al. (2002) Cost effectiveness of combined spa-exercise therapy in ankylosing spondylitis: a randomized controlled trial. *Arthritis Rheum* 47: 459-467.
46. Vijgen SM, Koopmans CM, Opmeer BC, et al. (2010) An economic analysis of induction of labour and expectant monitoring in women with gestational hypertension or pre-eclampsia at term (HYPITAT trial). *BJOG* 117: 1577-1585.

Appendix B.1 Chapter Four systematic review search strategies

Database(s): Ovid MEDLINE(R)

Search term

#1	presenteeism.mp.
#2	lost productivity.mp.
#3	reduced productivity.mp.
#4	work limitations.mp.
#5	productivity costs.mp.
#6	work performance.mp.
#7	#1 or #2 or #3 or #4 or #5 or #6
#8	Cost-Benefit Analysis/ or cost-utility analysis.mp.
#9	"Costs and Cost Analysis"/ or cost-effectiveness analysis.mp.
#10	#8 or #9
#11	#7 and #10

Database(s): PsycINFO

Search term

#1	presenteeism.mp.
#2	limit #1 to english language
#3	lost productivity.mp.
#4	limit #3 to english language
#5	reduced productivity.mp.
#6	limit #5 to english language
#7	work performance.mp.
#8	limit #7 to english language
#9	exp "Costs and Cost Analysis"/ or cost-utility analysis.mp.
#10	limit #9 to english language
#11	exp "Costs and Cost Analysis"/
#12	limit #11 to english language
#13	#2 or #4 or #6 or #8
#14	#10 or #12
#15	#13 and #14

Appendix B.2 Chapter Four references for individual studies identified

1. Boonen A, Brinkhuizen T, Landewe R, et al. (2010) Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. *Annals of the Rheumatic Diseases* 69: 1123-1128.
2. Braakman-Jansen LMA, Taal E, Kuper IH, et al. (2012) Productivity loss due to absenteeism and presenteeism by different instruments in patients with RA and subjects without RA. *Rheumatology* 51: 354-361.
3. Burton WN, Chen C-Y, Conti DJ, et al. (2005) The association of health risks with on-the-job productivity. *Journal of Occupational & Environmental Medicine* 47: 769-777.
4. Burton WN, Conti DJ, Chen C-Y, et al. (2002) The economic burden of lost productivity due to migraine headache: a specific worksite analysis. *Journal of Occupational & Environmental Medicine* 44: 523-529.
5. Cisternas MG, Blanc PD, Yen IH, et al. (2003) A comprehensive study of the direct and indirect costs of adult asthma. *J Allergy Clin Immunol* 111: 1212-1218.
6. Collins JJ, Baase CM, Sharda CE, et al. (2005) The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *Journal of Occupational & Environmental Medicine* 47: 547-557.
7. Daley M, Morin CM, LeBlanc M, et al. (2009) The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep* 32: 55-64.
8. Finkelstein EA, DiBonaventura Md, Burgess SM, et al. (2010) The costs of obesity in the workplace. *Journal of Occupational & Environmental Medicine* 52: 971-976.
9. Fishman P and Black L. (1999) Indirect costs of migraine in a managed care population. *Cephalalgia* 19: 50-57; discussion 51.
10. Goetzel RZ, Gibson TB, Short ME, et al. (2010) A multi-worksite analysis of the relationships among body mass index, medical utilization, and worker productivity. *Journal of Occupational & Environmental Medicine* 52 Suppl 1: S52-58.
11. Goetzel RZgSRi, R. J., kins K, g S, et al. (2004) Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. *J Occup Environ Med* 46: 398-412.
12. Hellgren J, Cervin A, Nordling S, et al. (2010) Allergic rhinitis and the common cold--high cost to society. *Allergy* 65: 776-783.
13. Henke CJ, Levin TR, Henning JM, et al. (2000) Work loss costs due to peptic ulcer disease and gastroesophageal reflux disease in a health maintenance organization. *Am J Gastroenterol* 95: 788-792.
14. Hilton MF, Scuffham PA, Sheridan J, et al. (2008) Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *Journal of Occupational & Environmental Medicine* 50: 1228-1243.
15. Lamb CE, Ratner PH, Johnson CE, et al. (2006) Economic impact of workplace productivity losses due to allergic rhinitis compared with select medical conditions in the United States from an employer perspective. *Current Medical Research & Opinion* 22: 1203-1210.

16. Lerner D, Mirza FG, Chang H, et al. (2008) Impaired work performance among women with symptomatic uterine fibroids. *J Occup Environ Med* 50: 1149-1157.
17. Li X, Gignac MA and Anis AH. (2006) The indirect costs of arthritis resulting from unemployment, reduced performance, and occupational changes while at work. *Med Care* 44: 304-310.
18. Linde M, Gustavsson A, Stovner LJ, et al. (2012) The cost of headache disorders in Europe: the Eurolight project. *Eur J Neurol* 19: 703-711.
19. Ricci JA and Chee E. (2005) Lost productive time associated with excess weight in the U.S. workforce. *Journal of Occupational & Environmental Medicine* 47: 1227-1234.
20. Smit F, Willemse G, Koopmanschap M, et al. (2006) Cost-effectiveness of preventing depression in primary care patients: randomised trial. *Br J Psychiatry* 188: 330-336.
21. Snedecor SJ, Botteman MF, Bojke C, et al. (2009) Cost-effectiveness of eszopiclone for the treatment of adults with primary chronic insomnia. *Sleep* 32: 817-824.
22. Stewart WF, Ricci JA, Chee E, et al. (2003a) Lost productive work time costs from health conditions in the United States: results from the American Productivity Audit. *J Occup Environ Med* 45: 1234-1246.
23. Stewart WF, Ricci JA, Chee E, et al. (2003b) Lost productive time and cost due to common pain conditions in the US workforce. *JAMA* 290: 2443-2454.
24. Thavorncharoensap M, Teerawattananon Y, Yothasamut J, et al. (2010) The economic costs of alcohol consumption in Thailand, 2006. *BMC Public Health* 10: 323.
25. Uegaki K, Stomp-van den Berg SGM, de Bruijne MC, et al. (2011) Cost-utility analysis of a one-time supervisor telephone contact at 6-weeks post-partum to prevent extended sick leave following maternity leave in The Netherlands: results of an economic evaluation alongside a randomized controlled trial. *BMC Public Health* 11: 57.
26. Wilson DA, Bork K, Shea EP, et al. (2010) Economic costs associated with acute attacks and long-term management of hereditary angioedema. *Ann Allergy Asthma Immunol* 104: 314-320.
27. Zhang W, Bansback N, Guh D, et al. (2008) Short-term influence of adalimumab on work productivity outcomes in patients with rheumatoid arthritis. *Journal of Rheumatology* 35: 1729-1736.

Appendix C.1

C.1.1 Healthcare resource use for the STarT Back Trial

Cost components analysed in the economic evaluation of the STarT Back trial. Values are mean (sd) back pain-related resource, use and costs (£) per patient, by treatment group, for patients providing healthcare utilisation data at 12 months (n=567), unless stated otherwise.				
Resource use (units)	Cost (£)			
Resource/cost component	Intervention (n = 386)		Control group (n = 181)	
Study back pain clinic and physiotherapy	107·50	92·8	92·77	83·2
Primary care contacts:				
General Practitioner	33·54	63·6	40·27	63·9
Practice Nurse	1·33	9·8	1·27	6·0
Hospital-based care:				
NHS Consultant	25·15	86·5	27·63	79·7
Private Consultant	4·82	36·1	4·72	30·5
NHS x-ray	1·66	7·5	3·18	11·2
NHS CT scan	1·04	12·4	1·10	10·5
NHS MRI scan	12·06	48·5	14·83	53·0
NHS blood tests	0·27	2·1	0·48	3·8
NHS epidural injections	2·72	23·2	2·31	21·4
Private diagnostic tests (combined)	0·93	12·9	0·99	13·3
Private epidural injections	0·55	10·4	1·14	15·2
Other healthcare professionals:				
Additional (non-study) NHS physiotherapy	17·25	57·5	33·13	75·6
Private physiotherapy	1·86	17·2	6·42	35·5
NHS 'other'	1·45	17·0	3·53	26·8
Private 'other'	8·52	44·6	4·38	27·2
Out-of-pocket treatments				
Non-opioid analgesics	1·23	4·7	1·14	3·9
Weak opioid analgesics	0·68	8·0	0·27	0·8

Source: (Whitehurst et al., 2012)

C.1.2 Unit costs for the STarT Back Trial

Details of the unit costs assigned to healthcare resource use data and periods of work absence collected at 12-month follow-up	
Healthcare resource	Unit Cost (£)
Study back pain clinic and physiotherapy:	
Initial clinic session (30 minutes)	21·50
First post-clinic session: high risk 'intervention' group (1 hour)	43·00
First post-clinic session: all other patients (45 minutes)	32·25
Follow-up sessions: high risk intervention group (45 minutes)	32·25
Follow-up sessions: all other patients (30 minutes)	21·50
Primary care contacts:	
General Practitioner: surgery consultation	31·00
General Practitioner: home visit	105·00
Practice Nurse: surgery consultation	11·00
Practice Nurse: home visit	20·00
'Other' healthcare professional: surgery consultation	16·00
Hospital-based care:	
Consultant: first attendance	124·00
Consultant: follow-up	103·00
Diagnostic tests: x-ray	31·99
Diagnostic tests: CT scan	100·00
Diagnostic tests: MRI scan	179·00
Diagnostic tests: blood test	17·28
Epidural injections	204·57
Other healthcare professionals	
First consultation	38·00
Follow-up consultation	27·00
Out-of-pocket treatments	Patient reported costs
Prescribed medication	Patient-specific

Source: (Whitehurst et al., 2012)

Appendix C.2

C.2.1 Healthcare resource use for the IMPaCT Back study

Back pain-related healthcare (£) per patient, by study phase, for participants providing responses to the resource use questions at 6 months (n=547). Values are mean (sd) costs unless stated otherwise.

Healthcare resource	Cost (£) Phase 1 (n = 233)		Phase 3 (n = 314)	
Primary care contacts				
General Practitioner: surgery	36.8	53.5	32.38	50.1
General Practitioner: home visit	0.45	6.9	2.04	20.5
Practice Nurse: surgery	1.51	6.2	0.80	3.8
Practice Nurse: home visit	0.09	1.3	0.06	1.1
Physiotherapy service				
NHS	33.15	60.8	45.09	77.2
Private healthcare	13.82	63.7	11.90	40.8
Hospital-based care				
NHS Consultant	33.64	84.5	29.95	80.7
NHS admissions	9.79	91.6	12.84	116.9
NHS x-ray	4.39	12.9	3.97	11.0
NHS CT scan	0.43	6.6	1.32	11.2
NHS MRI scan	16.90	55.0	17.38	53.9
NHS blood tests	0.30	2.2	0.06	1.0
NHS epidural injections	3.51	26.6	2.09	19.9
Private Consultant	15.54	71.5	9.25	55.2
Private admissions	4.97	75.8	3.68	65.3
Private diagnostic tests	3.07	23.3	3.80	25.4
Private epidural injections	0.88	13.4	0.02	0.1
Other healthcare professionals				
NHS acupuncture	2.44	17.5	1.83	20.6
NHS osteopathy	0.16	2.5	3.19	20.8
NHS 'other'	8.52	46.7	4.96	33.9
Private acupuncture	11.90	67.1	3.34	23.3
Private osteopathy	11.08	37.0	9.02	38.7
Private 'other'	15.41	63.0	4.85	24.7
Out-of-pocket treatments	15.69	53.0	17.50	83.7
Prescribed medication	17.32	80.1	6.56	19.5

C.2.2 Unit costs for the IMPaCT Back Trial

Details of the unit costs assigned to healthcare resource use data and periods of work absence collected at 6 month follow-up	
Healthcare resource	Unit Cost (£)
Primary care contacts:	
General Practitioner: surgery consultation	31.00
General Practitioner: home visit	105.00
Practice Nurse: surgery consultation	11.00
Practice Nurse: home visit	20.00
Physiotherapy sessions	
Initial 45-minute assessment	32.25
Initial 1-hour assessment (high risk, Phase 3 patients only)	43.00
Follow-up sessions	21.50
Hospital-based care	
Consultant: back pain first attendance	124.00
Consultant: back pain follow-up	103.00
Consultant: 'other' first attendance	190.00
Consultant: 'other' follow-up	130.00
NHS admission: day case	562.00
NHS admission: elective stay	1157.00
Diagnostic tests: x-ray	31.99
Diagnostic tests: CT scan	100
Diagnostic tests: MRI scan	179
Diagnostic tests: blood test	17.28
Epidural injections	204.57
Other healthcare professionals	
First consultation	38.00
Follow-up consultation	29.00
Out-of-pocket treatments	Patient reported costs
Prescribed medication	Patient-specific
Periods of work absence	Patient-specific

[a] Hospital-based or private practice, e.g. acupuncture, osteopathy, hydrotherapy etc.

LIST OF REFERENCES

Adams J, Greig M and McQuaid RW. (2002) Mismatch in Local Labour Markets in Central Scotland: The Neglected Role of Demand. *Urban Studies* 39: 1399-1416.

Adams N. (1997) *The Psychophysiology of Low Back Pain* Edinburgh: Church Hill Livingstone.

Aggarwal R, Caitlyn T, Wilke A, et al. (2009) Psychometric Properties of the EuroQol-5D and Short Form-6D in Patients with Systemic Lupus Erythematosus. *J Rheumatology* 36: 1209-1216.

Ahlstrom L, Grimby-Ekman A, Hagberg M, et al. (2010) The work ability index and single-item question: associations with sick leave, symptoms, and health--a prospective study of women on long-term sick leave. *Scand J Work Environ Health* 36: 404-412.

Airaksinen O, Brox JI, Cedraschi C, et al. (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15 Suppl 2: S192-300.

Altman DG. (1995) *Practical Statistics for Medical Research*: Chapman & Hall.

AMCP. (2012) The Academy of Managed Care Pharmacy (AMCP) Format for Formulary Submissions. Version 3.1.

Andrews MJ, Bradley S, Stott D, et al. (2008) Successful Employer Search? An Empirical Analysis of Vacancy Duration Using Micro Data. *Economica* 75: 455-480.

Appleby J, Devlin N and Parkin D. (2007) NICE's cost effectiveness threshold
BMJ. 335: 358-359.

Appleby J, Devlin N, Parkin D, et al. (2009) Searching for cost effectiveness thresholds in the NHS. *Health Policy* 91: 239-245.

Aronsson G, Gustafsson K and Dallner M. (2000) Sick but yet at work. An empirical study of sickness presenteeism. *Journal of Epidemiology & Community Health* 54: 502-509.

ASHE. (2008) Annual Survey of Hours and Earnings. Available at: http://www.ons.gov.uk

Barber JA and Thompson SG. (2000) Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 19: 3219-3236.

BEAM U. (2004) United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. *BMJ* 329: 1377.

Beaton D, Bombardier C, Escorpizo R, et al. (2009) Measuring worker productivity: frameworks and measures. *Journal of Rheumatology* 36: 2100-2109.

Beaton DE and Kennedy CA. (2005) Beyond return to work: testing a measure of at-work disability in workers with musculoskeletal pain. *Qual Life Res* 14: 1869-1879.

Beaton DE, Tang K, Gignac MAM, et al. (2010) Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. *Arthritis care & research* 62: 28-37.

Beaumont PB. (1978) The duration of registered vacancies: an exploratory exercise. *Scottish Journal of Political Economy* 25: 75-87.

Berger ML, Murray JF, Xu J, et al. (2001) Alternative valuations of work loss and productivity. *Journal of Occupational & Environmental Medicine* 43: 18-24.

Beurskens A, de Vet HCW and Koke AJA. (1996) Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*: 71-76.

Birch S and Donaldson C. (2003) Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Social Science & Medicine* 56: 1121-1133.

Bland JM and Altman DG. (1995) Multiple significance tests: the Bonferroni method. *BMJ* 310: 170.

Boadway R and Bruce N. (1984) *Welfare economics.,* Oxford: Basil Backwell

Boardman AE, Greenburg D, Vining AR, et al. (2006) Cost-benefit analysis. Concepts and practise. Third ed.: Pearson prentice hall.

Boersma K and Linton SJ. (2005) Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain* 21: 38-43; discussion 69-72.

Boonen A, Brinkhuizen T, Landewe R, et al. (2010) Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. *Annals of the Rheumatic Diseases* 69: 1123-1128.

Boonen A, van den Heuvel R, van Tubergen A, et al. (2005) Large differences in cost of illness and wellbeing between patients with fibromyalgia, chronic low back pain, or ankylosing spondylitis. *Ann Rheum Dis* 64: 396-402.

Boonen A, van der Heijde D, Landewe R, et al. (2007) How do the EQ-5D, SF-6D and the well-being rating scale compare in patients with ankylosing spondylitis? *Ann Rheum Dis* 66: 771-777.

Borghouts JAJ, Koes BW, Vondeling H, et al. (1999) Cost-of-illness of neck pain in The Netherlands in 1996. *Pain* 80: 629-636.

Bowling A. (2005) Just one question: If one question works, why ask several? *Journal of Epidemiology and Community Health* 59: 342-345.

Braakman-Jansen LMA, Taal E, Kuper IH, et al. (2012) Productivity loss due to absenteeism and presenteeism by different instruments in patients with RA and subjects without RA. *Rheumatology* 51: 354-361.

Brazier J, Deverill M, Green C, et al. (1999) A review of the use of health status measures in economic evaluation. *Health Technol Assess* 3: i-iv, 1-164.

Brazier J, Ratcliffe J, Salomon JA, et al. (2007) Measuring and Valuing Health Benefits for Economic Evaluation. New York: Oxford University Press.

Breen A, Langworthy J and Bagust J. (2005) Improved early pain management for musculoskeletal disorders. Research Report 399.

Brennan GP, Fritz JM, Hunter SJ, et al. (2006) Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. *Spine (Phila Pa 1976)* 31: 623-631.

Bridges JF, Onukwugha E and Mullins CD. (2010) Healthcare rationing by proxy: cost-effectiveness analysis and the misuse of the $50,000 threshold in the US. *Pharmacoeconomics* 28: 175-184.

Briggs A, Clark T, Wolstenholme J, et al. (2003) Missing... presumed at random: cost-analysis of incomplete data. *Health Econ* 12: 377-392.

Briggs A and Fenn P. (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ* 7: 723-740.

Briggs AH and O'Brien BJ. (2001) The death of cost-minimization analysis? *Health Economics* 10: 179-184.

Bronfort G, Haas M, Evans RL, et al. (2004) Efficacy of spinal manipulation and mobilization for low back pain and neck pain: a systematic review and best evidence synthesis. *Spine J* 4: 335-356.

Brooks A, Hagen SE, Sathyanarayanan S, et al. (2010) Presenteeism: critical issues. *Journal of Occupational & Environmental Medicine* 52: 1055-1067.

Brouwer W, Rutten F and Koopmanschap M. (2001) Costing in economic evaluations. In: Drummond M and McGuire A (eds) *Economic evaluation in health care: merging theory with practice.* Oxford: Oxford University Press.

Brouwer WB and Koopmanschap MA. (2000) On the economic foundations of CEA. Ladies and gentlemen, take your positions! *Journal of Health Economics* 19: 439-459.

Brouwer WB, Koopmanschap MA and Rutten FF. (1997a) Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion. *Health Economics* 6: 511-514.

Brouwer WB, Koopmanschap MA and Rutten FF. (1997b) Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. *Health Economics* 6: 253-259.

Brouwer WB, Koopmanschap MA and Rutten FF. (1998) Patient and informal caregiver time in cost-effectiveness analysis. A response to the recommendations of the Washington Panel. *Int J Technol Assess Health Care* 14: 505-513.

Brouwer WB, van Exel NJ, Baltussen RM, et al. (2006) A dollar is a dollar is a dollar--or is it? *Value Health* 9: 341-347.

Brouwer WBF, Culyer AJ, van Exel NJA, et al. (2008) Welfarism vs. extra-welfarism. *Journal of Health Economics* 27: 325-338.

Brouwer WBF and Koopmanschap MA. (2005) The friction-cost method - Replacement for nothing and leisure for free? *Pharmacoeconomics* 23: 105-111.

Brouwer WBF, van Exel NJA, Koopmanschap MA, et al. (2002) Productivity costs before and after absence from work: as important as common? *Health Policy* 61: 173-187.

Brouwers EP, de Bruijne MC, Terluin B, et al. (2007) Cost-effectiveness of an activating intervention by social workers for patients with minor mental disorders on sick leave: a randomized controlled trial. *Eur J Public Health* 17: 214-220.

Brunenberg DE, Steyn MJv, Sluimer JC, et al. (2005) Joint Recovery Programme versus Usual Care: An Economic Evaluation of a Clinical Pathway for Joint Replacement Surgery. *Medical Care* 43: 1018-1026.

Bukowski G, Jenkins S and Roberts H. (2010) A qualitative overview of vacancy filling services for employers: Target setting and performance management Available at: http://statistics.dwp.gov.uk/asd/asd5/rports2009-2010/rrep686.pdf, [Accessed 12/2012.

Burton AK, Waddell G, Tillotson KM, et al. (1999) Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine (Phila Pa 1976)* 24: 2484-2491.

Burton WN, Chen C-Y, Conti DJ, et al. (2005) The association of health risks with on-the-job productivity. *Journal of Occupational & Environmental Medicine* 47: 769-777.

Burton WN, Conti DJ, Chen C-Y, et al. (2002) The economic burden of lost productivity due to migraine headache: a specific worksite analysis. *Journal of Occupational & Environmental Medicine* 44: 523-529.

Buxton MJ. (2006) Economic evaluation and decision making in the UK. *Pharmacoeconomics* 24: 1133-1142.

Byford S and Raftery J. (1998) Perspectives in economic evaluation. *BMJ* 316: 1529-1530.

CADTH. (2006) Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the Economic Evaluation of Health Technologies: Canada. Ottawa: Canadian Agency for Drugs and

Technologies in Health.

Cassidy JD, Carroll LJ and Cote P. (1998) The Saskatchewan health and back pain survey. The prevalence of low back pain and related disability in Saskatchewan adults. *Spine (Phila Pa 1976)* 23: 1860-1866; discussion 1867.

Childs JD, Piva SR and Fritz Jm PT. (2005) Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine* 30: 1331-1334.

Chou R, Qaseem A, Snow V, et al. (2007) Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med* 147: 478-491.

CIPD. (2007) CIPD. Annual Survey report. Available at: http://www.cipd.co.uk/NR/rdonlyres/746F1183-3941-4E6A-9EF6-135C29AE22C9/0/recruitretntsurv07.pdf.

CIPD. (2008) CIPD. Annual Survey report. Recruitment, retention and turn over. Available at:
http://www.cipd.co.uk/NR/rdonlyres/BE3C57BF-91FF-4AD0-9656-FAC27E5398AA/0/recruitmentretentionturnover2008.pdf.

CIPD. (2009) CIPD. Annual Survey report. Recruitment, retention and turn over. Available at: http://www.cipd.co.uk/hr-resources/survey-reports/recruitment-retention-turnover-2009.aspx.

Cisternas MG, Blanc PD, Yen IH, et al. (2003) A comprehensive study of the direct and indirect costs of adult asthma. *J Allergy Clin Immunol* 111: 1212-1218.

Claxton K, Martin S, Soares M, et al. (2013) Methods for the estimation of the NICE cost effectiveness threshold. CHE Research Paper 81.

Coast J. (2004) Is economic evaluation in touch with society's health values? *BMJ* 329: 1233-1236.

Coast J, Smith R and Lorgelly P. (2008a) Should the capability approach be applied in health economics? *Health Economics* 17: 667-670.

Coast J, Smith RD and Lorgelly P. (2008b) Welfarism, extra-welfarism and capability: the spread of ideas in health economics. *Social Science & Medicine* 67: 1190-1198.

Coelho RA, Siqueira FB, Ferreira PH, et al. (2008) Responsiveness of the Brazilian-Portuguese version of the Oswestry Disability Index in subjects with low back pain. *Eur Spine J* 17: 1101-1106.

Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences. Second edition. Hillsdale.,* NJ: Lawrence Earlbaum Associates; .

Cohen SP, Argoff CE and Carragee EJ. (2008) Management of low back pain. *BMJ* 337.

Collins JJ, Baase CM, Sharda CE, et al. (2005) The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *Journal of Occupational & Environmental Medicine* 47: 547-557.

Cookson R. (2005) QALYs and the capability approach. *Health Economics* 14: 817-829.

Cooper C and Dewe P. (2008) Well-being--absenteeism, presenteeism, costs and challenges. *Occupational Medicine (Oxford)* 58: 522-524.

Critchley DJ, Ratcliffe J, Noonan S, et al. (2007) Effectiveness and cost-effectiveness of three types of physiotherapy used to reduce chronic low back pain disability: a pragmatic randomized trial with economic evaluation. *Spine (Phila Pa 1976)* 32: 1474-1481.

Croft P. (1998) Outcome of low back pain in general practice: a prospective study. *BMJ* 316: 1356-1359.

Crook J, Milner R, Schultz IZ, et al. (2002) Determinants of Occupational Disability Following a Low Back Injury: A Critical Review of the Literature. *Journal of Occupational Rehabilitation* 12: 277-295.

Culyer A, McCabe C, Briggs A, et al. (2007) Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy* 12: 56-58.

Culyer AJ. (1989) THE Normative Economics of Health Care Finance and Provision. *Oxford Review of Economic Policy* 5: 34-58.

Dagenais S, Caro J and Haldeman S. (2008) A systematic review of low back pain cost of illness studies in the United States and internationally. *The Spine Journal*: 8-20.

Dagenais S, Roffey DM, Wai EK, et al. (2009) Can cost utility evaluations inform decision making about interventions for low back pain? *Spine J* 9: 944-957.

Dahm KT, Brurberg KG, Jamtvedt G, et al. (2010) Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. *Cochrane Database Syst Rev*: CD007612.

Dakin H and Wordsworth S. (2013) Cost-minimisation analysis versus cost-effectiveness analysis, revisited. *Health Econ* 22: 22-34.

Daley M, Morin CM, LeBlanc M, et al. (2009) The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep* 32: 55-64.

de Koning J and Tuyl FAWM. (1984) The relation between labour time,production and employment (in Dutch). Netherlands Economic Institute,

Rotterdam.

Deyo RA and Phillips WR. (1996) Low Back Pain: A Primary Care Challenge. *Spine* 21: 2826-2832.

Dionne CE, Dunn KM and Croft PR. (2006) Does back pain prevalence really decrease with increasing age? A systematic review. *Age and Ageing* 35: 229-234.

DLA PIPER. (2012) DLA PIPER HR benchmarker.Vacancy duration and medical statistics report to the University of Birmingham, health economics unit. .

DOH. (1999) Department of Health. Statistical bulletin. .

Dolan P. (1997) Modelling valuations for EuroQol health states. *Med Care* 35: 1095-1108.

Dolan P and Edlin R. (2002) Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *Journal of Health Economics* 21: 827-843.

Dolan P, Gudex C, Kind P, et al. (1996) The time trade-off method: results from a general population study. *Health Econ* 5: 154.

Donaldson C, Baker R, Mason H, et al. (2011) The social value of a QALY: raising the bar or barring the raise? *BMC Health Services Research* 11: 8.

Drummond M and Rutten FF. (2008) New Guidelines for Economic Evaluation in Germany and the United Kingdom are we any closer to developing international standards? , 46.

Drummond M, Sculpher M, Torrance G, et al. (2005) Methods for the economic evaluation of health care programmes. NewYork: Oxford University Press.

Drummond M, Sculpher M, Torrance G, et al. (2005) Methods for the economic evaluation of health care programmes. NewYork: Oxford University Press.

Drummond M, Weatherly H and Ferguson B. (2008) Economic evaluation of health interventions. *BMJ* 337.

Dunn KM, Jordan KP and Croft PR. (2011) Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain* 15: 313-319.

DWP. (2013) Department of Work and Pensions. Fit note: guidance for GPs. London: Department for Work and Pensions; https://www.gov.uk/government/publications/fit-note-guidance-for-gps.

Eichler HG, Kong SX, Gerth WC, et al. (2004) Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? *Value Health* 7: 518-528.

Ekman M, Johnell O and Lidgren L. (2005) The economic cost of low back pain in Sweden in 2001. *Acta Orthop* 76: 275-284.

Engers A, Jellema P, Wensing M, et al. (2008) Individual patient education for low back pain. *Cochrane Database Syst Rev*: CD004057.

Erdogan-Ciftci E and Koopmanschap M. (2011) Estimation of productivity costs using the friction cost method: new evidence using national data.

Evanoff B, Abedin S, Grayson D, et al. (2002) Is disability underreported following work injury? *J Occup Rehabil* 12: 139-150.

Evans CJ. (2004) Health and work productivity assessment: state of the art or state of flux? *J Occup Environ Med* 46: S3-11.

Feleus A, Tineke van D, Sita MAB-Z, et al. (2007) Kinesiophobia in patients with non-traumatic arm, neck and shoulder complaints: a prospective cohort study in general practice. *BMC Musculoskeletal Disorders*.

Fenwick E and Byford S. (2005) A guide to cost-effectiveness acceptability curves. *The British Journal of Psychiatry* 187: 106-108.

Ferrie JE, Kivimaki M, Head J, et al. (2005) A comparison of self-reported sickness absence with absences recorded in employers' registers: evidence from the Whitehall II study. *Occup Environ Med* 62: 74-79.

Finkelstein EA, DiBonaventura Md, Burgess SM, et al. (2010) The costs of obesity in the workplace. *Journal of Occupational & Environmental Medicine* 52: 971-976.

Fishman P and Black L. (1999) Indirect costs of migraine in a managed care population. *Cephalalgia* 19: 50-57; discussion 51.

Folland S and Goodman A Cea. (2001) *The economics of health and health care. Upper Saddle River, NJ.,* NJ: Prentice Hall.

Foster N, Mullis R, Young J, et al. (2010) IMPaCT Back study protocol. Implementation of subgrouping for targeted treatment systems for low back pain patients in primary care: a prospective population-based sequential comparison. *BMC Musculoskeletal Disorders* 11: 186.

Foster NE, Bishop A, Thomas E, et al. (2008) Illness perceptions of low back pain patients in primary care: what are they, do they change and are they associated with outcome? *Pain* 136: 177-187.

Freburger JK, Holmes GM, Agans RP, et al. (2009) The rising prevalence of chronic low back pain. *Arch Intern Med* 169: 251-258.

Frew E. (2010) Benefit assessement for cost-benefit analysis studies in health care: A guide to carrying a stated preference willingness to pay survey in healthcare. In: McIntosh E, Clarke PM, Frew EJ, et al. (eds) *Applied methods of cost-benefit analysis in healthcare. Oxford handbooks of health economics.*: Oxford university press.

Furlan AD, Imamura M, Dryden T, et al. (2009) Massage for low back pain: an updated systematic review within the framework of the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 34: 1669-1684.

Furlan AD, van Tulder M, Cherkin D, et al. (2005) Acupuncture and dry-needling for low back pain: an updated systematic review within the framework of the cochrane collaboration. *Spine (Phila Pa 1976)* 30: 944-963.

Gafni A. (2006) Economic Evaluation of Health-care Programmes: Is CEA Better than CBA? *Environmental & Resource Economics* 34: 407-418.

Gafni A, Walter S and Birch S. (2012) Uncertainity and the Decision Maker: Assessing and Managing the Risk of Undesirable Outcomes. *Health Econ*.

Gallefoss F and Bakke PS. (2002) Cost-benefit and cost-effectiveness analysis of self-management in patients with COPD--a 1-year follow-up randomized, controlled trial. *Respir Med* 96: 424-431.

Garber AM and Phelps CE. (1997) Economic foundations of cost-effectiveness analysis. *J Health Econ* 16: 1-31.

Giovannetti ER, Wolff JL, Frick KD, et al. (2009) Construct Validity of the Work Productivity and Activity Impairment Questionnaire across Informal Caregivers of Chronically Ill Older Patients. *Value in Health* 12: 1011.

Goetzel RZ, Gibson TB, Short ME, et al. (2010) A multi-worksite analysis of the relationships among body mass index, medical utilization, and worker productivity. *Journal of Occupational & Environmental Medicine* 52 Suppl 1: S52-58.

Goetzel RZL, Stacey R; Ozminkowski, Ronald J; Hawkins, Kevin; Wang, Shaohung; Lynch, Wendy. (2004) Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. *Journal of Occupational & Environmental Medicine* 46: 398-412.

Gold M, Siegel J, Russell L, et al. (1996) *Cost-effectiveness in health and medicine.,* New York: Oxford University Press.

Graf von der Schulenberg JG, W ; Jost, F et al (2008) German recommendations on health economic evaluation: third and updated version of the Hanover Consensus. . *Value Health* 11: 539.

Gray AM, Clarke PM, Wolstenholme JL, et al. (2010) *Applied Methods of Cost-effectiveness Analysis in Healthcare. Handbooks in Health Economic Evaluation, volume 3* New York: Oxford University Press.

Grotle M, Foster NE, Dunn KM, et al. (2010) Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain* 151: 790-797.

Grovle L, Haugen AJ, Keller A, et al. (2012) Poor agreement found between self-report and a public registry on duration of sickness absence. *J Clin Epidemiol* 65: 212-218.
Guyatt G, Walter S and Norman G. (1987) Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 40: 171-178.
Hagberg M, Tornqvist EW and Toomingas A. (2002) Self-reported reduced productivity due to musculoskeletal symptoms: associations with workplace and individual factors among white-collar computer users. *J Occup Rehabil* 12: 151-162.
Hagen KB, Hilde G, Jamtvedt G, et al. (2000) Bed rest for acute low back pain and sciatica. *Cochrane Database Syst Rev*: CD001254.
Hakkaart-van Roijen L, Tan S and Bouwmans C. Handleiding voor kostenonderzoek. (2010) Methoden en standaard kostprijzen voor economische evaluaties in degezondheidszorg [Manual for Cost Research, Methods and Standard Cost Prices for Economic Evaluations in Health Care]. Diemen: College voor Zorgverzekeringen; .
Hakkaart-van Roijen L, Hoeijenbos MB, Regeer EJ, et al. (2004) The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta Psychiatrica Scandinavica* 110: 383-392.
Hanley N, Ryan M and Wright R. (2003) Estimating the monetary value of health care: lessons from environmental economics. *Health Econ* 12: 3-16.
Hanly P, Timmons A, Walsh PM, et al. (2012) Breast and prostate cancer productivity costs: a comparison of the human capital approach and the friction cost approach. *Value in Health* 15: 429-436.
Hansson EK and Hansson TH. (2005) The costs for persons sick-listed more than one month because of low back or neck problems. A two-year prospective study of Swedish patients. *Eur Spine J* 14: 337-345.
Harkness EF, Macfarlane GJ, Nahit ES, et al. (2003) Risk factors for new-onset low back pain amongst cohorts of newly employed workers. *Rheumatology (Oxford)* 42: 959-968.
Harrison MJ, Lunt M, Verstappen SM, et al. (2010) Exploring the validity of estimating EQ-5D and SF-6D utility values from the health assessment questionnaire in patients with inflammatory arthritis. *Health Qual Life Outcomes* 8: 21.
Hauck K, Smith PC and and Maria G. (2004) The Economics of Priority Setting for Health Care: A Literature Review

Hay E, Dunn K, Hill J, et al. (2008) A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. The STarT Back Trial Study Protocol". *Musculoskeletal Disorders* 9: 58.
Hayden JA, van Tulder MW and Tomlinson G. (2005) Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 142: 776-785.
Hays R and Revicki D. (2005) *Reliability and validity (including responsiveness). In Assessing Quality of Life in Clinical Trials. Second edition. Edited by Fayers P, Hays R. ; 2005. OpenURL,* New York:: Oxford University Press.
Hellgren J, Cervin A, Nordling S, et al. (2010) Allergic rhinitis and the common cold--high cost to society. *Allergy* 65: 776-783.
Hemp P. (2004) Presenteeism: at work--but out of it. *Harvard Business Review* 82: 49-58.

Henke CJ, Levin TR, Henning JM, et al. (2000) Work loss costs due to peptic ulcer disease and gastroesophageal reflux disease in a health maintenance organization. *Am J Gastroenterol* 95: 788-792.

Henke RM, Carls GS, Short ME, et al. (2010) The relationship between health risks and health and productivity costs among employees at Pepsi Bottling Group. *Journal of Occupational & Environmental Medicine* 52: 519-527.

Henschke N, Ostelo RW, van Tulder MW, et al. (2010) Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev*: CD002014.

Herman PM, Szczurko O, Cooley K, et al. (2008) Cost-effectiveness of naturopathic care for chronic low back pain. *Altern Ther Health Med* 14: 32-39.

Heymans MW, van Tulder MW, Esmail R, et al. (2004) Back schools for non-specific low-back pain. *Cochrane Database Syst Rev*: CD000261.

Hicks JR. (1939) The Foundations of Welfare Economics. *The Economic Journal* 49: 696-712.

Hill JC, Dunn KM, Lewis M, et al. (2008) A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum* 59: 632-641.

Hill JC, Whitehurst DGT, Lewis M, et al. (2011) Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *The Lancet* 378: 1560-1571.

Hillman M, Wright A, Rajaratnam G, et al. (1996) Prevalence of low back pain in the community: implications for service provision in Bradford, UK. *Journal of Epidemiology and Community Health* 50: 347-352.

Hilton MF, Scuffham PA, Sheridan J, et al. (2008) Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *Journal of Occupational & Environmental Medicine* 50: 1228-1243.

Hirth RA, Chernew ME, Miller E, et al. (2000) Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making* 20: 332-342.

Hiscock J and Ritchie J. (2001) National Centre for Social Research, Department for Work and Pensions. The role of GPs in sickness certification; a report of research carried out by the National Centre for Social Research on behalf of the Department for Work and Pensions. Leeds: Department for Work and Pensions; Report No 148.

Hollinghurst S, Sharp D, Ballard K, et al. (2008) Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain: economic evaluation. *BMJ* 337.

Hoy D, Bain C, Williams G, et al. (2012) A systematic review of the global prevalence of low back pain. *Arthritis Rheum* 64: 2028-2037.

Hoy D, Brooks P, Blyth F, et al. (2010a) The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol* 24: 769-781.

Hoy D, March L, Brooks P, et al. (2010b) Measuring the global burden of low back pain. *Best Practice & Research Clinical Rheumatology* 24: 155-165.

HSE. (2003/2004) Health and Safety Executive. http://www.hse.gov.uk/press/2005/e05077.htm.

Hudes K. (2011) The Tampa Scale of Kinesiophobia and neck pain, disability and range of motion: a narrative review of the literature. *J Can Chiropr Assoc.* 55: 222-232.

Hussey S, Hoddinott P, Wilson P, et al. (2004) Sickness certification system in the United Kingdom: qualitative study of views of general practitioners in Scotland. *BMJ* 328: 88.

Hutubessy RC, van Tulder MW, Vondeling H, et al. (1999) Indirect costs of back pain in the Netherlands: a comparison of the human capital method with the friction cost method. *Pain* 80: 201-207.

Ihlebaek C, Hansson TH, Laerum E, et al. (2006) Prevalence of low back pain and sickness absence: a "borderline" study in Norway and Sweden. *Scand J Public Health* 34: 555-558.

Jacob-Tacken KHM, Koopmanschap MA, Meerding WJ, et al. (2005) Correcting for compensating mechanisms related to productivity costs in economic evaluations of health care programmes. *Health Economics* 14: 435-443.

Jacobs P and Fassbender K. (1998) The measurement of indirect costs in the health economics evaluation literature. A review. *Int J Technol Assess Health Care* 14: 799-808.

Johannesson M. (1995) A note on the depreciation of the societal perspective in economic evaluation of health care. *Health Policy* 33: 59-66.

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

Johannesson M and Jonsson B. (1991) Economic evaluation in health care: is there a role for cost-benefit analysis? *Health Policy* 17: 1-23.

Johannesson M and Karlsson G. (1997) The friction cost method: a comment. *J Health Economics* 16: 249-255.

Johannesson MJ, Bengt ;Jönsson, Linus; Kobelt, Gisela ; Zethraeus,Niklas. (2009) Why Should Economic Evaluations of Medical Innovations Have a Societal Perspective? In: Economics OoH (ed).

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

Johnson RE, Jones GT, Wiles NJ, et al. (2007) Active exercise, education, and cognitive behavioral therapy for persistent disabling low back pain: a randomized controlled trial. *Spine (Phila Pa 1976)* 32: 1578-1585.

Jones GT and Macfarlane GJ. (2005) Epidemiology of low back pain in children and adolescents. *Archives of Disease in Childhood* 90: 312-316.

Jonsson D and Husberg M. (2000) Socioeconomic costs of rheumatic diseases. Implications for technology assessment. *Int J Technol Assess Health Care* 16: 1193-1200.

Jordan K, Dunn KM, Lewis M, et al. (2006) A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. *Journal of Clinical Epidemiology* 59: 52.

Kaldor N. (1939) Welfare propositions of economics and interpersonal comparisons of utility. *The Economic Journal* 49: 549-552.

Kamper SJ, Maher CG, Hancock MJ, et al. (2010) Treatment-based subgroups of low back pain: A guide to appraisal of research studies and a summary of current evidence. *Best Practice & Research Clinical Rheumatology* 24: 181-191.

Kazis LE, Anderson JJ and Meenan RF. (1989) Effect sizes for interpreting changes in health status. *Medical Care* 27: S178.

Kessler RC, Ames M, Hymel PA, et al. (2004) Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *Journal of Occupational & Environmental Medicine* 46: S23-37.

Kessler RC, Barber C, Beck A, et al. (2003) The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med* 45: 156-174.

Klaber-Moffett J, Richardson G, Sheldon T, et al. (1995) Back Pain its Management and Cost to Society. Discussion Paper. Centre for Health Economics. The University of York.

Knies S, Severens JL, Ament AJ, et al. (2010) The transferability of valuing lost productivity across jurisdictions. differences between national pharmacoeconomic guidelines. *Value.Health* 13: 519-527.

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

Koes BW, van Tulder M, Lin CW, et al. (2010) An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J* 19: 2075-2094.

Koopmanschap M, Burdorf A, Jacob K, et al. (2005) Measuring productivity changes in economic evaluation: setting the research agenda. *Pharmacoeconomics* 23: 47-54.

Koopmanschap M and van Ineveld B. (1992) Towards a new approach for estimating indirect costs of disease. *Soc Sci Med* 34: 1005 - 1010.

Koopmanschap MA and Rutten FF. (1993) Indirect costs in economic studies: confronting the confusion. *Pharmacoeconomics* 4: 446-454.

Koopmanschap MA, Rutten FF, van Ineveld BM, et al. (1995) The friction cost method for measuring indirect costs of disease. *J Health Econ* 14: 171-189.

Koopmanschap MA and Rutten FFH. (1996) A practical guide for calculating indirect costs of disease. *Pharmacoeconomics* 10: 460-466.

Koopmanschap MA, Rutten FFH, vanIneveld BM, et al. (1997) The friction cost method: Reply. *Journal of Health Economics* 16: 257-259.

Krismer M and van Tulder M. (2007) Low back pain (non-specific). *Best Practice & Research Clinical Rheumatology* 21: 77-91.

Krol M, Brouwer WBF, Severens JL, et al. (2012) Productivity cost calculations in health economic evaluations: correcting for compensation mechanisms and multiplier effects. *Social Science & Medicine* 75: 1981-1988.

Krol M, Papenburg J, Koopmanschap M, et al. (2011) Do productivity costs matter?: the impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders. *Pharmacoeconomics* 29: 601-619.

Krol M, Sendi P and Brouwer W. (2009) Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. *Value in Health* 12: 172-180.

Lamb S, Lall R, Hansen Z, et al. (2010a) A multicentred randomised controlled trial of a primary care-based cognitive behavioural programme for low back pain. The Back Skills Training (BeST) trial.

Lamb SE, Hansen Z, Lall R, et al. (2010b) Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. *The Lancet* 375: 916-923.

Lamers LM, Meerding W-J, Severens JL, et al. (2005) The relationship between productivity and health-related quality of life: an empirical exploration in persons with low back pain. *Quality of Life Research* 14: 805-813.

Layard R and Glaister S. (1994) Cost-Benefit Analysis. Cambridge, UK: Cambridge University Press.

Leboeuf-Yde C. (1999) Smoking and low back pain. A systematic literature review of 41 journal articles reporting 47 epidemiologic studies. *Spine (Phila Pa 1976)* 24: 1463-1470.

Lerner D, Amick BC, III, Rogers WH, et al. (2001) The Work Limitations Questionnaire. *Med Care.*: 72-85.

Lerner D, Mirza FG, Chang H, et al. (2008) Impaired work performance among women with symptomatic uterine fibroids. *J Occup Environ Med* 50: 1149-1157.

Lewis M, James M, Stokes E, et al. (2007) An economic evaluation of three physiotherapy treatments for non-specific neck disorders alongside a randomized trial. *Rheumatology (Oxford)* 46: 1701-1708.

Li X, Gignac MA and Anis AH. (2006) The indirect costs of arthritis resulting from unemployment, reduced performance, and occupational changes while at work. *Med Care* 44: 304-310.

Liem MS, Halsema JA, van der Graaf Y, et al. (1997) Cost-effectiveness of extraperitoneal laparoscopic inguinal hernia repair: a randomized comparison with conventional herniorrhaphy. Coala trial group. *Ann Surg* 226: 668-675; discussion 675-666.

Liljas B. (1998) How to calculate indirect costs in economic evaluations. *Pharmacoeconomics* 13: 1-7.

Lin CC, McAuley JH, Macedo L, et al. (2011) Relationship between physical activity and disability in low back pain: A systematic review and meta-analysis. *Pain* 152: 607-613.

Linde L, Sorensen J, Ostergaard M, et al. (2008) Health-Related Quality of Life: Validity, Reliability, and Responsiveness of SF-36, EQ-15D, EQ-5D, RAQoL, and HAQ in Patients with Rheumatoid Arthritis. *J Rheumatology.* 35: 1528-1537.

Linde M, Gustavsson A, Stovner LJ, et al. (2012) The cost of headache disorders in Europe: the Eurolight project. *Eur J Neurol* 19: 703-711.

Lindeboom R, Sprangers MA and Zwinderman AH. (2005) Responsiveness: a reinvention of the wheel? *Health Qual Life Outcomes* 3: 8.

Linton SJ and Nordin E. (2006) A 5-year follow-up evaluation of the health and economic consequences of an early cognitive behavioral intervention for back pain: a randomized, controlled trial. *Spine* 31: 853-858.

Little P, Roberts L, Blowers H, et al. (2001) Should we give detailed advice and information booklets to patients with back pain? A randomized controlled factorial trial of a self-management booklet and doctor advice to take exercise for back pain. *Spine (Phila Pa 1976)* 26: 2065-2072.

Liu JLY, Maniadakis N, Gray A, et al. (2002) The economic burden of coronary heart disease in the UK. *Heart* 88: 597-603.

Loeppke R, Taitel M, Richling D, et al. (2007) Health and Productivity as a Business Strategy. *Journal of Occupational and Environmental Medicine* 49: 712-721.

Lofland JH, Pizzi L and Frick KD. (2004) A review of health-related workplace productivity loss instruments. *Pharmacoeconomics* 22: 165-184.

Louise MA, Braakman J, Erik T, et al. (2012) Productivity loss due to absenteeism and presenteeism by different instruments in patients with RA and subjects without RA. *Rheumatology*: 354-361.

Luce BM, WG; Siegel, JE ; Lipscomb, J. (1996) Estimating costs in cost-effectiveness analysis. In Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-Effectiveness in Health and Medicine.. New York: : Oxford University Press.

Luengo-Fernandez R, Leal J, Gray A, et al. (2006) Cost of cardiovascular diseases in the United Kingdom. *Heart* 92: 1384 - 1389.

Macfarlane GJ, Beasley M, Jones EA, et al. (2012) The prevalence and management of low back pain across adulthood: results from a population-based cross-sectional study (the MUSICIAN study). *Pain* 153: 27-32.

Macfarlane GJ, Jones GT and Hannaford PC. (2006) Managing low back pain presenting to primary care: where do we go from here? *Pain* 122: 219-222.

Macfarlane GJ, Pallewatte N, Paudyal P, et al. (2009) Evaluation of work-related psychosocial factors and regional musculoskeletal pain: results from a EULAR Task Force. *Annals of the Rheumatic Diseases* 68: 885-891.

Maetzel A and Li L. (2002) The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Pract Res Clin Rheumatol* 16: 23-30.

Mallen CD, Peat G, Elaine T, et al. (2007) Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract* 57: 655-661.

Manca A, Dumville JC, Torgerson DJ, et al. (2007) Randomized trial of two physiotherapy interventions for primary care back and neck pain patients: cost–effectiveness analysis. *Rheumatology* 46: 1495-1501.

Manca A, Hawkins N and Sculpher MJ. (2005) Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 14: 487-496.

Maniadakis N and Gray A. (2000) The economic burden of back pain in the UK. *Pain* 84: 95-103.

Mannion AF, Horisberger B, Eisenring C, et al. (2009) The association between beliefs about low back pain and work presenteeism. *Journal of Occupational & Environmental Medicine* 51: 1256-1266.

Marra CA, Woolcott JC, Kopecc JA, et al. (2005) comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Social Science & Medicine* 60: 1571-1582.

Mattke S, Balakrishnan A, Bergamo G, et al. (2007) A review of methods to measure health-related productivity loss. *American Journal of Managed Care* 13: 211-217.

Mauskopf JA, Paul JE, Grant DM, et al. (1998) The role of cost-consequence analysis in healthcare decision-making. *Pharmacoeconomics* 13: 277-288.

May S. (2010) Self-management of chronic low back pain and osteoarthritis. *Nat Rev Rheumatol* 6: 199-209.

McCabe C, Claxton K and Culyer AJ. (2008) The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 26: 733-744.

McEachan R, Lawton R, Jackson C, et al. (2011) Testing a workplace physical activity intervention: a cluster randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity* 8: 29.

McIntosh E, Donaldson C and Ryan M. (1999) Recent advances in the methods of cost-benefit analysis in healthcare. Matching the art to the science. *Pharmacoeconomics* 15: 357-367.

McIntosh E, M.Clarker. P, Emma .J F, et al. (2010) Applied methods of cost-benefit analysis in healthcare. Edited by Oxford handbooks of health economics.: Oxford university press.

McIntosh G, Frank J, Hogg-Johnson S, et al. (2000) Low Back Pain Prognosis: Structured Review of the Literature. *Journal of Occupational Rehabilitation* 10: 101-115.

Meerding WJ, Ijzelenberg W, Koopmanschap MA, et al. (2005) Health problems lead to considerable productivity loss at work among workers with high physical load jobs. *Journal of Clinical Epidemiology* 58: 517-523.

Morris S, Devlin N and David. P. (2007) *Economic Analysis in Health Care.* : John Wiley and Sons Ltd.

Morrison J. (2011) The role of the GP in keeping people in work. *Occupational Medicine* 61: 74-75.

Mortimer M, Pernold G and Wiktorin C. (2006) Low back pain in a general population. Natural course and influence of physical exercise--a 5-year follow-up of the Musculoskeletal Intervention Center-Norrtalje Study. *Spine (Phila Pa 1976)* 31: 3045-3051.

NICE.(2008) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal.

NICE.(2009a) National Institute for Health and Care Excellence: Low back pain Early management of persistent non-specific low back pain.

NICE.(2009b) National Institute for Health and Clinical Excellence: Low Back Pain. Costing Report. Implementing Nice Guidance. .

NICE.(2011) National Institute of Clinical Excellence. National costing report: chronic obstructive pulmonary disease.

NICE.(2013) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal.

Nicholson S, Pauly MV, Polsky D, et al. (2006) Measuring the effects of work loss on productivity with team production. *Health Econ* 15: 111-123.

NMA.(2005) Norwegian Medicine Agency. Norwegian Guidelines for Pharmacoeconomic Analysis in Connection with Applications for Reimbursement.

Oslo: Department for Pharmacoeconomics, Ministry of Health and Social Affaires, .

Nyman JA. (2012) Productivity Costs Revisted: Toward a new US policy. *Health Economics* 21: 1387-1401.

Obradovic M, Lal A and Liedgens H. (2013) Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. *Health and Quality of Life Outcomes* 11: 110.

Olsen JA and Richardson J. (1999) Production gains from health care: what should be included in cost-effectiveness analyses? *Soc Sci Med* 49: 17-26.

Ostelo RW, Deyo RA, Stratford P, et al. (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)* 33: 90-94.

Ostelo RW, van Tulder MW, Vlaeyen JW, et al. (2005) Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev*: CD002014.
Ozminkowski RJ, Goetzel RZ, Chang S, et al. (2004) The application of two health and productivity instruments at a large employer. *Journal of Occupational & Environmental Medicine* 46: 635-648.
Palmer KT, Walsh K, Bendall H, et al. (2000) Back pain in Britain: comparison of two prevalence surveys at an interval of 10 years. *BMJ* 320: 1577-1578.
Papageorgiou AC, Croft PR, Thomas E, et al. (1996) Influence of previous pain experience on the episode incidence of low back pain: results from the South Manchester Back Pain Study. *Pain* 66: 181-185.
Pauly MV, Nicholson S, Polsky D, et al. (2008) Valuing reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. *Health Economics* 17: 469-485.
Pauly MV, Nicholson S, Xu J, et al. (2002) A general model of the impact of absenteeism on employers and employees. *Health Economics* 11: 221-231.
PBAC. (2008) Pharmaceutical Benefits Advisory Committee, Department of Health and Ageing, Australian Government
http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-pubs-pharmpac-gusubpac.htm. (accessed April, 2013).
Pearce D. (1971) *Macmillan Studies in Economics: Cost Benefit Analysis*: McMillan Press Ltd
Pengel LH, Refshauge KM and Maher CG. (2004) Responsiveness of pain, disability, and physical impairment outcomes in patients with low back pain. *Spine (Phila Pa 1976)* 29: 879-883.
Pincus T, Burton AK, Vogel S, et al. (2002) A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)* 27: E109-120.
Pole JD, Franche RL, Hogg-Johnson S, et al. (2006) Duration of work disability: a comparison of self-report and administrative data. *Am J Ind Med* 49: 394-401.
Ponto KA, Merkesdal S, Hommel G, et al. (2013) Public health relevance of Graves' orbitopathy. *J Clin Endocrinol Metab* 98: 145-152.
Prasad M, Shikiar R, Shih YC, et al. (2004) A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. . *Pharmacoeconomics* 22: 225-244.
Pritchard C and Sculpher M. (2000) *Productivity Costs: Principles and Practise in Economic Evaluation,* Office of Health Economics.
Punnett L and Wegman DH. (2004) Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *J Electromyogr Kinesiol* 14: 13-23.
Raftery J. (2009) Should NICE's threshold range for cost per QALY be raised? No. *BMJ* 338.
Ratcliffe J, Thomas KJ, MacPherson H, et al. (2006) A randomised controlled trial of acupuncture care for persistent low back pain: cost effectiveness analysis. *BMJ* 333: 626.
Rawlins MD and Culyer AJ. (2004) National Institute for Clinical Excellence and its value judgments. *BMJ* 329: 224-227.

Reilly MC, Bracco A, Ricci LF, et al. (2004) he validity and accuracy of the Work Productivity and Activity Impairment Questionnaire-irritable bowel syndrome version WPAI: IBS. *T Aliment Pharmacol Ther* 20: 459-467.

Reilly MC, Gerlier L, Brabant Y, et al. (2008) Validity, reliability, and responsiveness of the work productivity and activity impairment questionnaire in Crohn's disease. *Clinical Therapeutics* 30: 393-404.

Reilly MC, Gooch KL, Wong RL, et al. (2010) Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. *Rheumatology (Oxford)* 49: 812-819.

Reilly MC, Lavin PT, Kahler KH, et al. (2003) Validation of the Dermatology Life Quality Index and the Work Productivity and Activity Impairment-Chronic Hand Dermatitis questionnaire in chronic hand dermatitis. *J Am Acad Dermatol* 48: 128-130.

Reilly MC, Zbrozek AS and Dukes EM. (1993) The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 4: 353-365.

Revicki D, Cella D, Hays R, et al. (2006) Responsiveness and minimal important differences for patient reported outcomes. *Health and Quality of Life Outcomes* 4: 70.

Ricci JA and Chee E. (2005) Lost productive time associated with excess weight in the U.S. workforce. *Journal of Occupational & Environmental Medicine* 47: 1227-1234.

Riddle DL, Lee KT and Stratford PW. (2001) Use of SF-36 and SF-12 health status measures: a quantitative comparison for groups versus individual patients. *Med Care* 39: 867-878.

Rivero-Arias O, Gray A, Frost H, et al. (2006) Cost-utility analysis of physiotherapy treatment compared with physiotherapy advice in low back pain. *Spine (Phila Pa 1976)* 31: 1381-1387.

Rivero-Arias O, Gray A and Wolstenholme J. (2010) Burden of disease and costs of aneurysmal subarachnoid haemorrhage (aSAH) in the United Kingdom. *Cost effectiveness and resource allocation* 8: 6.

Roberts L, Little P, Chapman J, et al. (2002) The back home trial: general practitioner-supported leaflets may change back pain behavior. *Spine (Phila Pa 1976)* 27: 1821-1828.

Roberts TE, Tsourapas A, Sutcliffe L, et al. (2012) Is Accelerated Partner Therapy (APT) a cost-effective alternative to routine patient referral partner notification in the UK? Preliminary cost–consequence analysis of an exploratory trial. *Sexually Transmitted Infections* 88: 16-20.

Roelofs J, Goubert L, Peters ML, et al. (2004) The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *Eur J Pain* 8: 495-502.

Roelofs PD, Deyo RA, Koes BW, et al. (2008) Non-steroidal anti-inflammatory drugs for low back pain. *Cochrane Database Syst Rev*: CD000396.

Roland M and Morris R. (1983) A study of the natural history of low back pain: Part 1. Development of a reliable and sensitive measure of disability in low-back pain. *Spine* 8: 141-144.

Roper S. (1988) RECRUITMENT METHODS AND VACANCY DURATION*. *Scottish Journal of Political Economy* 35: 51-64.

Rossignol M, Rozenberg S and Leclerc A. (2009) Epidemiology of low back pain: what's new? *Joint Bone Spine* 76: 608-613.

Rubinstein SM, van Middelkoop M, Assendelft WJ, et al. (2011) Spinal manipulative therapy for chronic low-back pain: an update of a Cochrane review. *Spine (Phila Pa 1976)* 36: E825-846.

Russell LB, Gold MR, Siegel JE, et al. (1996) The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 276: 1172-1177.

Rutten-van Molken MP, van Nooten FE, Lindemann M, et al. (2007) A 1-year prospective cost-effectiveness analysis of roflumilast for the treatment of patients with severe chronic obstructive pulmonary disease. *Pharmacoeconomics* 25: 695-711.

Ryan M, Scott DA, Reeves C, et al. (2001) Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technol Assess* 5: 1-186.

Ryan M, Watson V and Amaya-Amaya M. (2003) Methodological issues in the monetary valuation of benefits in healthcare. *Expert Review of Pharmacoeconomics & Outcomes Research* 3: 717-727.

Sainsbury Centre for Mental Health. (2003) Sainsbury Centre for Mental Health. The economic and social costs of mental illness .Policy paper 3. London: Sainsbury Centre for Mental Health.

Sainsbury Centre for Mental Health. (2010) Centre for Mental Health.The economic and social costs of mental health problems in 2009/10. http://www.centreformentalhealth.org.uk/pdfs/Economic_and_social_costs_2010.pdf.

Saka m, McGuire A and Wolfe C. (2009) Cost of stroke in the United Kingdom. *Age and Ageing* 38: 27-32.

Schultz AB, Chen C-Y and Edington DW. (2009) The cost and impact of health conditions on presenteeism to employers: a review of the literature. *Pharmacoeconomics* 27: 365-378.

Schultz AB and Edington DW. (2007) Employee health and presenteeism: a systematic review. *Journal of Occupational Rehabilitation* 17: 547-579.

Schwappach DL and Boluarte TA. (2007) HEE-GER: a systematic review of German economic evaluations of health care published 1990-2004. *BMC Health Serv Res* 7: 7.

SCMH. (2007) Sainsbury Centre for Mental Health. Mental Health at Work: Developing the business Case. Policy Paper 8: London. http://www.centreformentalhealth.org.uk/pdfs/mental_health_at_work.pdf.

Sculpher M. (2001) The role and estimation of productivity costs in economic evaluation; in: Drummond MF, McGuire A (eds): Economic evaluation in health care: merging theory with practice. Oxford:: Oxford University Press, 94-112.

Severens JL, Laheij RJ, Jansen JB, et al. (1998) Estimating the cost of lost productivity in dyspepsia. *Alimentary Pharmacology & Therapeutics* 12: 919-923.

Severens JL, Mulder J, Laheij RJ, et al. (2000) Precision and accuracy in measuring absence from work as a basis for calculating productivity costs in The Netherlands. *Social Science & Medicine* 51: 243-249.

Shaw WS, van der Windt DA, Main CJ, et al. (2009) Early patient screening and intervention to address individual-level occupational factors ("blue flags") in back disability. *J Occup Rehabil* 19: 64-80.

Shinohara S, Okada M, Keira T, et al. (1998) Prognosis of accidental low back pain at work. *Tohoku J Exp Med* 186: 291-302.

Shiri R, Karppinen J, Leino-Arjas P, et al. (2010) The association between smoking and low back pain: a meta-analysis. *Am J Med* 123: 87 e87-35.

Smit F, Willemse G, Koopmanschap M, et al. (2006) Cost-effectiveness of preventing depression in primary care patients: randomised trial. *Br J Psychiatry* 188: 330-336.

Snaith RP. (2003) The Hospital Anxiety and Depression Scale. *Health and Quality of Life Outcomes* 1: 29.

Snedecor SJ, Botteman MF, Bojke C, et al. (2009) Cost-effectiveness of eszopiclone for the treatment of adults with primary chronic insomnia. *Sleep* 32: 817-824.

SOC. (2000) Standard Occupational Classification. www.ons.co.uk.

Soegaard R, Buenger CE, Christiansen T, et al. (2007) Circumferential fusion is dominant over posterolateral fusion in a long-term perspective - Cost-utility evaluation of a randomized controlled trial in severe, chronic low back pain. *Spine* 32: 2405-2414.

Steuten LMG, Bruijsten M and Vrijhoef HJM. (2007) Economic evaluation of a diabetes disease management programme with a central role for the diabetes nurse specialist. *European Diabetes Nursing* 4: 64-71.

Stewart WF, Lipton RB, Kolodner KB, et al. (2000) Validity of the Migraine Disability Assessment (MIDAS) score in comparison to a diary-based measure in a population sample of migraine sufferers. *Pain* 88: 41-52.

Stewart WF, Ricci JA, Chee E, et al. (2003a) Lost productive work time costs from health conditions in the United States: results from the American Productivity Audit. *J Occup Environ Med* 45: 1234-1246.

Stewart WF, Ricci JA, Chee E, et al. (2003b) Lost productive time and cost due to common pain conditions in the US workforce. *JAMA* 290: 2443-2454.

Streiner DL and Norman GR. (2006) *Health Measurement Scales: A Practical Guide to Their Development and Use*: Oxford University Press.

Sullivan MJL, Bishop SR and and Pivik J. (1995) The Pain Catastrophizing Scale: development and validation. *Psychological assessment* 7: 524-532.

Tang K, Pitts S, Solway S, et al. (2009) Comparison of the psychometric properties of four at-work disability measures in workers with shoulder or elbow disorders. *Journal of Occupational Rehabilitation* 19: 142-154.

Taylor SJ, Taylor AE, Foy MA, et al. (1999) Responsiveness of common outcome measures for patients with low back pain. *Spine (Phila Pa 1976)* 24: 1805-1812.

Thavorncharoensap M, Teerawattananon Y, Yothasamut J, et al. (2010) The economic costs of alcohol consumption in Thailand, 2006. *BMC Public Health* 10: 323.

Thomas CM and Morris S. (2003) Cost of depression among adults in England in 2000. *The British Journal of Psychiatry* 183: 514-519.

Thomas E, Silman AJ, Croft PR, et al. (1999) Predicting who develops chronic low back pain in primary care: a prospective study. *BMJ* 318: 1662-1667.

Thomas KJ, MacPherson H, Ratcliffe J, et al. (2005) Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain. *Health Technol Assess* 9: iii-iv, ix-x, 1-109.

Thomas L, Merete L, Karl Bang C, et al. (2006) Physical work environment risk factors for long term sickness absence: prospective findings among a cohort of 5357 employees in Denmark. *BMJ* 332: 449-452.

Torrance GW, Blaker D, Detsky A, et al. (1996) Canadian guidelines for economic evaluation of pharmaceuticals. Canadian Collaborative Workshop for Pharmacoeconomics. *Pharmacoeconomics* 9: 535-559.

Towse A. (2009) Should NICE's threshold range for cost per QALY be raised? Yes. *BMJ* 338.

Tranmer JE, Guerriere DN, Ungar WJ, et al. (2005) Valuing patient and caregiver time: a review of the literature. *Pharmacoeconomics* 23: 449-459.

Tsuchiya A and Williams A. (2001) Welfare economic s and economic evaluation In: Drummond M, McGuire A (eds) Economic Evaluation in Health Care Oxford University Press.

Turner J, Aaron A and Leslie A. (2001) Pain-Related Catastrophizing: What Is It? *The Clinical Journal of Pain Issue* 17: 65.

Turner JA, Franklin G and Turk DC. (2000) Predictors of chronic disability in injured workers: a systematic literature synthesis. *Am J Ind Med* 38: 707-722.

Turpin RS, Ozminkowski RJ, Sharda CE, et al. (2004) Reliability and validity of the Stanford Presenteeism Scale. *Journal of Occupational & Environmental Medicine* 46: 1123-1133.

U. K. Beam Trial T. (2004) United Kingdom Back Pain Exercise And Manipulation (Uk Beam) Randomised Trial: Cost Effectiveness Of Physical Treatments For Back Pain In Primary Care. *BMJ: British Medical Journal* 329: 1381-1385.

Uegaki K, Stomp-van den Berg SGM, de Bruijne MC, et al. (2011) Cost-utility analysis of a one-time supervisor telephone contact at 6-weeks post-partum to prevent extended sick leave following maternity leave in The Netherlands: results of an economic evaluation alongside a randomized controlled trial. *BMC Public Health* 11: 57.

van den Berg B, Brouwer W, van Exel J, et al. (2006) Economic valuation of informal care: Lessons from the application of the opportunity costs and proxy good methods. *Social Science & Medicine* 62: 835-845.

van den Brink M, van den Hout WB, Stiggelbout AM, et al. (2005) Self-reports of health-care utilization: Diary or questionnaire? *International Journal of Technology Assessment in Health Care* 21: 298-304.

van den Heuvel SG, Geuskens GA, Hooftman WE, et al. (2010) Productivity loss at work; health-related and work-related factors. *J Occup Rehabil* 20: 331-339.

van den Hout WB. (2010) The value of productivity: human-capital versus friction-cost method. . *Ann Rheum Dis* 69: i89-i91.

van den Hout WB, Goekoop-Ruiterman YPM, Allaart CF, et al. (2009) Cost-Utility Analysis of Treatment Strategies in Patients With Recent-Onset Rheumatoid Arthritis. *Arthritis & Rheumatism-Arthritis Care & Research* 61: 291-299.

van Hout BA, Al MJ, Gordon GS, et al. (1994) Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 3: 309-319.

van Ours J and Ridder G. (1991) Cyclical variation in vacancy durations and vacancy flows: An empirical analysis. *European Economic Review* 35: 1143-1155.

van Poppel MN, de Vet HC, Koes BW, et al. (2002) Measuring sick leave: a comparison of self-reported data on sick leave and data from company records. *Occup Med (Lond)* 52: 485-490.

van Roijen LH, van Straten A, Al M, et al. (2006) Cost-utility of brief psychological treatment for depression and anxiety. *Br J Psychiatry* 188: 323-329.

Van Schayck CP, Kaper J, Wagena EJ, et al. (2009) The cost-effectiveness of antidepressants for smoking cessation in chronic obstructive pulmonary disease (COPD) patients. *Addiction* 104: 2110-2117.

Van Tubergen A, Boonen A, Landewe R, et al. (2002) Cost effectiveness of combined spa-exercise therapy in ankylosing spondylitis: a randomized controlled trial. *Arthritis Rheum* 47: 459-467.

van Tulder M, Becker A, Bekkering T, et al. (2006a) Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 15 Suppl 2: S169-191.

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

van Tulder MW, Koes B and Malmivaara A. (2006b) Outcome of non-invasive treatment modalities on back pain: an evidence-based review. *Eur Spine J* 15 Suppl 1: S64-81.

van Tulder MW, Touray T, Furlan AD, et al. (2003) Muscle relaxants for nonspecific low back pain: a systematic review within the framework of the cochrane collaboration. *Spine (Phila Pa 1976)* 28: 1978-1992.

van Tulder MW and Waddell G. (2005) Evidence-based medicine for non-specific low back pain. *Best Practice & Research Clinical Rheumatology* 19: vii-ix.

van Zundert J and van Kleef M. (2005) Low back pain: from algorithm to cost-effectiveness? *Pain Pract* 5: 179-189.

Vlad I. (2003) Obesity costs UK economy £2bn a year. *BMJ* 327: 1308.

Vogt MT, Kwoh CK, Cope DK, et al. (2005) Analgesic usage for low back pain: impact on health care costs and service use. *Spine (Phila Pa 1976)* 30: 1075-1081.

von Korff M. (1999) Pain management in primary care: an individualized stepped care approach. In: Psychosocial Factors in Pain: Critical Perspectives, edited by R. Gatchel & D. Turk pp. . New York: : Guilford Press, 360 – 373.

Von Korff M and Moore JC. (2001) Stepped care for back pain: activating approaches for primary care. *Ann Intern Med* 134: 911-917.

Wahlqvist P, Carlsson J, Stalhammar NO, et al. (2002) Validity of a Work Productivity And Activity Impairment Questionnaire for patients with symptoms of gastroesophageal reflux disease WPAI-GERD: results from a cross sectional study. *Value Health* 5: 106-113.

Wahlqvist P, Guyatt GH, David A, et al. (2007) The Work Productivity and ActivityImpairment Questionnaire for Patients with Gastroesophageal Reflux Disease (WPAI-GERD). Responsiveness to Change and English Language Validation. *Pharmacoeconomics* 25: 385-396.

Walker BF, Muller R and Grant WD. (2003) Low back pain in Australian adults: the economic burden. *Asia Pac J Public Health* 15: 79-87.

Walker BF, Muller R and Grant WD. (2004) Low back pain in Australian adults: prevalence and associated disability. *J Manipulative Physiol Ther* 27: 238-244.

Wand BM and O'Connell NE. (2008) Chronic non-specific low back pain - sub-groups or a single mechanism? *BMC Musculoskelet Disord* 9: 11.
Ware JE, Jr., Kosinski M and Keller SD. (1996) A 12 Item Short Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Med Care* 34: 220-233.
Waxman R, Tennant A and Helliwell P. (2000) A prospective follow-up study of low back pain in the community. *Spine (Phila Pa 1976)* 25: 2085-2090.
Weiner SS and Nordin M. (2010) Prevention and management of chronic back pain. *Best Pract Res Clin Rheumatol* 24: 267-279.
Weinstein MC, Siegel JE, Garber AM, et al. (1997) Productivity costs, time costs and health-related quality of life: a response to the Erasmus Group. *Health Economics* 6: 505-510.
Weinstein MC, Siegel JE, Gold MR, et al. (1996) Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 276: 1253-1258.
Weinstein MC and Stason WB. (1977) Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 296: 716-721.
Weisbrod BA. (1961) The Valuation of Human Capital. *The Journal of Political Economy* 69: 425-436.
Whitehurst DG, Bryan S, Lewis M, et al. (2012) Exploring the cost-utility of stratified primary care management for low back pain compared with current best practice within risk-defined subgroups. *Ann Rheum Dis* 71: 1796-1802.
Whitehurst DG, Lewis M, Yao GL, et al. (2007) A brief pain management program compared with physical therapy for low back pain: results from an economic analysis alongside a randomized clinical trial. *Arthritis Rheum* 57: 466-473.
Wieser S, Horisberger B, Schmidhauser S, et al. (2011) Cost of low back pain in Switzerland in 2005. *European Journal of Health Economics* 12: 455-467.
Wilkinson D. (1999) Cost-Benefit Analysis Versus Cost-Consequences Analysis. *Performance Improvement Quarterly* 12: 71-81.
Williams A. (1983) The economic role of ‘health indicators’. In Measuring the Social Benefits of Medicine, Teeling SE (ed),Office of Health Economics: London; 63–67.
Williams A. (1985) Economics Of Coronary Artery Bypass Grafting. *British Medical Journal (Clinical Research Edition)* 291: 326-329.
Williams I, McIver S, Moore D, et al. (2008) The use of economic evaluations in NHS decision-making: a review and empirical investigation. *Health Technol Assess* 12: iii, ix-x, 1-175.
Wilson DA, Bork K, Shea EP, et al. (2010) Economic costs associated with acute attacks and long-term management of hereditary angioedema. *Ann Allergy Asthma Immunol* 104: 314-320.
Witt CM, Jena S, Selim D, et al. (2006) Pragmatic Randomized Trial Evaluating the Clinical and Economic Effectiveness of Acupuncture for Chronic Low Back Pain. *American Journal of Epidemiology* 164: 487-496.
Witter SE, T; Jowett, M and Thompson, R. (2000) Health Economics for Developing Countries: A practical guide. Malaysia: Macmillan Publication.
Wynne-Jones G, Mallen CD, Main CJ, et al. (2008) Sickness certification in general practice: a comparison of electronic records with self-reported work absence. *Primary Health Care Research & Development* 9: 113-118.

Wynne-Jones G, Mallen CD, Mottram S, et al. (2009) Identification of UK sickness certification rates, standardised for age and sex. *Br J Gen Pract* 59: 510-516.

Wynne-Jones GM, Christian D ; Main Chris J and Dunn Kate M. (2008) Sickness certification in general practice:a comparison of electronic records with self-reported work absence. *Primary Health Care Research & Development*

9: 113-118.

Yost KJ and Eton DT. (2005) Combining Distribution- and Anchor-Based Approaches to Determine Minimally Important Differences: The FACIT Experience. *Eval Health Prof* 28: 172.

Young A, Pransky G and van Mechelen W. (2002) Introduction to the Special Issue on Measurement of Work Outcomes. *Journal of Occupational Rehabilitation* 12: 115-117.

Zhang W, Bansback N and Anis AH. (2011) Measuring and valuing productivity loss due to poor health: A critical review. *Social Science & Medicine* 72: 185-192.

Zhang W, Bansback N, Boonen A, et al. (2012) Development of a composite questionnaire, the valuation of lost productivity, to value productivity losses: application in rheumatoid arthritis. *Value in Health* 15: 46-54.

Zhang W, Bansback N, Boonen A, et al. (2010a) Validity of the work productivity and activity impairment questionnaire--general health version in patients with rheumatoid arthritis. *Arthritis Research & Therapy* 12: R177.

Zhang W, Bansback N, Guh D, et al. (2008) Short-term influence of adalimumab on work productivity outcomes in patients with rheumatoid arthritis. *Journal of Rheumatology* 35: 1729-1736.

Zhang W, Gignac MAM, Beaton D, et al. (2010b) Productivity loss due to presenteeism among patients with arthritis: estimates from 4 instruments. *Journal of Rheumatology* 37: 1805-1814.

Zorgverzerkeringen Cv. (2006) Guidelines for Pharmacoeconomic Research, Updated Version. Diemen, The Netherlands: College voor Zorgverzekeringen (CVZ).