Productivity costs in economic evaluations

Marieke Krol

Productivity costs in economic evaluations

Marieke Krol

ISBN: 978-94-6169-327-3

Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Productivity costs in economic evaluations

Productiviteitskosten in economische evaluaties

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. H.G. Schmidt
en volgens besluit
van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 21 december 2012
om 11.30 uur

door

Hendrika Maria Krol
geboren te Hoogezand-Sappemeer

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor
Prof.dr. W.B.F. Brouwer

Overige leden
Prof.dr. C.D. Dirksen
Prof.dr. F.F.H. Rutten
Prof.dr. J.L. Severens

Contents

1.

Introduction

1.1 Introduction

The increase in health expenditures has raised important questions about the appropriate height of health care spending as well as the justification of these expenditures. One tool in the search of ensuring the optimal allocation of scarce societal and health care resources is economic evaluation of health care interventions, such as new pharmaceuticals, diagnostics or preventive measures. In economic evaluations, the costs of an intervention are compared to its benefits, expressed in some meaningful manner. Consistently applying these evaluations, in theory, would ensure an optimal level of spending in the health care sector (that is, the size of the budget, or how much to spend on health) as well as an optimal use of the available resources within the budget (that is, on *what* the budget is spent). This optimal spending can be defined in light of the twin goals of health care policy; efficiency and equity. (1) As such, economic evaluations can be seen as applied welfare economics, aimed at informing social choices to come to a maximization of broadly defined welfare. (2)

In recent years the role of economic evaluation in health care decisions has become more prominent. (3) For instance, pharmaceutical companies may be obliged to provide information on the costs and effects of new drugs when applying for coverage within a collectively funded health care system. While the influence of this information on subsequent decisions may still be limited (3-5), an increasing number of countries appear to wish to include this information in the process of reimbursement decisions (6). The increased use of economic evaluation in health care also increases the need to ensure a sound methodology of these evaluations. Obviously, this need increases with the influence of outcomes on subsequent decision making. While important developments have been achieved in the methodology of economic evaluations over the past decades (7), it is important to note that numerous methodological debates regarding how to best perform economic evaluations are, nonetheless, ongoing. Often, these disputes strike at the heart of economic evaluations and concern (normative) methodological choices that can have a very large influence on subsequent outcomes (and, potentially, decision making).

One of the areas of debate concerns (the inclusion of) productivity costs, which can be defined as *"Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid."* (8). This thesis deals with the methodology of including productivity costs in economic evaluations of health care interventions. It will address several issues related to the inclusion of productivity costs, as further highlighted below.

First, however, some background to the thesis will be provided by introducing the more general methodology of economic evaluations. Subsequently, the place of productivity costs in economic evaluations is highlighted, as well as important debates regarding whether and how to include them in the preferred type of economic evaluation (cost-utility analysis).

1.2 Economic evaluations

Economic evaluations[1] can be defined as *'the comparative analysis of alternative courses of action in terms of both their costs and consequences'* [(7)]. As such, they have a clear root in welfare economics, where changes in 'states of the world' (i.e. reimbursing some new drug or not) are compared in terms of their effects on welfare. Hence, an economic evaluation must always compare at least two alternative courses of action (or states of the world), by which it can be determined whether moving for instance from the current state (not yet funding the drug) to an alternative state (funding the drug) would improve welfare. As illustrated in figure 1.1, in performing economic evaluations, the costs and benefits (often confined to health effects) of an old intervention are commonly compared with the costs and health effects of a new intervention. In doing so, economic evaluations aim to provide information on whether the improvement in health status justifies the additional resources required for the new intervention as compared to the old intervention.

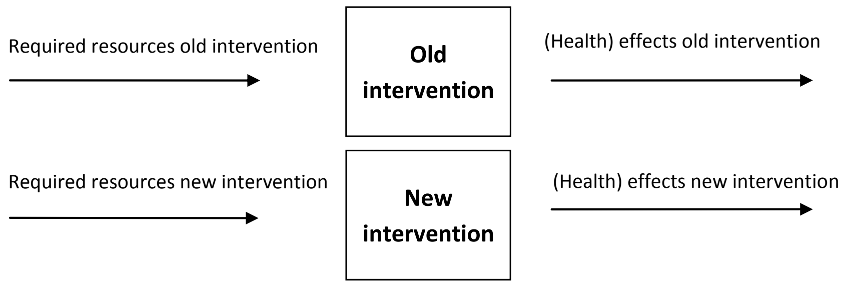

Figure 1.1 A graphical representation of an economic evaluation.

[1] Although strictly speaking these terms are not necessarily synonyms, in this thesis the terms 'economic evaluation', 'health technology assessment', 'cost-effectiveness study' and cost-effectiveness analysis' are used interchangeably.

Comparing health interventions can be done in several ways. Four types of economic evaluations are commonly distinguished. All four express costs in monetary terms, but differ in whether and how they express the benefits. The first type of economic evaluation is the traditional cost-benefit analysis in which both costs and effects are expressed in monetary terms. This type of evaluation is relatively uncommon in the field of health care, which appears to have to do with difficulties of finding an appropriate monetary valuation of health gains as well as the fact that 'pricing life' may be considered undesirable (or even 'unethical'). This paved the way for a second type of economic evaluation: cost-effectiveness analysis. In such an analysis costs are still expressed in monetary terms but the effects are expressed in natural (clinically relevant) units such as hip fractures avoided, number of successfully treated patients, or life years gained. While this type of economic evaluation relates well to clinical practice, it makes societal decisions difficult due to the incomparability of outcome measures and results. Hence, the third type of economic evaluation, cost-utility analysis, became popular. This is a specific type of cost-effectiveness analysis in which effects are expressed in quality adjusted life years (QALYs). The QALY is a comprehensive generic preference-based health measure encompassing both length of life and quality of life, as further explained below. The fourth type is cost-minimization analysis (CMA). In such an analysis only costs are considered, not health effects. The use of CMA is only appropriate when there is no (significant) difference in terms of health benefits. [(7)]

Cost-utility analysis

Cost-utility analysis currently appears to be the preferred type of economic evaluation [(7)], since it allows comparing outcomes of economic evaluations across diverse interventions, without expressing them in monetary terms. Rather, the effects are expressed in terms of QALYs. To calculate QALYs life years are multiplied with the quality of life value during those years. These values are commonly referred to as QALY weights. QALY weights are anchored at full health (or best imaginable health state) and the state 'dead'. The value for full health is normalized at '1', while that for 'dead' is normalized at '0'. One year in perfect health thus equals 1 QALY, while one year in the state 'dead' equals 0 QALYs. Likewise, (abstracting from discounting) ten years in perfect health represents ten QALYs. The values of all other health states are determined based on their relative position compared to full health and dead. States considered to be 'worse than dead' are assigned values below zero. [(7)] To illustrate, a simple QALY calculation looks as follows. If an individual will live for an additional 30 years with a quality of life of 0.8, the remaining life time QALYs of that individual would equal 0.8 * 30 = 24. If that person is treated for some illness, improving his health to 0.9 for the full 30 year period, this means an improvement of 3 QALYs (since he now will live 30*0.9 = 27 QALYs).

The outcomes of cost-utility analyses are commonly expressed as the costs per QALY gained, or, correspondingly, the incremental cost-utility ratio (ICUR)[2]. The ICUR is calculated by dividing the difference in costs by the difference in effects of the old and new intervention. This is expressed in equation 1.1, where B represents the new health intervention and A the old intervention.

Equation 1.1

$$\text{ICUR} = \frac{(\textit{Costs B} - \textit{Costs A})}{(\textit{QALYs B} - \textit{QALYs A})}$$

The comparability of ICURs is a pivotal advantage over outcomes of cost-effectiveness analyses, since decision makers typically need to make choices between different technologies and at different points in time. In relation to productivity costs two important issues need to be discussed regarding equation 1.1: (i) which costs are relevant to include and (ii) are productivity costs (best) placed in the numerator or denominator of the equation. These questions are addressed below.

Defining relevant costs

Costs in economic evaluations of health care interventions can be divided in direct medical and non-medical costs and indirect medical and non-medical costs. Direct medical costs refer to health care costs directly associated with the illness and treatment of interest. Examples of direct medical costs are drug costs, costs of health care personnel and hospital costs. Direct non-medical costs refer to costs that accrue to patients and their families while receiving health care, such as costs of traveling to the hospital, but these also include time-costs, such as the time significant others spent on providing informal care. Indirect medical costs refer to medical costs indirectly associated with the illness and intervention of interest, for instance medical costs occurring in life years gained as a result of the intervention. Indirect non-medical costs are costs emerging outside the health care system and indirectly associated with

2 In this thesis instead of ICUR, the more general term incremental cost-effectiveness ratio (ICER) is regularly.

the illness and related intervention. These costs are commonly restricted to productivity costs.[2]

Importantly, not all economic evaluations include all cost-categories. The issue of which costs to include in economic evaluations is strongly related to the perspective adopted in the evaluation. Most commonly, economic evaluations are performed from either the health care (payer) perspective or the societal perspective. Evaluations taking a health care perspective limit the costs included to costs falling on the health care budget (i.e. medical costs). By contrast, evaluations from a societal perspective aim to include all relevant societal costs (and effects), regardless of where these costs fall [7]. There seems to be no consensus on which perspective to take in economic evaluations. The normative choice of the appropriate perspective relates to the assumed context of health care decision making. If the main aim of economic evaluations is considered to be informing decision makers who wish to maximize (or optimize) population health from a given health care budget, only costs falling on the health care budget are relevant. Costs falling outside the health care sector, including productivity costs, then are irrelevant. However, if it is assumed that the decision maker has the broader objective of contributing to maximizing social welfare, costs falling outside the health care budget can be considered equally important as those falling on the health care budget. Then, obviously, productivity costs are fully relevant, and included in the analysis. At present, there is no theoretical or practical consensus on which perspective is most appropriate. This lack of consensus regarding the appropriate perspective in economic evaluations, and the relevance thereof for productivity costs, is further explained and discussed in chapter 2 of this thesis.

Inclusion of productivity costs

Even when a societal perspective is taken, productivity costs are not systematically included in economic evaluations. This is disturbing since productivity costs may strongly affect cost-effectiveness outcomes and their absolute height can exceed medical costs [9-11]. The reasons for omitting productivity costs are diverse. First, concerns have been raised regarding the ethical implications of including productivity costs [12,13]. Inclusion of these costs could lead to favoring interventions targeted at the working population, since health interventions successfully increasing the (paid) productivity of patients could result in substantial societal savings. Consequently, these savings could result in more favorable ICURs than equally effective treatments targeted at older patients.

Second, there is dissensus regarding the appropriate methodology for productivity cost inclusion. Many debates, which are far from being settled, have focused on the question how to derive valid productivity cost estimates for use in economic evaluations. The most important debates are related to the three main aspects of costing methodology in economic evaluations: identification, measurement and valuation of costs [(7)]. These debates are recaptured in this thesis and some issues are empirically investigated and discussed more extensively. Examples of ongoing productivity cost debates are the relevance of productivity cost inclusion in (individual) economic evaluations, the appropriate valuation approach, and the identification of commonly ignored relevant health related productivity changes, such as the effects of ill-health on unpaid labor and effects on coworkers' productivity. Most commonly, productivity costs are estimated by placing a wage-based value on an ill worker's absenteeism and (occasionally) presenteeism (i.e. productivity losses related to being ill at work). However, it has been suggested that common productivity cost estimates are, in fact, underestimations of actual costs [(14-16)], since these do not account for the negative effect of absenteeism and presenteeism on coworkers' productivity in case of team-dependent production. At most, an entire team's productive output can be at risk when one of its members experiences health problems. Oppositely, it has been suggested that common estimates may be an overestimation of true productivity costs [(17,18)], since productivity losses may partly be compensated during normal working hours. Although compensation of lost productivity in normal hours is not costless 'true' productivity costs may then be lower than common productivity cost estimates.

Third, productivity cost inclusion in economic evaluations may be hampered by practical limitations. In some particular circumstances (e.g. when researchers are highly dependent on retrospective data) it may not be possible to collect data on the effects of ill-health and treatment on patients' productivity. Currently, in such situations productivity costs are typically ignored regardless of whether the inclusion of these costs would be appropriate. Given the potential effect of productivity costs on cost-effectiveness outcomes, excluding these costs may lead to a misrepresentation of actual costs and benefits and hence to misinforming decision makers. It has been suggested that in cases where collecting productivity data proves impossible, productivity may be predicted based on information on quality of life. Although the significant correlation between productivity and quality of life has been previously investigated [(19,20)], it is unclear whether adequately predicting productivity based on quality of life data is possible.

Productivity costs in QALYs?

One of the unresolved productivity cost issues that stirred quite some debate is whether productivity costs should be captured as costs or in terms of QALYs [2,8,21-23]. Importantly, QALY weights for use in economic evaluations are obtained by means of valuation surveys among the general public. These weights are thus usually based on societal preferences regarding health states rather than patients' preferences [2]. In such valuation surveys, a representative sample from the general public is asked to value a variety of distinct health states. These health states are often described by using level and domain descriptions of generic quality of life questionnaires, for instance the EQ-5D [24] or the Health Utility Index [25]. The health states are commonly valued with techniques such as the Visual Analogue Scale, the Standard Gamble method or the Time Trade-Off method [26-30]. Based on the respondents' answers, a value set can be constructed producing QALY weights for all possible health states described with the relevant generic quality of life instrument. Using such QALY weights, health gains due to interventions can be quantified as QALY gains and used in cost-utility analyses.

What the health state values underlying QALY gains exactly represent is unclear. In a typical Time Trade-Off exercise to obtain health state valuations, respondents are required to make trade-offs between living in a better health state for a shorter period or in a worse health state for a longer period of time. It has been suggested that respondents in such exercises may not only consider health itself but also broader effects, such as effects of ill-health on income [2]. Traditionally, productivity costs (if included) in economic evaluations were commonly included in monetary terms on the cost side of the cost-utility ratio. In 1996 the Washington Panel, however, stated that productivity was already included on the effect side of the ratio (that is, in health state values), and should therefore be excluded from the cost side to avoid double counting of the effect of ill-health on productivity and income. Although this recommendation was strongly criticized [22,31], it was never empirically investigated whether respondents to health state valuation exercises in fact included income effects.

1.3 Productivity costs questions addressed in this thesis

The main questions addressed in this thesis relate to unresolved issues that have emerged from earlier debates regarding productivity cost inclusion in economic evaluations. These questions are:

1. Do respondents to health state valuations include the effects of ill-health on income and if so, how does such inclusion affect valuations?
2. How common is productivity cost inclusion in economic evaluations?
3. How does productivity cost inclusion or exclusion affect cost-effectiveness outcomes?
4. How is reduced productivity related to ill-health compensated and how do compensation mechanisms and multiplier effects potentially affect productivity cost estimates?
5. Can productivity losses be predicted based on quality of life estimates?

By answering these questions, this thesis seeks to contribute to the ongoing debates regarding theoretical and practical aspects of productivity cost inclusion in economic evaluations. In doing so, it aims to stimulate the process towards standardizing productivity cost inclusion and methodology.

1.4 Outline of this thesis

Chapter 2 provides an overview of the most important past and present debates and developments regarding productivity cost inclusion, identification, measurement and valuation. Moreover, this chapter describes important underexplored research questions for future research regarding productivity costs in economic evaluations. The future research topics are based on present-day issues discussed in recent literature.

Chapters 3, 4 and 5 describe empirical work on the question whether productivity costs are or even should be included on the effect side of the cost-effectiveness ratio. It is investigated whether respondents to health state valuations take the effects of ill-health on income into account when participating in health state valuation exercises. Furthermore, it is investigated how spontaneous inclusion or exclusion of income-effects influences final health state valuations. Moreover, the effects of explicitly instructing respondents to either include or exclude income-effects are explored. The effects of such instructions are important to identify, considering that if respondents to health state valuations do not consistently include or exclude income effects spontaneously, explicit instructions may be required to ensure consistency.

In chapters 6 and 7 insight is provided in the current role of productivity costs in economic evaluations. It is investigated to what extent productivity costs are included in economic evaluations. Moreover, in evaluations where these costs are included it is examined how these costs were identified, measured and valued. Additionally, the amount of productivity costs within studies including these costs is discussed (both in absolute terms and as a fraction of total costs) as is the impact of including productivity costs on cost-effectiveness outcomes.

In chapter 8 it is discussed whether common productivity cost estimates accurately reflect actual productivity costs. In this chapter it is investigated how ill workers' absenteeism is compensated. Additionally, it is explored how compensation mechanisms and effects of ill-health on team workers potentially affect common productivity cost estimates.

In chapter 9 of this thesis it is explored whether patients' productivity levels can be estimated based on EQ-5D quality of life data. It does so by asking respondents in an online survey about whether they would be able to attend to work in different health states and if so, how they would expect to function in that health state. Subsequently, based on these responses two prediction models are presented to enable productivity predictions based on EQ-5D information. The validity of the models is tested and discussed by means of comparing predictive productivity levels with measured productivity levels in a group of patients suffering from low back pain.

Chapter 10 draws together the results of the previous chapters. It provides a discussion of the results and explores the implications and the limitations of this thesis.

To note, the chapters of this thesis are based on research articles published, or planned to be published in scientific peer reviewed journals. As a result, the chapters of this thesis can be read independently and some overlap exists between some of the chapters.

2.

Productivity costs in economic evaluations: Past, present, future

Based on: Krol M, Brouwer WBF, Rutten FFH.

Productivity costs in economic evaluations: Past, present, future.

Accepted for publication in Pharmacoeconomics.

Summary

The inclusion of productivity costs and its respective methodologies have been fiercely debated in the literature. Productivity costs occur when the productivity of individuals is affected by illness, treatment, disability, or premature death. The main debates and divides in theory and practice of economic evaluations of health technologies have centered on the questions of *whether* and *how* to include productivity costs, especially productivity costs related to paid work. We focus on past and current developments related to productivity costs and address its role in economic evaluations. We highlight how debates in the area have evolved and where we currently stand in theory and practice of including productivity costs in economic evaluations. The objectives of this chapter are (i) to discuss the main debates and developments regarding the inclusion, identification, measurement, and valuation of productivity costs; and (ii) to summarize the most topical unresolved issues regarding productivity costs in economic evaluations, and (iii) to encourage future research in the area.

2.1 Introduction

Economic evaluations are increasingly used to aid decision makers in allocating scarce health care resources. Given the direct influence the results of economic evaluations may have on health care decisions, ensuring a sound methodology is critical. Many fierce debates regarding the methodology of economic evaluations are ongoing. Importantly, these debates normally do not relate to methodological choices that marginally affect final outcomes of economic evaluations. In many cases, their influence can be profound. Examples of such debates relate to the normative foundations of economic evaluations, the perspective of economic evaluations (i.e., what costs and benefits to include), how to discount future costs and benefits, and the nature and height of the threshold value used in judging cost-effectiveness results (e.g. [32-38]). One area of debate and dissensus closely related to the appropriate perspective of economic evaluations is the inclusion of productivity costs and the methodologies used to measure and value them. Productivity costs can be defined as the costs associated with paid and unpaid production loss and replacement due to illness, disability, or death of productive persons.[8] The main debates and divides in theory and practice of economic evaluations of health technologies have centered on *whether* and *how* to include these costs, especially those related to paid work.

Several important topics regarding the questions above have been recently reviewed and discussed by Zhang et al.,[39] who pointed out, '*more attention should be paid to the methodologies of measuring and valuing productivity loss,*' since many concerns regarding productivity costs are yet to be resolved. Table 2.1 presents a summary of important questions addressed in the literature and a selection of literature references pertaining to them, highlighting the attention the topic has captured.

In this chapter, we focus on past and current developments related to the inclusion of productivity costs and address their inclusion in economic evaluations. This highlights how the debates in this area have evolved and where we currently stand in the theory and practice of including productivity costs in economic evaluations. The objectives of this chapter are (i) to discuss the main debates and developments regarding the inclusion, identification, measurement, and valuation of productivity costs; and (ii) to summarize the most topical unresolved issues regarding productivity costs in economic evaluations, and (iii) to encourage future research in this area.

Table 2.1 Important aspects related to the estimation of productivity costs within economic evaluations

Aspects	Main questions	Selection of literature discussing these issues
Inclusion of PC	Should productivity costs be included in economic evaluations?	
Ethics	Is it ethical to include productivity costs?	(12,13)
Perspective	Which perspective should be taken in economic evaluations?	(2,7,33-35,40,41)
Relevance	Are productivity costs relevant to include?	(9,42-45)
Identification	How can health related productivity changes be identified?	
Input and output	How do health changes affect productive input and output?	(46,47)
Paid and unpaid labor	How can and should productivity changes in paid and unpaid labor be identified?	(43,46,48-50)
Ill workers and coworkers	How can the effects of productivity changes on coworkers be identified?	(14-16)
Measurement	How can productivity changes best be measured?	
Data collection	How and when should data on productivity be collected?	(39)
Content of instruments	Which productivity related information should be collected?	(17,39,51)
Availability of instruments	Which instruments are available and how is their performance?	(52-54)
Indirect measurement	Is it possible to adequately measure productivity changes indirectly?	(19,20,55)
Valuation	How can productivity changes best be valued?	
Approach: theory	Which valuation approach should be applied?	(8,11,21-23,47,56-59)
Approach: empirical	Are income effects already captured in the QALY?	(60-67)
Values	Which values should be placed on productivity changes?	(46,68)
Reporting outcomes	How should productivity cost estimates be reported in economic evaluations?	(41,44)

PC = productivity costs
QALY = quality adjusted life year

2.2 Inclusion

The question of *whether* to include productivity costs has been addressed in the literature mainly by focusing on the ethical consequences of inclusion (e.g.[12,13]) and taking the appropriate perspective in economic evaluation (e.g. [33-35,40,41]). Moreover, whether the costs are relevant [9,23,42-45] is a matter of debate.

Ethical concerns

Some authors (e.g. Williams and Olsen and Richardson [12,13]) have raised concerns regarding the equity implications of including productivity costs in economic evaluations. One issue is that health care programs aimed at productive people (especially in exchange for pay) can produce substantial societal gains due to improved productivity. Inclusion of these costs would therefore result in more favorable cost-effectiveness results of interventions aimed at employed persons than those aimed at the unemployed or retirees. This could, in turn, lead to fewer resources allocated to interventions aimed at unemployed individuals, which could be considered inequitable. While inclusion of productivity costs can indeed have such distributional consequences, several issues temper the subsequent conclusion that their outright exclusion is warranted.

First, economic evaluations are commonly used to inform decisions on reimbursement of interventions through the inclusion in some basic insurance package (which covers all individuals, regardless of productivity). They are normally not used to decide *which groups* have access to the interventions within the total patient group.

Second, inclusion of productivity costs does not necessarily lower cost-effectiveness estimates or favors interventions for productive individuals.[9,43-45] When interventions require substantial time from productive individuals (e.g., intensive psychotherapy), inclusion of productivity costs may cause interventions to become *less* cost-effective and the treatment of individuals without paid work would thus be more cost-effective than that of those with paid work.[9,44]

Third, complete exclusion of costs based on equity considerations seems difficult and dangerous, since many types of costs (including medical) may be perceived to 'discriminate' between different groups (e.g., young and old). Therefore, health policy makers have to balance the desirable effects of freeing additional resources due to productivity gains with the equity implications of including productivity costs. As recommended (e.g. by Brouwer et al. and Pritchard and Sculpher [41,43]), final cost-effectiveness ratios may best presented with both scenarios, i.e., with and without productivity costs. Moreover, as recently argued,[69] equity considerations should have a prominent place in deliberations on final policy decisions based on results of economic evaluations.

Perspective

The debate on in- or excluding productivity costs in an economic evaluation is closely related to the issue of the appropriate perspective to take in economic evaluations and (therefore) the relevant decision context and decision rule [2]. Two prominent perspectives are the health care perspective and the societal perspective. Evaluations performed from a health care perspective do not include productivity costs, since these costs (or savings) do not fall within the health care budget. Economic evaluations conducted from the societal perspective, however, include all costs – including productivity costs – whenever relevant.

The important normative choice of perspective relates to the context of economic evaluations (e.g., health care) and the objective(s) of the decision maker. (See Claxton et al. and Brouwer et al.) [33,41] for more elaborate discussion.) If we assume that economic evaluations are to inform a decision maker with the aim of maximizing health from a given health care budget, costs falling outside of that budget (and non-health benefits) can be left out of the evaluation. Influential textbooks (e.g. [2,7]), however, suggest that taking a broader, societal perspective in economic evaluations is appropriate. Such thinking furthermore seems to align with the welfare economic roots of economic evaluations.[70] Adopting a societal perspective all costs and effects directly or indirectly induced by the intervention are incorporated (when significantly present) in the evaluation, *regardless of where the burden (or benefit) falls.* [2] This relates to the decision maker's broader underlying objective of maximizing social welfare.[34] When adopting such a perspective, productivity costs are clearly relevant and should be included in the evaluation.

While the discussions regarding appropriate perspectives are ongoing at the theoretical level, different viewpoints have been taken in practice. For instance, when investigating the various national guidelines for (pharmaco-) economic evaluations, we notice that many countries (e.g., the UK, Belgium, New Zealand [71,72]) require analysts to take a narrower, health care perspective, encouraging analysts to ignore costs outside the health care sector, including productivity costs (or at least not to report them). Other countries (e.g., Sweden and the Netherlands [73,74]) do prescribe a societal perspective and thus require analysts to consider productivity costs whenever relevant.

Although practical and theoretical consensus regarding this important issue of perspective remains out of reach, recent developments in the literature may represent steps toward it. Brouwer et al.[41] have attempted to reconcile both viewpoints to some extent. They argued that, while productivity costs represent real societal impacts on welfare and should not be systematically ignored in economic evaluations, they can carry a different weight in health care decisions because their opportunity costs may be different from costs falling on the health care budget. Moreover, for costs outside their budget health care policy makers are not directly or formally responsible, and may weigh any adverse equity implications that productivity cost inclusion may have, indicating that the context in which health economic evaluations are performed is relevant. This does not imply, however, that a health care policy maker should be left ignorant of costs and savings external to the health care sector that occur due to actions in the health care sector. As several authors (e.g. [75,76]) have argued, a health care program improving productivity of patients produces additional savings to society that free resources for welfare-improving purposes, which may include health care. Such a program therefore uses fewer societal resources to produce the same benefit than a comparable program without improved productivity or, when the savings are used to produce additional health benefits (e.g., added to the health care budget[3]), it leads to higher benefits for the same health care costs. While these broader costs and benefits may be viewed as externalities from the viewpoint of a health care decision maker, having decision makers ignorant of them may result in non-optimal decisions from a societal perspective.(34,40) In that sense, Brouwer et al.(41) refer to the concept of "local optimality." (77) The puzzle of maximizing welfare is too complex for single decision making bodies; all solve a part (education, health care, defense, and so on) of the puzzle, each with its own budget and objectives. Such local optimality, however, does not necessarily translate into an overall optimal outcome, especially when externalities exist. A two-perspective approach was therefore proposed as a new standard in economic evaluations, whereby the cost-effectiveness ratios from *both* the health care and the full societal perspective are presented. (41) Such an approach would highlight discrepancies (i.e., when an intervention is deemed cost-effective from a societal perspective but not from a health care perspective and vice versa) that inspire further attention. The approach also allows decision makers to attach more weight to some costs than to others. In some ways, the suggestion aligns with recent work on appropriate perspectives (33) in the sense that costs outside the health care sector could be accounted for but weighted differently.

The two-perspective approach is already recommended in several national health economic guidelines (e.g., Italy and Norway.(78,79)) Moreover, numerous country guidelines prescribing a health care perspective for the base-case analysis allow presenting results from additional perspectives. Whether and how this affects (or improves) decision making, however, remains unclear. In general, balancing outcomes from evaluations using different perspectives may be a demanding task for decision makers. Future research could be aimed at investigating (i) how to support decision makers in this context, (ii) the extent to which external costs and savings are important, and (iii) whether they can be directly or indirectly transferred into health care funds.

3 Reallocation of savings outside the health care sector to the health care budget is likely to be difficult to accomplish in practice. (43) Such reallocation would, at the least, involve transaction costs (33)

Relevance

While systematically ignoring productivity costs seems hard to defend, their inclusion is not necessarily influential. In general, the inclusion of productivity costs related to paid work is especially relevant when the intervention is targeted at patients of working age or younger patients whose future productivity is at stake.[4] Moreover, productivity cost inclusion is only relevant if an intervention indeed affects patients' work ability and, in turn, productivity. Interventions may not affect productivity in cases involving very mild conditions (patients continue to work) or very severe conditions (patients work neither before nor after treatment). Finally, for productivity costs to affect incremental cost-effectiveness, the intervention and comparator should affect productivity differently. If we can reasonably expect cost-effectiveness outcomes to be unaffected by productivity costs, they can be ignored even when adopting a societal perspective.(45)

Hence, in some cases productivity cost inclusion seems irrelevant. However, a decision to exclude productivity costs in an economic evaluation should be taken with caution. Several studies have shown that, if work ability of productive, working age individuals is indeed affected, the choice of including or excluding productivity costs in economic evaluations can strongly influence both incremental costs and incremental cost-effectiveness (e.g. (9,42,43,45)). Nyman has recently suggested that excluding productivity costs in the friction cost approach may not substantially bias incremental cost-effectiveness estimates since "*under the frictional approach, the omitted costs are likely to be relatively small.*"(23) We urge caution here, since it seems that in practice such biases can still be considerable. Krol and colleagues have shown that even with the more conservative friction cost estimates, productivity costs on average reflected 23% of total costs in studies considering expensive drugs administered in a hospital setting (45) and 56% among interventions targeted at depressive disorders.(44) The precise impact of productivity cost inclusion varied strongly in individual economic evaluations, both as a fraction of total costs and incremental cost-effectiveness outcomes. Nevertheless, the strong potential impact of productivity costs on cost-effectiveness outcomes confirms that structurally ignoring them may lead to suboptimal decision making (in terms of welfare) and inefficient use of societal resources.

4 Whether it is relevant to include potential effects on future productivity depends on the valuation approach. If the friction cost approach is adopted, these costs normally need not be included, since they will commonly fall beyond the friction period. Their inclusion is important, however, when using the human capital approach.

In sum, while there may be reasons for decision makers to assign different weights to productivity costs depending on where they fall, a case for their full and systematic exclusion is hard to defend. The question of *how* these costs need to be included is then the relevant one.

2.3 Identification of factors influencing productivity costs

To be able to identify health-related productivity changes it is important to understand where and when they can emerge. In general, the several ways are: (i) lower production levels and hence welfare decrease; (ii) higher cost of production levels (e.g., hiring additional labor); or (iii) some combination of both (e.g., when a replacement is not as productive as the worker replaced).[47] In this section we highlight the identification of relevant factors driving productivity costs.

Paid labor

In the context of paid labor, several aspects need to be addressed.

Absenteeism

Traditionally, patients having paid work *and* experiencing absenteeism due to illness, disability, or premature death would be a reason to consider productivity costs in economic evaluations. Absenteeism from paid work clearly represents a major source of productivity costs as emphasized in the literature.

Presenteeism

Reduced productivity at work, also called presenteeism, can be important. Presenteeism may occur without absenteeism, but can also precede absenteeism (e.g., a progressive illness whose onset is mild) or follow absenteeism (e.g., a partial recovery that allows return to less productive work).[20] For some illnesses (e.g., depression or migraine) presenteeism may be particularly important and may, in fact, be more important (cost-wise) than absenteeism. Moreover, we should note that in jurisdictions where sickness benefits for absenteeism are less generous, presenteeism (and its related costs) may be more prominent.

Compensation mechanisms

Not all reduced productivity results in reduced production. Compensating for lost work is quite common.[17,18,80] For instance, an ill worker may compensate for lower productivity in normal or extra hours (potentially after returning to work); colleagues may take over some or all tasks (in normal or additional hours); new personnel may be hired to compensate for losses. Compensation mechanisms (even those involving compensation during normal work hours) are not costless, but little is known about their actual costs.

Multiplier effects

Productivity losses in one individual may negatively affect coworkers' productivity in case of team-dependent production. The productive output of a full team can be jeopardized by one member's illness. This is especially relevant when substitutes are less equipped or unavailable. The effect of ill health on coworkers' productivity ('the multiplier effect' (14-16)) can be non-negligible.

Unpaid labor

That productivity costs involve costs related to paid *and* unpaid labor is often ignored. Productivity costs related to unpaid work are rarely included in economic evaluations and, as a consequence, have little influence in decision-making processes. From a societal viewpoint productivity losses related to unpaid labor are important. Swiebel (81) defines unpaid labor as *"all productive activities outside the official labor market done by individuals for their own households or for others."* Inter alia, this includes household work, caring for significant others, and volunteer work. Productivity costs related to unpaid work may result from loss of unpaid production or replacement costs. Little is known about lost unpaid work involving absenteeism and presenteeism. Along with the growing number of interventions targeted at the elderly, unpaid productivity should have a more prominent role in economic evaluations.

In sum, productivity costs related to paid and unpaid work are complex. They involve costs related to lost production or replacement costs due to absenteeism and presenteeism. Diminished productivity of ill workers can negatively affect coworkers' productivity, increasing productivity costs. Compensation mechanisms can reduce production losses, but their costs are underexplored.

2.4 Measurement

Once it is clear what needs to be included in the analysis (identification), the next issue is to measure the relevant quantities. In this section we highlight measurement of different aspects related to productivity costs.

Paid productivity

A sound estimation of productivity costs requires sound measurement of the relevant components. Unfortunately, there is an apparent lack of standardization of measurement methods for productivity costs and no consensus on the best instruments to reliably capture changes in productivity. Ideally, information is also gathered on the effects of productivity changes on coworkers and the relevance and costs of compensation mechanisms. When using the friction cost method (section 5), total absenteeism, its frequency and the duration of each separate period of absenteeism must be captured. Actual, objective measurement of productivity changes (absenteeism and presenteeism) at the work sites of the relevant patients is normally impossible in the context of health economic evaluations. Commonly, patients are required to self-assess diminished productivity in terms of absenteeism and, occasionally, presenteeism. The reliability of such measures has not often been tested.(82)
Many instruments (mostly written questionnaires suitable for self-administration by ill workers) have been developed to measure health-related productivity changes. Selecting an appropriate instrument among the available instruments is challenging.(52) Empirical research has shown that the use of different instruments can lead to large differences in outcomes.(83) Such differences are worrisome, since they seriously hamper comparability of productivity cost outcomes and can lead to justifiable doubt regarding the validity of productivity cost estimates. Differences in outcomes secondary to instrument choice are likely to be influenced by how questions are framed and the applied recall period. Zhang et al.(39) pointed out in a recent review that outcomes of scientific research on the appropriate recall period for absenteeism and presenteeism are inconclusive. Still, given the available evidence, they recommend applying a three-month recall period for questions regarding absenteeism and a one-week time period for questions related to presenteeism. The feasibility of different time-frames within one questionnaire for self-completion, however, deserves attention.

It is not clear which of the currently available instruments provides the best estimates of absenteeism and presenteeism.(39) Most instruments focus on a particular part of productivity losses, such absenteeism. Other, typically more recently developed instruments, capture both diminished productivity related to presenteeism and absenteeism. (For reviews of productivity instruments see [52-54].) Few instruments seem to include measurement of unpaid productivity losses by, for instance, questioning the ability to perform household tasks (e.g. the Health and Labor Questionnaire [84]). Furthermore, the validity of the answers to such questions has rarely been investigated. Zhang et al.[85] recently developed a new productivity cost questionnaire, which seems to be the first to combine absenteeism, presenteeism, paid and unpaid labor, job characteristics, compensation mechanisms, and team dynamics. While being the most comprehensive instrument to date, to our knowledge, its validity needs to be further explored.

Some elements relevant to productivity costs, especially compensation mechanisms and multiplier effects, are inherently difficult to measure by patient-completed questionnaires. Patients may not have sufficient knowledge of how their health problems affect team output and how their diminished productivity is compensated for. Employees and their supervisors appear to provide different answers on how health-related absenteeism is compensated.[18] Regarding the effects of illness on coworkers and team output some attempts have been made to construct 'job-dependent multipliers' that take account of the (average) effect of co-worker absenteeism and presenteeism in specific job types. The idea behind them – which is worthy of future research – is that if multipliers are robust and transferable it would be possible to include the effects on coworkers in economic evaluations solely based on information of the ill worker's job-type. Investigating how compensation of health-related productivity losses interacts with the effects of the losses on coworkers' productivity is also important.

Another measurement difficulty is deciding on how often to repeat measurement. The appropriate frequency obviously depends on the type of intervention and when and how often changes in productivity are expected. A trade-off must be made between the costs of measurement, study compliance, and accuracy of outcomes. Accuracy is likely to increase with the number of measurements but at the expense of evaluation costs and patient burden, the latter of which is likely to negatively affect compliance rates. In this light, it is important to note that it is unclear how to estimate productivity when measurement gaps exceed the recall period.

Unpaid productivity

Two approaches seem most commonly used in measuring health-related productivity changes in the context of unpaid labor.[2,7] The first measures the changes in time spent on unpaid labor and the second measures the additional time others spend on unpaid labor tasks not performed by the patient due to illness. Both approaches have difficulties. In the first, distinguishing between time spent on unpaid labor and leisure time is challenging. Time spent with one's children, for example, may be considered leisure time but it is also childcare; should the parent become ill someone else would have to look after them, a task that, at least in part, could be considered unpaid work. In 1934 Reid [86] introduced the 'third person criterion', whereby all output replaceable by a third person can be considered unpaid labor. Elements people cannot take over (e.g., the enjoyment of playing tennis) are then considered leisure. The distinction is useful, but can be difficult to apply in practice. Especially tasks such as caring for children may involve elements of both leisure and unpaid labor, which may be difficult to disentangle.

This difficulty is avoided to some extent by applying the second approach, measuring the time spent by others who take over the patient's unpaid work. In that case, however, all tasks that are simply forgone or later compensated for remain unvalued, potentially leading to an underestimation of productivity loss related to unpaid labor. A solution may be to ask patients how much time others would have to spend to perform all lost activities, rather than how much time others actually did spend taking over tasks. To date, clear guidance and sufficient knowledge on how unpaid labor changes can be best measured is lacking, introducing another potential for undesirable variation in measurement.

Indirect estimation

A practical concern related to productivity cost measurement is that collecting all required information related to productivity changes might not be feasible. Researchers are sometimes fully dependent on retrospective data or data collection within clinical trials that do not include questions concerning patients' productivity. In these situations, occasionally published estimates of previous studies are used (if available), but productivity costs seem largely to be ignored under such circumstances.[10,43-45,87] One partial remedy is to indirectly estimate productivity losses based on other available patient information, especially quality of life measures,[19] which have been shown to be significantly associated with productivity.[19] In a recent exploration of this possibility, the results of a study by Krol et al.[55] were promising, but require additional confirmation. Applying prediction models to estimate productivity losses in cases where no actual data is available may be preferable to ignoring the effect of illness on productivity. Whether it is possible to predict productivity related to unpaid work based on quality of life data has not yet been explored.

2.5 Valuation

The valuation of productivity changes, especially those concerning paid work, has been most prominently debated in the literature. The following section highlights the main issues of debate and indicates where we stand in theory and practice.

Valuation approaches of lost paid work

Valuation methods for productivity have been fiercely debated in the health economic literature. Early work on the valuation of reduced productivity was based on the theory of human capital [(88)] whereby, in the case of total unproductivity due to illness, the relevant value of the production loss was assumed to equal the present value of all lost future earnings of the individual. That is, income acts as a proxy for the production value of the individual and all production not produced by this person is counted as production loss. This obviously produces relatively large estimates of productivity costs in case of long term absenteeism, disability, and premature death. Consider, for instance, a person, who would have retired at age 65, but dies at age 35 from acute heart failure. The human capital approach then counts all lost production time (i.e., from age 35 to 65) to calculate productivity costs and commonly multiplies it with the expected average annual wage rate. For an annual wage of € 40,000, for example, total value of lost productivity would be 30 x € 40,000, or € 1,200,000 (without discounting).

The very high valuations resulting from the human capital approach induced important criticism. [(47,56,89,90)] The approach was importantly challenged by Koopmanschap and Van Ineveld, [(90)] who argued that an important, implicit underlying assumption of this approach is that all markets – including the labor market – clear, such that no involuntary unemployment ever occurs. In reality, involuntary unemployment is rather common; ill workers are often replaced by healthy, (unemployed) persons. In that case, Koopmanschap and Van Ineveld argued, productivity losses due to long term absence would be limited to the 'friction period', or the period it takes to replace the ill worker by a formerly unemployed person and, hence, to restore production to its initial level. Production losses and transaction costs (related to advertising, hiring, training, etc.) occur during the friction period only. Moreover, since a reduction in labor time is often assumed to cause a less than proportional decrease in production, Koopmanschap et al. [47] originally proposed to apply an elasticity factor of 0.8, which is often used in empirical studies applying the friction cost approach. Productivity costs using this method are markedly lower than using the human capital approach, especially in the case of long term absence. On the basis of an econometric model of a national economy, Koopmanschap et al. [47] showed that it is possible to focus on the short term costs incurred during the friction period to measure productivity costs. Long term indirect costs were estimated to be relatively small in relation to the friction costs, and relate to the effects of changes in the number of ill and unemployed persons in the population and the corresponding effects, for example, on social insurance premiums for unemployment and disability benefits.

The friction cost approach was subsequently criticized by proponents of the human capital approach (e.g., [58,91]). Important criticisms were that it lacked a sound theoretical underpinning and would render leisure time valueless. Although an unemployed person replacing the ill worker sacrifices leisure time to perform paid work, the friction cost approach gives it no value. Friction cost proponents indicated that leisure was not treated as 'costless', but rather that gains and losses of leisure at a societal level even out between the sick worker and the replacing worker. Moreover, differences in ability to enjoy leisure time are commonly valued in terms of quality of life. [57,59] Although the debate between proponents of the two approaches is likely to continue, Nyman [23] recently concluded on theoretical grounds that *"to be consistent with the societal perspective … frictional accounting of productivity costs should be adopted rather than the human capital accounting."*

This recommendation, made especially for the US context, was important, since during the debate between proponents of the two approaches in the 1990s, the US Panel on Cost-Effectiveness in Health[2] implicitly criticized both by advocating the inclusion of productivity costs in economic evaluations as *effects* rather than costs. The Panel suggested that the effect of disease on productivity is (and should be) valued through a preference-based measure of health like the quality adjusted life year (QALY) as health effects. They moreover indicated that respondents in health state valuations underlying QALY calculations would consider effects on productivity and income when valuing health states, unless explicitly instructed otherwise. This implies that any additional monetary valuation of the impacts on productivity would be double counting.

The US Panel method was criticized soon after it was proposed. It was suggested that its estimates would be unreliable, since people may be unable to translate a hypothetical health state, described in a health state valuation exercise, to an estimate of productivity. Moreover, the link between productivity and income on an individual level is weak because of private and social security systems. (22,31) In spite of the theoretical criticisms of the US Panel approach, the suggestion of double counting when also applying the human capital or friction cost approach could not be easily rejected. Further empirical research followed and is discussed in the next section.

Empirical evidence and the theoretical debate

The theoretical debates between the three valuation approaches could benefit, to some extent at least, from more empirical research. An important issue in the US Panel method was the assertion that individual respondents would include income consequences of reduced productivity as a result of poor health in their valuations of health states, for instance, when answering common time trade-off (TTO) questions. While such considerations may not (necessarily) result in a valid valuation of productivity losses, they could result in (partial) double counting. Such partial double counting casts doubts on including productivity costs in the numerator of a CE ratio, using either the friction cost or human capital approach. This issue initiated empirical studies focusing on whether respondents in health state valuations indeed consider income changes and, if so, how it affects valuations.

Recently, Tilling et al. (92) reviewed the seven studies thus far published on the topic, which varied substantially in design and approach. Their review indicated that, while a non-negligible minority of respondents may indeed include income changes in health state valuations when not explicitly instructed not to do so, inclusion did not appear to significantly affect health state valuations. Tilling et al. (92) concluded that this evidence suggests that the US Panel method is not a reliable valuation method for productivity costs. This underlines the theoretical criticism it received. Moreover, in terms of the risk of double counting, Tilling et al. (92) indicated that: "*the evidence seems to suggest that population value sets derived using generic instruments such as the EQ-5D are not noticeably influenced or 'polluted' by income effects and therefore can be used alongside monetary valuation methods of productivity costs, as well as in contexts where income effects are to be excluded from the analysis completely*". Although they indicated that more research is warranted to fully exclude the risk of double counting, it seems that a monetary valuation method (human capital or friction cost) is necessary to adequately value productivity costs and can be used in combination with conventional health state valuations. Nyman, after considering these issues, recently [23] proposed to change US guidelines and move away from the US Panel approach. He moreover suggested explicitly instructing respondents in health state valuation exercises to assume income would *not* be affected by illness. Since empirical evidence suggests that spontaneous inclusion of income effects does not seem to affect health state valuations, [92] it is uncertain whether such explicit instruction (which may unintentionally over-emphasize income in health state valuations) [68] results in 'better' health state valuations. Future research is thus required.

The debate between the proponents of the human capital and friction cost approaches has generated less empirical research, partly related to the debates' more fundamental differences that render empirical research difficult at best. For example, the existence of involuntary unemployment may be undisputable, but it need not be viewed as complete evidence that the friction cost approach is appropriate. More empirical research, such as on exactly how firms replace ill workers, where they are recruited from, and what it implies for the friction period (one period or a chain of periods) would be an important area for future research. It is, moreover, important to have estimates of transaction costs of replacement of employees, since they can be substantial. [93] In addition, for most countries the appropriate friction period to apply is unknown. Without establishing (and regularly updating) these data, an accurate calculation of productivity costs is hampered.

Other valuation issues

Common standards regarding the appropriate operationalization of valuation methods are lacking, resulting in practice variation even between studies claiming to use the same method. An important source of variation is the exact value attached to lost time from work. In published economic evaluations values have been based on, for example, the employees' added value to the firm, the employees' gross income, average national gross income, or, less frequently, GDP per capita, insurance payments, or doctors' opinions. Patients' gross wage, however, or the (age and gender-dependent) average national gross wage seem the most commonly used valuation sources. [44] Which of the two is more appropriate depends on the study sample. Posnett and Jan [49] claim that: *"Unless the study is designed in such a way that the study population can be taken to be fully representative of the population of all potential patients, the* *results of the evaluation will be more readily generalized if market values for potential patients (of wage rates for example) are used (…..).”*

An unresolved valuation issue is whether and how to adjust conventional productivity cost estimates for compensation mechanisms. Currently, such adjustment is uncommon in economic evaluations although estimates using the friction cost approach may already include a correction for compensation of lost work during normal working hours by factoring in the 0.8 elasticity. What this elasticity exactly represents is unclear, however, and to our knowledge its appropriateness has not been extensively investigated. If the elasticity indeed already represents compensation in normal hours, additionally correcting for compensation mechanisms could be a ‘double correction’ that artificially lowers productivity cost estimates. Adjusting estimates for multiplier effects is also uncommon. Important research questions linger in this area too, like whether the influence of one person’s productivity on a team is not yet reflected in his wage.

To date, few attempts have been made to directly correct productivity cost estimates for either team effects or compensation mechanisms. (15-18) One recent study corrected for these effects simultaneously, (80) but how compensation mechanisms and team effects interact in practice has not been explored. Examining this interaction is relevant, since complementary team work could increase the opportunity to compensate for a missing member in some cases, but lead to larger losses in others.

Valuation of unpaid labor

Although some attention has been paid to the valuation of unpaid work related to informal care (94-99), little attention has been paid to valuing patients’ unpaid work. Time spent on unpaid labor such as household work can be considered a non-marketed use of time, which is commonly valued by assigning a shadow price based on the opportunity cost method or the replacement cost method. (7,96) With the replacement cost method (also called the proxy good method), the value of unpaid labor is determined by valuing the time it would take others to perform the tasks normally performed by the patient, determined by the cost of hiring a paid worker to perform the tasks. The value of an hour of household work, for example, would be determined by the (average) hourly gross wage of a paid household help.

Alternatively, the opportunity cost method could be applied, where the value of unpaid work is commonly set equal to the value of competing use of time spent on paid labor. For the employed, the value of net wage could be used as a basis for computation; for the unemployed, potential wage.

In practice, most economic evaluations ignore unpaid labor and little attention is paid to further develop and implement its valuation methods. A lack of theoretical attention and practical inclusion, however, are likely to interact. An increase in scientific interest in unpaid labor may eventually lead to an increase of inclusion of unpaid labor in economic evaluations.

2.6 Productivity costs in practice

So far, debates regarding whether and how to include productivity costs in economic evaluation have not led to an increase of their inclusion in economic evaluations since the 1980s. As seen in table 2.2, productivity costs related to paid work are neglected in the vast majority of cases and their exclusion rate seems to be rather stable over time. Productivity costs related to unpaid labor are rarely included. Moreover, there is a clear lack of standardization regarding methodology when including productivity costs. Exclusion may be partly explained first by the fact that half the national health economic guidelines prescribe a health care perspective. Second, productivity costs may not always be relevant to include, such as those related to paid labor in studies including elderly patients. The explanations nonetheless leave a considerable number of cases unexplained. [(45)] Productivity costs seem somewhat randomly included or excluded. Krol et al. [(45)] suggest this may indicate the existence of some kind of perspective selection bias; i.e., decisions regarding productivity cost inclusion and exclusion could be based on the expected effects on incremental cost-effectiveness outcomes. To avoid such bias, decision-making bodies should be more aware of their in- or exclusion whenever a societal perspective is (or ought to be) taken.

2.7 Conclusions

Past

A variety of aspects regarding the inclusion of productivity costs in economic evaluations have received attention in the scientific literature. The past few decades have seen important progress by, for example, acknowledging that productivity costs can be identified in cases other than absenteeism, such as presenteeism and effects on coworkers' productivity, and that the costs of

Table 2.2 Productivity cost inclusion

Study	Source	Type of studies	N	PC included N (%)	%PC of TC	Methodology		
						Identification	Measurement	Valuation
Gerard [87]	Inter alia, CD-ROM, BIDS, Excerpta Medica	CUA published 1980-1991	51	5 (9.8%)	NA	NA	NA	NA
Stone et al. [10]	PubMed, HealthSTAR, CancerLit, Current Content & EconLit	CUA 1975-1997	228	19 (8.3%)	NA	NA	13 studies applied literature estimates, 5 expert opinion, 1 NS	17 HCA, 1 HCA + FCA, 1 FCA
Pritchard & Sculpher [43]	HEED	Economic evaluations up to and including 2000	1086	120 (11%) 40 CUAs and CEAs were further analyzed	Wide variation	Absenteeism paid work + 5 out of 40 studies included unpaid work	NA	26 HCA, 7 WPA, 7 FCA
Krol et al. [44]	PubMed & Cochrane	CEA + CUA + CMA 1997-2008 of therapies for depressive disorders, working age	81	25 (31%)	3-92% Mean HCA: 61% Mean FCA: 56%	Absenteeism paid work. One study included presenteeism, one study included presenteeism + unpaid labor	10 studies used questionnaires, 10 used literature estimates, 3 doctor estimates, 5 NS, 2 other	24 HCA, 6 FCA, 13 average wages, 5 GDP based values, 4 insurance payments, 5 NS, 3 other
Krol et al. [45]	PubMed	CEA + CUA + CMA of expensive intramural drugs published 1998-2009	249	22 (9%)	0.3-83% Mean HCA: 49%, Mean FCA 23%	Absenteeism paid work. None of the studies reported presenteeism and unpaid labor	None of the studies specified the measurement instrument	22 HCA 3 HCA + FCA 10 studies used average wage rates, 10 did not specify values, 2 used other values

PC = productivity costs, TC = total costs, FCA = friction cost approach, HCA = human capital approach, WPA = Washington panel approach, CEA = cost-effectiveness analysis, CUA = cost-utility analysis, CMA = cost-minimization analysis, NA = not applicable health-related productivity changes depend on how they are compensated and, subsequently, the costs of the compensation. Numerous productivity costs measurement instruments have been developed over the years. Most of them focus on health-related productivity changes in the context of paid work but neglect unpaid labor and coworkers' productivity effects. Productivity cost estimates can vary substantially with the instrument, however, and which one provides the most accurate estimates has not been established. The valuation of health-related productivity changes has been fiercely debated. Empirical research initiated by the debate has illustrated that productivity changes are not sufficiently captured on the effects side of the cost-effectiveness ratio and therefore are better placed on the costs side. Whether productivity costs should be valued according to the human capital or friction cost approach remains a matter of debate.

Present

Despite the progress and the substantial amount of scientific research, consensus has not been reached on either the inclusion of productivity costs in economic evaluations or the methods used to produce productivity cost estimates. Such a lack of consensus has likely contributed to ignoring productivity costs in economic evaluations and is reflected in variations in national health economic guidelines.[10,43-45,87] Given the current variety in applied productivity cost methodology, accurately reporting the methods used to estimate productivity costs in economic evaluations is critical to understand the nature of potential differences in outcomes. The reported level of detail is often poor. [45] Reported outcomes should also consider different types of production losses, such as absenteeism, presenteeism, and lost unpaid labor. Reporting productivity costs separately from direct costs to enable comparison of outcomes between studies, as is sometimes done, is recommended.

Future

Further scientific research is needed to lessen the controversy regarding the estimation of health-related productivity costs. In table 2.3 we provide an overview of the most topical issues inspired by questions and concerns of previous sections. The table partly reiterates earlier pleas. [51] The activities in table 2.3 may serve a number of general objectives: (i) to increase the inclusion of productivity costs in economic evaluations whenever (deemed) relevant; (ii) to increase scientific research regarding (the inclusion of) health-related changes in unpaid labor; (iii) to promote the comprehensive inclusion of the costs of health-related productivity changes in paid labor (i.e., absenteeism, presen-

Table 2.3 Future research topics

Unpaid labor	Further develop and test methods to identify, measure and value relevant unpaid productivity changes.
Paid labor: Inclusion	Reduce the lack of standardization and guidance regarding the inclusion of productivity costs.
Paid labor: Identification	Explore how health related productivity changes affect coworkers' productivity and team output. Develop methods to identify and include these effects in economic evaluations.
Paid labor: Measurement	Develop and validate measures of compensation. Develop and validate measures of multiplier effects. Investigate the relationship between wages and multiplier effects. Investigate the robustness and transferability of multiplier effect estimates. Further investigate whether estimated productivity changes based on quality of life data can be useful for productivity cost calculations in cases where direct productivity measurement is not feasible.
Paid labor: Valuation	Investigate the costs of compensation mechanisms. Investigate the friction cost period for different countries. Investigate replacement costs of employees.
Other	Explore the actual and potential redistribution (mechanisms) of productivity costs and savings within and outside the health care sector.

teeism, and the effects on coworkers productivity); (iv) to increase knowledge on how health-related productivity changes lead to societal costs or savings; (v) to investigate how employee replacement costs can be measured and included in economic evaluations when relevant; and (vi) to explore the theoretical and practical potential to transfer societal savings due to productivity increases induced by medical treatment into the health care sector.

As has been shown, productivity costs can be substantial and the decision to include or exclude them can strongly affect cost-effectiveness outcomes. [9,43-45] This implies that if productivity costs are excluded where they are relevant, decisions regarding reimbursement of health interventions may be based on inaccurate incremental cost-effectiveness ratios. Decision makers aiming to make well informed decisions on reimbursement from a societal perspective might wish to emphasize the inclusion of productivity costs in future cost-effectiveness studies.

The ultimate aim would be to develop common standards regarding the identification, measurement, and valuation of productivity costs. Although consensus on methodology is lacking, more standardization is essential to increase the comparability and credibility of economic evaluations taking a societal perspective. We hope to stimulate additional research in the important area of productivity costs.

Acknowledgements

The study on which this chapter is based is part of a larger project investigating the broader societal benefits of health care, which was financially supported by AstraZeneca, GlaxoSmithKline, Janssen, Merck and Pfizer. The views expressed in this chapter are those of the authors. The authors have no conflicts of interest to declare.

3.

Productivity costs in health state valuations: Does explicit instruction matter?

Based on: Krol M, Brouwer W, Sendi P.

Productivity costs in health-state valuations : does explicit instruction matter?

Pharmacoeconomics 2006;24(4):401-414.

Summary

Background: There has been considerable debate on whether productivity costs should be captured in the numerator or the denominator of the cost-effectiveness ratio. That debate cannot be resolved on the basis of theoretical arguments alone because the final choice also depends on what is incorporated in health state valuations by respondents and how this influences outcomes. At the moment, little is known about whether the effects of ill health on income are included in health state valuations, and how instructions on including or excluding the effects on income influence health state valuations. Aim: To conduct an empirical study of health state valuations to test: (i) whether or not respondents spontaneously include the effect of ill health on income and leisure time; (ii) the impact on the valuation of inclusion (or exclusion) of such effects; and (iii) the influence of explicit instructions on this matter. Methods: Three questionnaires were developed and administered to the general public. Health state valuations were conducted by visual analogue scale scoring of three health states of differing severity taken from the EQ-5D. Version 1 had no directions regarding inclusion/exclusion of effects of ill health on income. Those respondents who spontaneously included effects on income were subsequently asked to value the same three health states again, excluding these effects. Version 2 had explicit instructions to incorporate the effects on income. Version 3 stated that income was assumed to not change as a result of ill health. Respondents for versions 2 and 3 were also questioned about inclusion of effects on leisure time. Results: Giving explicit instructions on the incorporation or exclusion of effects of ill health on income did not lead to significant differences in subsequent health state valuations. In the absence of instruction, 36% of respondents included and 64% excluded effects on income, but the health state valuations of the two groups were not significantly different. Eighty-four percent of respondents included the effects of ill health on leisure activities, and again this had no significant impact on the resulting health state valuations. Conclusions: It appears that neither spontaneous differences in incorporation of effects on income, nor explicit instructions will yield significantly different health state valuations. This may suggest that QALY measures are insensitive to concerns regarding effects on income even when these are (explicitly) incorporated, and these effects may therefore be best placed on the cost side of the cost-effectiveness ratio.

3.1 Introduction

The inclusion of productivity costs in economic evaluations of health care has always been an area of controversy.[5] Some have argued that productivity costs should not be part of a health economic evaluation, since these costs are not relevant for health care decisions, do not fall under the health care budget and their inclusion may have adverse distributional consequences (i.e. favoring interventions aimed at productive individuals). (12,100-102) Still, the normally advocated societal perspective dictates the incorporation of all societal costs significantly influenced by the intervention under study, regardless of where these costs fall. (2,103) This implies that productivity costs should be captured in a health economic evaluation.

Yet, even if we would agree on taking the societal perspective (and a practical consensus on this matter seems lacking, judging from the different national guidelines; see www.ispor.org), then the debate becomes whether productivity costs should be captured in the cost or the effect side of an economic evaluation, especially when effects are expressed as QALYs (8,20-22,31,104,105). Although the distinction between costs and effects may be thought of as rather straightforward, some argue that the consequences of reduced productivity should be, and normally are, captured in health state valuations, (2) thus in terms of effects, while others argue that these consequences need to be seen and expressed as costs in an economic evaluation. (8,22,31,68,91) If one opts to incorporate productivity costs in terms of costs, a further debate will be whether the human capital or the friction cost method is most appropriate to determine productivity costs. (47,57-59,102)

In this chapter, we are especially concerned with the debate on whether productivity costs should be placed at the cost or effect side of an economic evaluation. This debate was stimulated by the publication of the guidelines of the US Panel on Cost-Effectiveness in Health and Medicine, which recommended that for "… *the reference case analysis, health-related quality of life should be captured by an instrument that, at minimum, implicitly incorporates the effects of morbidity on productivity and leisure.*" (2) Although theoretical

5 In fact, even the definition of the term productivity costs differs in the literature. Brouwer et al. (22) refer to productivity costs as "*costs associated with production loss and replacement costs due to illness, disability and death of productive persons both paid and unpaid*," while the US Panel on Cost-Effectiveness in Health and Medicine refers to productivity costs as "*the cost associated with lost or impaired ability to work or engage in leisure activities due to morbidity and lost economic productivity due to death*." (2)

objections against this recommendation were put forward, the recommendations clearly brought to light the fact that currently there is little knowledge on which aspects of ill health respondents actually incorporate in health state valuations. It is especially uncertain whether respondents include productivity costs in terms of income effects in health state valuations. Although income losses are not synonymous with the broader concept of productivity costs, they are sometimes seen as a proxy for productivity costs related to paid work, (2) although this has been disputed.[6] (22,68) However, it is clear that if respondents incorporate income changes in health state valuations, while recognizing the general difference between income losses and productivity costs, this implies a clear risk of double counting if productivity costs were also to be counted as costs. In the context of this chapter, therefore, we will use the terms interchangeably.

Given the recent debate in the literature on whether productivity costs should be seen as costs or effects, more knowledge on this issue is essential. If it turns out that some respondents do and some do not include income effects, productivity costs might be underestimated or double counted, depending on whether productivity costs are placed at the cost or effect side of the cost-effectiveness ratio and the effect incorporation of these effects has on subsequent health state valuations. Moreover, it has been argued that to be sure that all respondents include the same aspects in health state valuations, explicit instruction to either include or exclude productivity costs might be necessary. (68) Yet, to our knowledge, the effects of such instructions have never been investigated.

In this chapter we present the results from an empirical study with the aim: (i) to find out whether people include potential income losses (and leisure-time loss) in health state valuations that are silent on effects of ill health on income (and leisure); and (ii) to find out what the effects of explicit instructions on including or excluding the effects of ill health on income are on health state valuations.

First, we provide some background on the debate on whether productivity costs should be in the numerator or the denominator of the cost-effectiveness ratio and present the scarce previous empirical studies on whether respondents include health-related income loss in health state valuations. We then

6 Note that the perspective of the study is relevant here as well. Income losses will be incorporated to the extent that the individual experiences them, while from a societal perspective this may not be a valid approach to measuring and valuing productivity costs. (22)

present the study we performed to test the incorporation of effects of ill health on income and leisure time in health state valuations and the influence of explicit instructions on this matter.

3.2 Productivity costs and health state valuations

In 1996, the US Panel on Cost Effectiveness in Health and Medicine recommended that productivity costs be included in a non-monetary value, such as QALYs, on the effect side of the cost-effectiveness ratio. (2) The Panel suggested doing this by explicitly mentioning that respondents should include the effect of ill health on income in health state valuations, but assumed that silence on this subject is enough to make people incorporate income effects. (2,21)

An important objection against the recommendations of the US Panel is that measuring the effects of ill health on income through QALY elicitation will not lead to accurate measurement of productivity costs. (8,22,31,91) For example, individuals who are insured against income losses experience little or no income effects of ill health, while their productivity may be significantly affected. Moreover, differences may exist between the societal valuation of production loss and an individual's valuation of income loss. Such differences may arise from various factors, for instance, differences in the period considered to be relevant for calculating what losses occur, whether gross wages are used to reflect the added value of a worker to society, or whether net income is used to reflect the outcome of interest for the individual, etc.

Even if it were possible to measure all income effects accurately at the effect side, these are not the only relevant costs. Indeed, the additional costs for the employer (e.g. hiring and training new employees) would still have to be calculated at the cost side of the cost-effectiveness ratio, since an individual will normally not consider such costs. (2) Such a separation of productivity costs (partly incorporated at the effect side and partly at the cost side) can cause an increase in the costs of conducting economic evaluations. (31) However, it needs to be noted that the loss of the ability to perform (un)paid work is usually seen as part of quality of life, for instance in terms of role functioning, but this needs to be distinguished from the societal costs resulting from this inability to function in a paid job. (22) In this chapter, we only address the latter aspect.

Capturing productivity costs at the cost side of the cost-effectiveness ratio (e.g. using the human capital or friction cost approach) (7,102,105) has also been criticized. Besides discussion on which of the two methods is superior, an important question in this context is whether incorporation at the cost side of a cost-effectiveness ratio could lead to double counting. [20] This would happen if respondents already take productivity costs into account when valuing health states, which is currently largely unknown. From this it follows that costs and effects in a cost-effectiveness analysis are complements and one needs to ensure that all relevant items are captured on both sides, not leaving out anything important and not double counting anything by including the item on both sides. [68] This indicates that knowledge of what respondents incorporate in health state valuations and how this alters results is instrumental for correct analysis, as is the knowledge of how to better instruct respondents to consistently include or exclude relevant aspects.

To date, there is little empirical evidence on what respondents consider in health state valuations. Only two studies have been conducted to find out if respondents included the effects of ill health on income in health state valuations. Meltzer et al. [60] tested how giving information about the financial consequences of ill health affected respondents' valuations using time trade-off (TTO) questions. However, they did not give explicit instructions to respondents on what to do with this information. Respondents were asked afterwards if they had thought of financial consequences of illness when answering the TTO questions. Even in the group in which respondents were informed that no form of disability payment existed, <25% of the respondents considered the financial consequences of illness. Without any information on income losses, <15% thought of the negative effects of ill health on income. Within the group without information on income loss, respondents who indicated that they had (spontaneously) considered effects on income showed lower TTO scores than respondents who did not consider effects of ill health on income.

Sendi and Brouwer [64] recently performed a small sample test among 20 health professionals to find out whether the effects of ill health on income and leisure time were included in health state valuations when no explicit instructions were provided. Respondents were asked to value a health state on a visual analogue scale (VAS). Forty percent of the respondents included the effects of ill health on income, whereas 60% did not. The former respondents had a significantly lower VAS score than the respondents who did not include effects of ill health on income.

Both studies [60,64] show that, without instruction, many respondents do not consider effects on income in health state valuations, although a reasonable proportion of respondents do consider such effects. To ensure that respondents are consistent in including or excluding effects of ill health on income in health state valuations and thus avoid double counts or omissions of important costs (regardless of the position one wants to take in terms of where they should be placed in the cost-effectiveness ratio), giving people explicit instructions in health state valuation exercises has been proposed. [68] Given the theoretical debate, two types of explicit instruction appear most relevant. First, one may opt to instruct respondents to incorporate all possible effects of ill health on income in their valuations (following the recommendations of the US Panel). Second, one could instruct respondents to exclude these effects and measure all productivity costs in a monetary value (using the human capital or friction cost method). Whether such instructions indeed result in changes in health state valuations has not been investigated to our knowledge.

Brouwer and Rutten [68] have questioned whether instructing respondents not to consider specific costs or effects is feasible, as it resembles instructing people not to think of a 'pink elephant'. It is therefore essential to have more information on what respondents do without instruction and what the influence on health state valuations is of explicit instruction on the inclusion or exclusion of income effects of ill health.

In the study by Sendi and Brouwer [64] respondents were also asked whether they had considered effects of ill health on leisure time. It transpired that (only) 75% of the respondents indicated that they had considered these effects. Although in this chapter the main focus is on productivity costs, and despite the fact that there is little debate on whether leisure time should be incorporated in the QALY measure, [2,20,64,68,91,106] it needs to be noted that it is equally uncertain, and therefore worthwhile investigating further, whether respondents in fact include leisure time in health state valuations (given the small sample size in previous research). [64] If not all respondents include these effects, even though we expect them to, it is also important to see what effect this has on subsequent health state valuations.

3.3 Methods

To test the effects of including loss of income and leisure time in health state valuations, as well as to test the influence of instructions regarding how to include such effects, three different questionnaires were developed. In all, respondents to the questionnaires were first asked some background questions about sex, age, income, education and their current health state using the EuroQoL descriptive system (EQ-5D) and the EuroQoL VAS. [107]

Next, the respondents were asked to value on a VAS, three ill-health states that differed in severity (see table 3.1). These three health states were chosen from 243 possible health state scenarios within the EQ-5D. (108) Our main selection criterion was to have three distinct health states in terms of severity. A VAS was used since it requires a minimum of explanation to respondents and lends itself to self-administration. (109)

In version 1 of the questionnaires, respondents did not receive directions on whether to incorporate or to ignore the effects of ill health on income. After having valued the three health states, the respondents were asked whether they included effects on income. If they indicated that they had included these effects, they were asked to value the same three health states again, but now with the explicit instruction to ignore the effects of ill health on income. In version 2, respondents were explicitly instructed upfront to incorporate the possible effects of ill health on income. In version 3, respondents were instructed to assume that income would not change as a result of ill health, and thus to ignore the effects of ill health on income. This phrasing was chosen in an attempt to avoid the 'pink elephant' trap described in the previous section. In all questionnaires, respondents were asked whether they thought being in the different health states would decrease their income and, if so, to what extent. This was always done after the valuation process, so as not to 'contaminate' the experiment. In versions 2 and 3, respondents were furthermore asked if they thought the described health states would affect their leisure time and whether they had included the possible effects on leisure time in their valuations. The latter questions were not included in version 1 (with the double health state valuations) to limit the size of the questionnaire.

Based on earlier findings and the debate in the literature, we formulated several hypotheses that this study sought to confirm:

Table 3.1 Health states described for valuation

Health state	EQ-5D code	Standard MVH_A1 score (108)	Description
1	21211	0.88	Some problems in walking, no problems with self-care, some problems performing usual activities, no pain or discomfort, not anxious or depressed
2	22221	0.587	Some problems in walking, some problems with self-care, some problems performing usual activities, moderate pain or discomfort, not anxious or depressed
3	33312	–0.043	Confined to bed, extreme problems with self-care, not able to perform usual activities, no pain or discomfort, moderately anxious or depressed

MVH = Measurement and Valuation of Health group.

1. In health state valuations, if respondents do not receive directions on including or excluding the effects of ill health on income and leisure, some respondents will include while others will exclude these effects in their valuations. [60,64]
2. Respondents who spontaneously include the effects of ill health on income in health state valuations give lower valuations than respondents who do not include these effects, *ceteris paribus* (i.e. all other things being equal). [60,64]
3. Similarly, respondents who spontaneously include the effects of ill health on leisure in health state valuations give lower valuations than respondents who do not include these effects, *ceteris paribus*.
4. Respondents who include the effects of ill health on income in health state valuations without explicit instructions will value the same health states higher the second time when asked explicitly to exclude these effects. [64]
5. Respondents who are explicitly asked to include the effects of ill health on income in health state valuations value health states lower than respondents who are explicitly asked to exclude the effects.

The study sample was drawn from the general public in 2004 in various public places such as people working at a hospital, people travelling by public transport, etc. Respondents were approached and asked whether they were willing to participate in this study. If they agreed to participate, the respondents in our convenience sample were randomly handed one of the three different versions of the questionnaires for self-completion.

Statistical analyses

Categorical variables were compared using the Chi-square test. Continuous variables were compared using the t-test, one-way ANOVA or Wilcoxon signed rank test where appropriate. We used a multivariate general estimation equation (GEE) model with a logit link and an unstructured correlation matrix to estimate the association between health state, the version of questionnaire administered, employment status and education (adjusted for age and sex) and the perceived effects of ill health on income by the respondents. A GEE model was used to account for the within-individual correlation of the respondents when judging the three health states described. Robust standard errors and z-values were used to calculate confidence intervals for odds ratios and P-values. A linear mixed-effects model was used to estimate the association between the valuation of the health states and whether respondents did consider effects of ill health on income in version 1 of the questionnaire, employment status and education, adjusted for age and sex. Respondents valuing the three health states were treated as random effects to account for within-respondent variability and correlation using an unstructured correlation matrix. A linear mixed-effects model was also used to evaluate whether there were differences in health state valuation between the three versions, adjusted for age, sex, employment status and education. Income was not included in the multivariate analysis. It was considered that income is a function of both education and employment status, and that both these variables would lead to a better fit than income alone. Moreover, including all three variables did not lead to a gain in information, rather to a loss of power. The data were analyzed with use of SPSS Statistical Software Package 11.5 and S-Plus 6.2 Professional.

3.4 Results

Sample population

Of the 227 completed questionnaires, 185 were useful for further analysis (version 1, n = 59; version 2, n = 64; version 3, n = 62). The exclusion of questionnaires was based on inconsistencies suggesting that respondents did not understand the purpose of the questionnaire. Most of these respondents gave all three health states the same value as their own. Some drew several lines from the health state description to the scale. The respondents of the excluded questionnaires were significantly older (P = 0.01), less educated (P = 0.00) and, at a 10% confidence level, of lower income (P = 0.06) than the respondents of the questionnaires used for further analysis.

Respondents were more often well-educated and younger than a representative sample would be expected to be. [110] As seen in table 3.2, the average age was 35.2 years (SD 13.1) and did not differ between groups (P = 0.84). The average self-reported health of the respondents was 0.81, which did not differ between the respondents of the different versions of the questionnaire (P = 0.92, SD 0.13). Respondents were more often female (54%). There were no significant differences between the groups in sex (P = 0.53), employment (P = 0.84) or income (P = 0.57).

Table 3.2 Characteristics of respondents by questionnaire group

Characteristic	Questionnaire administered			All respondents
	Version 1 (n = 59)	Version 2 (n = 64)	Version 3 (n = 62)	(n = 185)
% Females	59	53	49	54
Mean age [y] (SD)	36.0 (13.7)	34.7 (13.1)	35.0 (12.6)	35.2 (13.1)
% Employment				
No paid work	24	23	18	22
Fulltime employment	44	41	48	44
Part-time employment	32	36	35	34
% Education [a]				
Lower education	12	11	16	13
Medium education	39	42	28	36
Higher education	49	47	56	51
% Income [b]				
Lower income	16	21	17	18
Medium income	41	29	29	33
Higher income	43	51	54	49
Mean SAH with VAS (SD)	0.81 (0.12)	0.81 (0.15)	0.80 (0.12)	0.81 (0.13)

a These groups were defined according to the Dutch educational system. The category ‘lower education’ comprises elementary school, lower general secondary education (MAVO) and pre-vocational (VBO). ‘Medium education’ comprises higher general secondary education (HAVO, VWO) and senior vocational education (MBO). ‘Higher education’ comprises vocational colleges (HBO) and academic education. b Based on net income per month:
low = <€650, medium = €650–1900, high = >€1900, SAH = self-assessed health,
VAS = Visual Analogue Scale, y = years, SD = standard deviation

Effects of ill-health on income

As shown in table 3.3, 34%, 75% and 88% of the respondents thought that being in the health states 1, 2 or 3, respectively, would decrease their income. For all health states, a smaller percentage of respondents to questionnaire version 1 than to version 2 and 3 thought income would decrease (although only the differences between questionnaire versions for health state 3 were significant). In other words, the group of people that was not triggered upfront to think about income through a prior instruction was less likely to think income would decrease because of the ill-health states. Respondents stating that the very poor health state 3 would not cause income loss were, unsurprisingly, more often unemployed (since normally these respondents are not able to lose income because of ill health) than the people that did expect a decrease in income ($P < 0.001$). These differences in employment were not found between respondents that did and did not think health states 1 and 2 would decrease income. This might be explained by the fact that working in health states 1 and 2 may still be possible, while in health state 3 this is unlikely since one is confined to bed.

Table 3.3 Effects of ill-health on income and leisure time

	Health state 1 decreases income	Health state 2 decreases income	Health state 3 decreases income	Included effect on income	Included effect on leisure-time
Version 1 (n = 59) [%]	25	69	78	36	-
Version 2 (n = 64) [%]	43	79	90	-	81
Version 3 (n = 62) [%]	33	78	92	-	88
Mean (%)	34	75	88	-	84
P-value [a]	0.118	0.357	0.012	-	0.258
Mean estimated income loss [€] (SD)	298 (324)	473 (426)	737 (573)	-	-

[a] Pearson chi-square test
SD = standard deviation

Table 3.4 Multivariate analysis of the effects of health state, version of questionnaire administered, employment status and education on the likelihood of believing ill health would affect income, adjusted for age and sex

Variable	Odds Ratio	95% CI	P-Value
Health state 2 vs. 1	12.72	7.68-21.09	<0.001
Health state 3 vs. 1	19.06	11.20-32.41	<0.001
Fulltime vs. no job	3.96	1.70-9.18	0.001
Part-time vs. no job	3.58	1.47-8.72	0.005
Version 2 vs. 1	2.71	1.31-5.61	0.007
Version 3 vs. 1	2.06	1.03-4.10	0.039
Education (high vs. low)	0.50	0.25-1.00	0.050

The results in table 3.3 were largely confirmed in the multivariate analysis (table 3.4). Health states 2 and 3 were more likely thought to induce income losses than health state 1; respondents who received version 2 or 3 (therefore, with explicit instructions), and respondents with a full- or part-time job, were more likely to think that a certain health state would induce income losses. The variable education was on the border of significance at the 5% level (with P = 0.05), with the odds ratio indicating that people with a higher education were less likely to think a health state would induce income losses (perhaps because of differences in work type).

The effects of instruction

No instructions

Confirming hypothesis (1) that without instructions some respondents will and some will not include the effects of ill health on income in health state valuations, 36% of the respondents of version 1 of the questionnaire included the effects of ill health on income in their valuations. Although higher than the 15–25% in the study by Meltzer et al., [60] this result is comparable with the 40%

Table 3.5 Health state visual analogue scale valuations of respondents to version 1 of the questionnaire that spontaneously included or excluded the effects of ill health on income

Group	Health state 1 [Mean (SD)]	Health state 2 [Mean (SD)]	Health state 3 [Mean (SD)]
Included effect on income (n = 21)	0.68 (0.10)	0.53 (0.11)	0.32 (0.12)
Did not include effect on income (n = 38)	0.69 (0.13)	0.49 (0.17)	0.28 (0.12)
P-value [a]	0.722	0.295	0.261

[a] T-test equality of means
SD = standard deviation

in the small sample of Sendi and Brouwer [(64)] and differences in methods used may partly contribute to these differences.

For respondents to version 1 of the questionnaire, the group that did and the group that did not consider effects on income in their health state valuations did not differ in age, sex, education, employment, net income and self-reported present health state. Moreover, there were no significant differences between the health state valuations of respondents who considered income effects and respondents who did not (table 3.5). These findings contradict the findings of Meltzer et al. [(60)] and Sendi and Brouwer [(64)] and do not support hypothesis (2) that respondents who spontaneously consider effects on income will value health states lower than respondents who do not consider these effects.

These results were maintained in a multivariate analysis. Again, whether effect on income was considered did not significantly impact health state valuation (table 3.6). The only variable significantly associated with health state valuation was the health state presented. The more severe the health state, the lower the valuation of that health state by the respondent.

From no instructions to instructions

The respondents to version 1 of the questionnaire who included income in their valuations were asked to value the same health states again assuming these health states would not influence their income. It was expected that the respondents would value the same health states significantly higher the second time, similar to the findings of Sendi and Brouwer. [(64)] The results in table 3.7 show that changing the instructions did not significantly affect the valuation of health state 1, whereas it did change the valuations of health states 2 and 3. This is probably best explained by the fact that most respondents stated that health state 1 would not affect their income, as seen in table 3.3, whereas most did expect that health states 2 and 3 would decrease income (which logically translates in the results found). These results largely confirm hypothesis (4) that respondents change their valuations if they are asked to not consider a

Table 3.6 Multivariate analysis of the effects of consideration of impact on income, health state, employment status and education on health- state valuation in version 1 of the questionnaire, adjusted for age and sex

Variable	Coefficient	SE	P-Value
Intercept	0.679	0.058	<0.001*
Included effect on income	0.015	0.029	0.604
Health state 2 vs. 1	-0.183	0.016	<0.001*
Health state 3 vs. 1	-0.394	0.020	<0.001*
Fulltime vs. no job	-0.014	0.036	0.706
Part-time vs. no job	-0.044	0.041	0.288
Education (high vs. low)	0.000	0.032	0.991

* Significant at $P < 0.05$
SE = Standard error

Table 3.7 First and second health state visual analogue scale valuations of respondents (n = 21) who spontaneously included effects on income the first time and were asked to exclude these effects the second time

	Health State 1 [Mean (SD)]	Health State 2 [Mean (SD)]	Health State 3 [Mean (SD)]
First valuation	0.68 (0.10)	0.53 (0.11)	0.32 (0.13)
Second valuation	0.70 (0.11)	0.56 (0.10)	0.37 (0.13)
P-value [a]	0.167	0.027*	0.041*

[a] Wilcoxon signed rank test
* Significant at $P < 0.05$

negative aspect of ill health they did consider before, when this is applied to the most relevant scenarios (i.e. health states 2 and 3). Despite these significant differences, many respondents did not change their original valuation the second time, as was the case in the study by Sendi and Brouwer. Fourteen of 21 respondents valued health state 1 the same, 13 valued health state 2 the same and 11 valued health state 3 the same.

Explicit instructions

Table 3.8 shows the average VAS valuations of the three health states of the respondents to the three different questionnaires. In contrast to our expectations, the respondents of version 2 (who were explicitly asked to include the effect of ill health on income) did not value the health states lower than the respondents of version 3 (who were explicitly asked to exclude these effects). As table 3.8 shows, there were no significant differences between respondents in the valuations of each of the health states for the different versions of the questionnaire. The valuations of the respondents without instruction (version 1) did not differ from those in the other two groups. Thus, these results seem to imply that instruction does not matter in the sense that it does not change the overall results. Again, these results were maintained in a multivariate mixed effects model where there were no significant differences in the valuations of a given health state between respondents of the three versions of the questionnaire (P = 0.665), when adjusted for age, sex, employment status and education.

As an aside, it is worth noting that the health state valuations in table 3.8 are different from the EQ-5D MVH (Measurement and Valuation of Health group)-A1 scores shown in table 3.1. (108) This need not be surprising or worrying, since the MVH_A1 scores are elicited through TTO exercises, while we used VAS valuations, and the latter are known to differ from TTO or standard gamble scores. (109) On average the respondents in our sample valued health states 1 and 2 significantly lower than the corresponding MVH_A1 scores: 0.67 versus 0.88 and 0.49 versus 0.59, respectively. Health state 3 was valued significantly higher: 0.29 versus –0.04. The incongruent valuations of respondents with the corresponding MVH_A1 scores might be partly caused by context- and end-aversion bias related to the use of the VAS. (109) Context bias reflects the fact that the VAS score for a health state depends on the number of better and worse states presented at the same time. Respondents have the tendency to spread their valuations over the scale. End-aversion bias, the reluctance of people to use the ends of a scale, is a well-known bias that affects scales scored on a continuum. (109) The fact that the MVH_A1 scores are based on a wider scale (scores can be below zero) is also of significance.

Regression techniques were used to investigate which variables had a significant influence on the valuations of the different health states. The results are shown in table 3.9, and as can be seen, age is a significant predictor for the valuation of all three health states. Self-reported present health state had predictive value for health state 1 and 2, but lost its significance for the valuation of (the very poor) health state 3. Education (at P < 0.10) and whether respondents think a health state will decrease income (at P < 0.05) is only of

Table 3.8 Health state visual analogue scale valuations according to version of questionnaire administered

	Health State 1 [Mean (SD)]	Health State 2 [Mean (SD)]	Health State 3 [Mean (SD)]
Version 1 (n = 59)	0.69 (0.12)	0.50 (0.15)	0.29 (0.12)
Version 2 (n = 64)	0.68 (0.12)	0.49 (0.15)	0.27 (0.16)
Version 3 (n = 62)	0.65 (0.10)	0.48 (0.14)	0.30 (0.15)
Mean	0.67 (0.11)	0.49 (0.15)	0.29 (0.14)
P-value [a]	0.316	0.671	0.463

[a] One-way ANOVA

Table 3.9 Results of regression analyses of visual analogue scale health state valuations (n = 176)

	Valuation health state 1			Valuation health state 2			Valuation health state 3		
	Beta	SE	P-value	Beta	SE	P-value	Beta	SE	P-value
Constant		.068	< .001*		.094	.248		.103	.766
Education	-.122	.006	.092	.019	.008	.805	.114	.008	.161
Employment	.002	.012	.977	-.065	.016	.447	.014	.018	.880
Gender	.074	.015	.267	.092	.021	.200	.093	.022	.222
Age	.303	.001	.001*	.319	.001	.001*	.311	.001	.002*
Net income	-.017	.008	.865	.049	.011	.646	-.098	.011	.389
Self-reported health	.415	.056	< .001*	.269	.078	< .001*	.088	.081	.248
Version 1 (vs. 2)	-.004	.018	.957	-.011	.025	.897	.016	.027	.853
Version 3 (vs. 2)	-.057	.018	.455	-.049	.025	.551	.103	.026	.229
Does HS 1 decrease income	-.150	.016	.028*	-	-	-	-	-	-
Does HS 2 decrease income	-	-	-	-.081	.025	.265	-	-	-
Does HS 3 decrease income	-	-	-	-	-	-	-.072	.036	.377
Adjusted R^2			.291			.171			.058

* Significant at $P < 0.05$
HS = health state.

influence for health state 1. The version of the questionnaire has no significant influence on any of the three health state valuations. This implies that the differences in instruction do not appear to lead to differences in health state valuations. These findings contradict the findings of Meltzer et al. [(60)] and Sendi and Brouwer [(64)] who did find differences between groups of respondents based on different considerations about negative effects of ill health on income, assessed without explicit instruction. The findings in our study do not support hypothesis (5) that respondents who are explicitly asked to include effects on income will value health states lower than respondents who are explicitly asked to exclude these effects.

Leisure time

The results in table 3.3 show that 84% of the respondents to questionnaire version 2 and 3 included the effects of ill health on leisure time, confirming the first hypothesis that some respondents do and some do not include leisure time in health state valuations. The characteristics age, sex, level of education, employment, self-reported present health state and net income did not differ between respondents who did and those who did not think of the effects of ill health on leisure time in the health state valuations. Thus, even though it is usually assumed that respondents incorporate the effects of ill health on leisure time in health state valuations, (2,7,68) not all do so. This may lead to omission of leisure time effects in economic evaluations. The extent to which this poses a serious problem crucially depends on the influence the incorporation of leisure in health state valuations has on outcomes. However, interestingly, the health state valuations of the two groups (that did or did not incorporate effects on leisure time) did not differ significantly. These results therefore do not support the hypothesis (3) that respondents who include the effects of leisure time will value the health states lower than those who do not include these effects, and raises the question of whether leisure is adequately captured in health state valuations at the moment. The latter holds, because one might expect variations in health state valuations between people who do and do not incorporate leisure, depending on the relative importance of leisure. In addition, if leisure time is to be incorporated at the cost side of the cost-effectiveness ratio and has a noticeable effect on total costs, this raises the question of whether leisure should not be captured in monetary terms, to avoid ignoring these effects.

Of note, the percentage of respondents considering leisure in our sample was slightly higher than the percentage considering leisure time in the small sample of Sendi and Brouwer (64) (84% vs. 75%). This could be due to the use of the EQ-5D in our questionnaire to describe the ill-health states, as opposed to the description of a specific health state in the questionnaire used by Sendi and Brouwer, (64) since the EQ-5D focuses explicitly on the ability to perform usual activities.

3.5 Discussion

These results show that, without instruction on whether to consider effects of ill health on income in health state valuations, some respondents will include these effects and some will not. Moreover, our results confirm that respondents are inconsistent in including or excluding the effects of ill health on income. There are large differences in the perceptions of people as to whether and to what extent their income will be influenced by ill health. In light of the debate on productivity costs, this means that measuring productivity costs without instructing respondents can lead to double counting or underweighting of productivity costs. As noted in the introduction, the extent of double counting is also related to the extent to which one considers income losses to be a proxy for (societal) productivity costs. Therefore, as suggested before in the literature, giving explicit instructions to respondents on what to incorporate and what to exclude in health state valuations seems appropriate. However, our results indicate that explicit instructions do not lead to significant differences in health state valuations (with the exception of the respondents to version 1, who were asked to change their considerations and to exclude a negative aspect of ill health they did include before, which in itself likely prompted respondents to change their valuations).

What do these results imply? First of all, instructions on the inclusion or exclusion of effects of ill health on income do not appear to result in different valuations of health states. Our results suggest that respondents seem to give little weight to income loss of ill health in health state valuations. Although we did not directly assess the relative importance that respondents placed on effects on income relative to other aspects in health state valuations (and this requires further investigation), our results indicate that valuations are determined more by other aspects of ill health. Therefore, incorporating productivity costs (partly) in the effect side appears to have no noticeable effect on health valuations. Since the incorporation of productivity costs at the cost side does have a (clear) influence on the resulting cost-effectiveness ratio, one may feel that productivity costs are best placed at the cost side and excluded from the effect side of the cost-effectiveness ratio. If our results are also confirmed in other samples and using other valuation methods such as TTO, this exclusion need not necessarily be explicit, since instructions do not seem to alter the final valuations. This would also imply that current valuations (which are normally based on health state valuations that are silent in terms of incorporation of effects on income, for example the much used MVH_A1 scores [(108)]) can be used in combination with monetary valuations of productivity costs without risking double counting, even though some respondents may have included effects on income.

Our results confirm the results of Sendi and Brouwer [(64)] in that no differences in health state valuations were found between respondents who did and those who did not include the effects of ill health on leisure. There is consensus among health economists that respondents should include effects on leisure time in their valuations and it has been noted that without incorporating effects on leisure time a QALY seems to become a somewhat hollow concept. [(91)] In this light, it is remarkable that differences in including or excluding leisure time seem to have no significant influence on health state valuations. More research is necessary to understand why the valuations of the 16% who did not include leisure time did not differ from the 84% who did. In fact, our results, if confirmed in other research, may imply that losses in leisure time may need to be (partly) covered at the cost side of an economic evaluation, unlike current standards and practice. In that respect, Brouwer and Rutten [68] recently proposed that it may be possible to reserve the QALY measure for the valuation of the mere fact that one is impaired in terms of leisure activities, while a monetary value could be attached to each hour in which this is the case. They suggest that "*the quality of leisure is then captured in terms of the QALY, valuing health states using a general indication of impairment, while the quantity of leisure is captured in terms of money.*"

Study limitations

Some limitations of our study need to be noted. First, 18.5% of the respondents did not seem to understand the questionnaire. This suggests that the questionnaire is difficult to understand for a part of the general public (especially for older and less educated people). If respondents had received a comprehensive oral instruction, the number of questionnaires that could have been used for further analyses might have been higher.

Second, there may be differences between different quality of life measurements in terms of their sensitivity to changes regarding inclusion of effects on income. We used a VAS instrument and it is possible that the income insensitivity of this instrument was partly due to scale compatibility. It is well known that respondents tend to give more weight to the subject of focus (in our case health, since the VAS ranges from the best to the worst attainable health state) and therefore might under-weigh other aspects related to ill health. [111] Other valuation techniques may be more sensitive to changes in income considerations, but this needs confirmation. Although Meltzer et al., [60] using TTO questions, found differences in TTO scores of respondents who had and those who had not thought about financial consequences of illness, their study differs in many ways from ours. Moreover, it needs to be noted in this context that in contradiction to our findings, Sendi and Brouwer [64] did find some significant differences in health state valuations between respondents who had and had not included the effects of ill health on income, also using the VAS (although the sample size in that study makes conclusions vulnerable to outliers). Differences between that and this study may explain the different findings. For instance their respondents were health professionals who might expect a greater impact of ill health on income (based on their experience) than the general public. Furthermore Sendi and Brouwer [64] asked respondents to value a specific disease (multiple sclerosis [MS]), while here we used more general health state descriptions without specifying duration or type of illness.

The clear chronic nature of MS could have caused respondents to predict more income loss or attach a greater weight to it. More research is needed to further test the insensitivity of the VAS instrument and to investigate whether other QALY elicitation methods are equally insensitive to differences in income considerations. The effect of different instructions and information provided to the respondents may also be worthy of further investigation.

Another point worth investigating is whether respondents who are explicitly instructed to either include or exclude income effects really follow these instructions. If they do not, this may explain the fact that we did not find an influence of instruction. Moreover, while we have chosen the general public as our source of health state valuations (following a societal perspective), it could be considered worthwhile investigating whether the influence of income loss would be more pronounced in patients affected by illness.

Finally, it is possible that a similar study as the one presented in this chapter will lead to different outcomes when conducted in another country. Maybe respondents would be more aware of possible income losses due to ill health in countries with a less comprehensive social security system than in The Netherlands. This may lead to a higher percentage of respondents including the effects of ill health on income, a higher weight being given to income considerations in health state valuations and therefore to significant differences between the valuations of respondents who include and those who exclude the effects of ill health on income.

3.6 Conclusions

It appears that neither spontaneous differences in incorporation of effects on income, nor explicit instructions will yield significantly different health state valuations. This may suggest that QALY measures are insensitive to concerns regarding effects on income even when these are (explicitly) incorporated, and these effects may therefore be best placed on the cost side of the cost-effectiveness ratio. Obviously more research (with larger samples and other health state valuation techniques) should be conducted to see if the results presented in this chapter hold, but it is hoped that this chapter will give new empirical impetus to the debate on measuring productivity costs and contribute to increasing the knowledge of what people take into consideration in health state valuations and on how to move forward.

Acknowledgements

We are grateful to Marc Koopmanschap and Ken Redekop for their useful comments on an earlier version of the paper on which this chapter is based.

4.

Breaking the silence:
Exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises

Based on: Krol M, Sendi P, Brouwer W.

Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises.

Value in Health. 2009 Jan;12(1):172-180.

Summary

Objectives: An underexplored question in the debate on incorporating productivity costs as costs or effects in a cost-effectiveness analysis is whether people include effects of ill health on income in health state valuations (HSV). The same holds for the actual inclusion in HSV of the effects of ill health on leisure. This study aims to test whether respondents to HSV using time trade-off (TTO) questions include income and leisure effects without instructions. Moreover, it tests the consequences of explicit instructions to include or exclude income effects. Methods: Three questionnaires were administered among the general public. Respondents were asked to value three distinct EuroQol descriptive system health states using TTO. In version 1, respondents were asked afterwards whether they included income effects. In versions 2 and 3, respondents were instructed upfront to include or exclude income effects. They were furthermore asked whether they included leisure effects. Results: A total of 222 respondents completed the questionnaire. In version 1, 64% of the respondents spontaneously included income effects, not resulting in differences in mean valuations. In versions 2 and 3, 88% included leisure time, resulting in a significantly lower TTO value in one health state. With explicit instructions, respondents instructed to include income gave lower HSV for the worst health state presented. Conclusions: Respondents in our sample did not consistently include income effects and leisure effects. Including income effects only had (some) effect on TTO valuations after an explicit instruction. If these results are confirmed in future research, this implies that income effects may be best captured on the cost side of the cost-effectiveness ratio. Spontaneous inclusion or exclusion of leisure time appeared to be more influential.

4.1 Introduction

Productivity costs remain a controversial cost category in the context of health economic evaluations. Not only is their inclusion debated, mostly in relation to the perspective chosen in the evaluation [41], also the correct way to estimate productivity costs is a much debated topic [12,47,102,105]. One of the ongoing debates is whether productivity costs are best measured at the cost or the effect side of the cost-effectiveness ratio in a cost-utility analysis [2,7,8,22,31,68,91,105]. This debate was especially triggered by the publication of the US guidelines on cost-effectiveness in 1996 [2], which specified that productivity costs should not be valued in monetary terms (and captured at the cost side of the cost-effectiveness ratio) as is normally the case. Rather, they were to be included at the effect side of a cost-effectiveness ratio. To ensure this, respondents in health state valuations (HSV) should consider the possible income effects of ill health, which, according to the US guidelines, would be the case as long as respondents are not explicitly instructed to exclude these effects. Although explicit instructions to include income effects could be considered superior, silence on this topic would be sufficient to ensure the incorporation of income effects by respondents in HSV [2]. This recommendation received quite some attention and criticism [8,22,31,91,105]. One of the main objections to measuring the effects of ill health on income through quality adjusted life year (QALY) elicitation is that it will not lead to accurate measurement of productivity costs [8,22,31,91]. Moreover, despite the suggestion of the US Panel to instruct respondents to include income effects, all but one of the most popular multi-attribute utility instruments (EuroQol-5D, Health Utilities Index, Quality of Wellbeing Scale, Rosser and 15D) remains silent on the matter. Only the Health Utilities Index is explicit in this respect and excludes rather than includes these effects [112]. This, by the way, is consistent with the original instructions for use of the time trade-off (TTO) [113]. The silence of the other instruments has been explained as a deliberate attempt to avoid alluding to income rather than an implicit way to include these effects [8]. But still, the current silence in most HSV exercises does not ensure the exclusion or the inclusion of income effects by respondents. This means that to some extent, income effects might be included in HSV, which would result in double counting when productivity costs are also captured at the cost side of the cost-effectiveness ratio. On the other hand, unless all respondents would include correctly estimated income effects in HSV, the approach advocated in the US guidelines will not result in an accurate estimation of productivity costs—even apart from other objections [68].

Unsurprisingly, therefore, it has been suggested that research should be undertaken to find out whether respondents consider income effects of reduced health in HSV when these are silent in this respect, whether this affects their subsequent valuations, and what effect explicit instructions regarding inclusion or exclusion would have on HSV [112]. To date, however, few studies have been conducted in this area. The results seem to indicate that a majority of the respondents do not include income effects in HSV [60,64,65,67]. Moreover, the effect of incorporating income effects in HSV appears to be extremely limited [65]. Still, the available evidence is limited, often used a visual analog scale (VAS) as valuation instrument and focused on spontaneous inclusion rather than on (the effect of) instruction. Still, explicit instructions may be considered the appropriate way to prevent either an underestimation or double counting of productivity costs [7,31,68,91]. Given the theoretical debate, two types of explicit instruction appear most relevant. First, one could instruct respondents to incorporate all possible effects of ill health on income in their valuations (in line with the recommendations of the US Panel). Second, one could oppositely instruct respondents to exclude these effects and measure all productivity costs in a monetary value (using the human capital or friction cost method). To our knowledge, the effects of such instructions have been tested only once [65]. Different explicit instructions on including or excluding the effects of ill health on income, or giving no instructions at all, did not lead to significant differences in HSV on a VAS. The latter may have been a limitation in establishing effects of income considerations in HSV because scale compatibility (i.e., the tendency of respondents to give more weight to the subject of focus in the response scale, potentially underweighting other aspects) may result in relative insensitivity of the VAS instrument regarding income effects [65]: the respondents may have been inclined to focus mainly on health itself instead of focusing on other aspects related to ill health.

Although there is debate in this area as well [114], the TTO method currently seems to be the preferred valuation technique for health states rather than standard gamble techniques, VAS, or person trade-offs, and indeed, much used tariffs have been based on TTO elicitations [108,115]. It seems especially relevant therefore to investigate the extent to which respondents consider income changes in TTO exercises and whether the TTO method is equally insensitive as the VAS seems to be to differences in explicit instructions on including or excluding income effects. This chapter reports on a study investigating this.

Although the main focus is on productivity costs and despite the fact that there is little debate on whether leisure time *should* be incorporated in the

QALY measure [2,7,68,91], it needs noting that the amount of empirical research on whether respondents in fact include leisure time in HSV is equally limited [64,65,67]. If respondents do not consistently include these effects on leisure, even though we expect them to, it is also important to see what effect this has on subsequent HSV. We explicitly highlight the effects of including the negative consequences of ill health for leisure time in TTO valuations in this chapter.

The structure of this chapter is as follows. First, we highlight the current knowledge in this area. Then, we present the design of the new study. Next, we highlight the results and subsequently discuss them.

4.2 Previous studies

As mentioned, few empirical studies have been performed to assess respondents' income (and leisure) considerations in HSV. To our knowledge, four studies have been conducted to find out whether respondents include the effects of ill health on income in HSV; the studies of Meltzer et al. [60], Sendi and Brouwer [64], Brouwer et al. [67], and the study of Krol et al. [65]. The first three studies focused on whether respondents include income effects and the latter additionally focused on the effect of instructions on including and excluding these effects. In the studies of Sendi and Brouwer [64], Brouwer et al. [67], and Krol et al. [65], respondents were furthermore asked whether they had included the effects of leisure time in their HSV.

Meltzer et al. [60] tested how giving information about the financial consequences of ill health affects respondents' considerations and valuations using the TTO method. Respondents were asked whether they had thought of financial consequences of illness while answering the TTO questions. Even when respondents were informed that no form of disability payment existed, less than 25% of the respondents included the financial consequences of illness in their valuations. Without any information on income losses, less than15% included income effects. Respondents who indicated to have (spontaneously) included income, without previous information on income loss, had lower TTO scores than respondents who did not consider income changes.

Sendi and Brouwer [64] performed a small sample test among 20 health professionals to find out whether the effects of ill health on income and leisure time would be included in HSV when no explicit instructions are provided. Respondents were asked to value a disease-specific health state on a VAS. Forty percent of the respondents included the effects of ill health on income, whereas 60% did not. The respondents who included income effects had a significantly lower mean VAS score than the respondents who did not include these effects. In the experiment of Sendi and Brouwer [(64)], respondents were also asked whether they had considered effects of ill health on leisure time. It turned out that 75% of the respondents indicated to have considered these effects. Although the valuations of respondents who had not included leisure time were higher than the valuations of those who did, the difference was not significant.

Brouwer et al. [(67)] asked 75 respondents of the general public to value three health states on a VAS. The respondents were asked afterwards whether they had considered income effects and leisure time effects of ill health. Respectively, 31% included income effects and 61% had included leisure time effects. Although income consideration did not result in significant changes in HSV, the incorporation of leisure proved to be influential in the valuation, but only for the most severe health state.

Krol et al. [(65)] asked 185 respondents of the general public, divided into three groups, to value three health states on a VAS. The first group was asked afterwards whether they had considered income effects in their valuations. Thirty-six percent spontaneously included income. The second group was instructed to consider income effects in their valuations and the third group was instructed to assume that income would not change due to ill health. There were no differences in the valuations of the three groups of respondents. The second and the third group of respondents were furthermore asked whether they had included the effects of ill health on leisure time. Eighty-four percent included these effects. There were no significant differences between the valuations of respondents who included or excluded effects on leisure time.

The results of the above studies seem to indicate that inclusion of income in HSV is inconsistent without explicit instructions. Nevertheless, the results of the studies of Brouwer et al. [(67)] and Krol et al. [(65)] suggest that inclusion of income effects does not change HSV when using a VAS as elicitation method. This may reflect that the VAS is rather insensitive in capturing these effects of ill health on income. The results regarding the effects of including leisure time on HSV are inconclusive.

This study aims to further investigate respondents' considerations on income and leisure in HSV using TTO, and additionally focuses on the effects of explicit instructions regarding the inclusion of income.

4.3 A new study

To increase the comparability with previous studies, we drew our study sample in a similar way as was done in the research of Krol et al. [65] and (with exception of the QALY-elicitation method) we used similar questionnaires.

To test the inclusion of income and leisure time as well as the influence of instruction, three different questionnaires were constructed. The study sample was drawn from the general public in various public places (mainly individuals traveling by public transportation). Respondents were approached and asked whether they were willing to participate in this study. If they agreed to participate, the respondents in our convenience sample were randomly handed one of the three different versions of the questionnaires for self-completion. A total of 240 questionnaires were administered. In all three questionnaires, respondents were first asked some background questions about sex, age, income, education, and their current health state using the EuroQol descriptive system (EQ-5D) and the EuroQol VAS. Next, the respondents were asked to value three ill health states that differed in severity, and were chosen from the 243 possible health state scenarios with the EQ-5D, with the standard Measurement and Valuation of Health group (MVH)_A1 scores of, respectively, 0.88 (health state 21211), 0.587 (health state 22221), and -0.043 (health state 33312) to have a large spread in health states (table 4.1).

In the study of Krol et al. [65], respondents were asked to value the same three health states on a VAS. Now, respondents were asked to value the health states with the use of the TTO method. The respondents were asked to state how many years of full health they considered to be equally good as living 10 years in the given ill health state (see Appendix 4A). Respondents also had the opportunity to choose not to give up life years and live the full 10 years in the given ill health state. Respondents were informed that the ill health state would remain the same the full 10 years and that life would end after these 10 years in ill health

Table 4.1 The health states presented in the questionnaires

Health state	EQ-5D code	MVH_A1* score	Description
1	21211	0.880	Some problems in walking, no problems with self-care, some problems performing usual activities, no pain or discomfort, not anxious or depressed.
2	22221	0.587	Some problems in walking, some problems with self-care, some problems performing usual activities, moderate pain or discomfort, not anxious or depressed
3	33312	-0.043	Confined to bed, extreme problems with self-care, not able to perform usual activities, no pain or discomfort, moderate anxious or depressed

* MVH = Measurement and Valuation of Health group

or the (less than10) years chosen in full health. Respondents were instructed to make choices considering themselves and reflecting their own opinion.

In version 1 of the questionnaires, respondents did not receive directions on whether to incorporate or to ignore the effects of ill health on income while making the trade-offs. The respondents were asked whether they had included income effects, after the valuation process. If they indicated they had included these effects, they were asked to value the same three health states again, but now with the explicit instruction to ignore the effects of ill health on income by assuming that their income would not change due to ill health. In version 2, people were explicitly instructed upfront to assume that income would not change due to ill health, and thus to ignore the effects of ill health on income. In version 3, respondents were oppositely instructed to incorporate the possible effects of ill health on income. In all questionnaires, respondents were asked whether they thought being in the different health states would decrease their income and if so, to what extent. This was always done after the valuation process in order not to 'contaminate' the outcomes.

In versions 2 and 3, respondents were furthermore asked whether they had included the possible effects of ill health on leisure time in their valuations and if they thought the three health states presented would affect their leisure time. The latter aspect was asked because the differences in the extent to which respondents believe that the three ill health states will affect their leisure time can be of influence on their valuations. If, for instance, all respondents expect that the health states will have no effect on their leisure time, it could be expected that including or excluding leisure time considerations in the valuation process will have no effect on subsequent valuations. In version 1 (with the double HSV), the leisure time questions were not included to limit the size of the questionnaire.

The questionnaire was designed to confirm the following hypotheses, based on theoretical debate and the results of previous studies.

1. Without receiving directions on including or excluding the effects of ill health on income or leisure time, some respondents to HSV will and some will not include these effects in their valuations [60,64,65,67].
2. Respondents of HSV who spontaneously include the effects of ill health on income do not significantly value health states different from respondents who do not automatically include these effects [65,67].

3. Respondents of HSV who include the effects of ill health on income without instructions, will value the same health states higher the second time when asked explicitly to exclude these effects [64,65].
4. Respondents of HSV who are explicitly asked to include the effects of ill health on income value health states not significantly different from respondents who are explicitly asked to exclude the effects [65].
5. Respondents of HSV who include the effects of ill health on leisure time, value health states lower than respondents who do not include these effects [67].

Categorical variables were compared with the chi-square test, ordinal variables with the Mann–Whitney U-test (for two groups) or the Kruskal–Wallis test (for more than two groups). Continuous variables are compared with the Student's *t* test (two groups) and with one-way analysis of variance (two or more groups). Simple linear regression models and linear mixed effects (LME) models were used to estimate the association between the valuations of the health states and other variables such as the version of the questionnaire administered and characteristics of the respondents as age, sex, and income. The data were analyzed with use of SPSS Statistical Software Package 13.0 (SPSS, Chicago, IL) and S-Plus 6.2 Professional (Insightful S-Plus, Seattle, WA).

4.4 Results

Sample population

A total of 222 respondents completed the questionnaire. Eighteen of the administered questionnaires were not returned or returned blank. Twelve questionnaires were excluded from further analysis based on comments of the respondents that they did not understand the TTO method or based on evidence that instructions were not followed (e.g., respondents who did not trade-off healthy years, but instead wanted more than 10 years in full health in return). A total of 210 questionnaires were used for further analysis (version 1; n = 72, version 2; n = 75, version 3; n = 63). As seen in table 4.2, the average age of the respondents was 34.8 years and did not differ between the versions of the questionnaire. Fifty percent of the respondents were female. The average self-reported health of the respondents was 0.82 reported on a VAS. Higher educated people and those with higher incomes were overrepresented in our sample. The respondents with higher incomes were significantly older (P <

Table 4.2 Characteristics of respondents (SD between brackets)

Characteristic	Questionnaire administered			All respondents
	Version 1 (n = 59)	Version 2 (n = 64)	Version 3 (n = 62)	(n = 185)
% Females	54	44	54	50
Age	36.0 (14.4)	35.1 (14.4)	33.0 (14.2)	34.8 (14.3)
% Employment				
No paid work	36	23	29	29
Fulltime employment	31	43	40	38
Part-time employment	33	34	31	33
% Education [a]				
Lower education	11	13	21	15
Medium education	39	39	41	40
Higher education	50	48	38	46
% Income [b]				
Lower income	26	21	25	24
Medium income	26	34	33	35
Higher income	47	45	41	45
Mean SAH with VAS (SD)	0.82 (0.14)	0.83 (0.12)	0.82 (0.15)	0.82 (0.14)

[a] These groups were defined according to the Dutch educational system. The category 'lower education' comprises elementary school, lower general secondary education (MAVO) and pre-vocational (VBO). 'Medium education' comprises higher general secondary education (HAVO, VWO) and senior vocational education (MBO). 'Higher education' comprises vocational colleges (HBO) and academic education.

[b] Based on net income per month: low = <€650, medium = €650–1900, high = >€1900,
SAH = self-assessed health, VAS = Visual Analogue Scale, y = years, SD = standard deviation

0.001). The respondents of the different versions of the questionnaire did not differ in income, education, sex, employment, or self-reported health.

Effects of ill health on income

As seen in table 4.3, 28% of the respondents stated that ill health state 1 (the 'mildest' state) would decrease their income. Respectively, 71% and 89% expected income loss in health state 2 or 3. Although especially for less severe health states explicit instructions may have induced more respondents to expect income changes, these differences were never significant at the 5% confidence level. Respondents who thought that the health states would decrease their income were younger than those who did not expect income influence ($P < 0.01$ for all health states). Respondents who stated that health state 2 or health state 3 would not change their income were more often unemployed ($P < 0.01$). Respondents who stated that health state 3 would decrease their income were better educated ($P = 0.05$) as opposed to those who did not.

The respondents who expected income loss to occur expected on average that health state 1 would lead to a loss of 284 euros per month, while this

Table 4.3 Effects of ill health on income and leisure-time (SD between brackets)

	Health state 1 decreases income	Health state 2 decreases income	Health state 3 decreases income	Included income	Included leisure time
% Version 1 (n = 72)	21	61	88	64	-
% Version 2 (n = 75)	32	76	91	-	90
% Version 3 (n = 63)	32	77	86	-	85
Mean %	28	71	89	-	88
P-value [a]	0.228	0.077	0.758	-	0.371
Expected income loss (€)	284 (205)	421 (330)	787 (461)	-	-

[a] Chi-Square Test

amount was 421 and 787 for health states 2 and 3, respectively. We used regression techniques to see which variables are associated with expected income loss. The valuations of respondents were not significantly associated with the expected amount of income loss. A higher income was associated with a higher expected income loss, but only in the worst health state 3 (P = 0.05). Other demographic variables such as age, sex, and education turned out not to be significant for all the three health states.

The effects of instructions

No instructions

Confirming our hypothesis (1) that without instructions, some respondents will and some respondents will not include the effects of ill health on income in HSV, 64% of the respondents of version 1 automatically included the possible effects of ill health on income in their HSV and 36% did not. This percentage of inclusion is higher than those found in previous studies, where it never exceeded 40% [60,64,65,67]. The respondents who automatically included the effects of ill health on income did not differ from those who did not, in education, income, employment, sex, age, or self-reported health.

Although Meltzer et al. [60] and Sendi and Brouwer [64] report differences in HSV based on different considerations about negative effects of ill health on income (assessed without explicit instruction), our overall results do not show a significant influence of including or excluding of the effects of ill health on income without instructions on the subject (see table 4.4). These findings support the hypothesis (2) that respondents of HSV who spontaneously include the effects of ill health on income do not significantly value health states differently than respondents who do not spontaneously include these effects. These findings are similar to the findings in the studies of Krol et al. [65] and Brouwer et al. [67].

Table 4.4 health state valuations

		Valuations					
Version	Group	HS1	n	HS2	n	HS3	N
	The effects of instructions						
1	All (no instructions)	0.88	72	0.72	72	0.39	71
2	All (excluding income)	0.91	74	0.74	74	0.45	73
3	All (including income)	0.91	62	0.72	62	0.38	60
	P-value [a] (1 vs. 2 vs. 3)	*0.329*		*0.820*		*0.279*	
	P-value [b] (1 vs. 2)	*0.201*		*0.537*		*0.188*	
	P-value [b] (1 vs. 3)	*0.209*		*0.867*		*0.911*	
	P-value [b] (2 vs. 3)	*0.985*		*0.676*		*0.181*	
	Spontaneous inclusion/exclusion of income						
1	Including income	0.89	45	0.72	45	0.41	45
1	Excluding income	0.86	25	0.70	25	0.35	24
	P-value [b]	*0.494*		*0.752*		*0.364*	
	Excluding income after spontaneous inclusion of income						
1	1st valuation (+ income)	0.88	34	0.73	34	0.39	34
1	2nd valuation (- income)	0.87	34	0.75	34	0.42	34
	P-value [c]	0.343		0.226		0.189	
	Income affected & income included/excluded						
1	Income affected [d]	0.79	14	0.70	41	0.40	60
1	Income not affected	0.90	56	0.74	28	0.34	9
	P-value [b]	*0.008*		*0.498*			*0.508*
1	Income included + affected [d]	0.78	6	0.71	24	0.42	40
1	Income excluded + affected [d]	0.79	8	0.70	17	0.38	20
	P-value [b]	*0.955*		*0.908*			*0.579*
2	Income affected [d]	0.93	23	0.72	56	0.44	64
3	Income affected [d]	0.87	19	0.68	43	0.34	50
	P-value [b]	*0.173*		*0.433*		*0.055*	
	Leisure-time (LT) affected & leisure-time included/excluded						
2 & 3	Including LT	0.90	113	0.71	113	0.39	110
2 & 3	Excluding LT	0.94	16	0.84	16	0.48	16
	P-value [b]	*0.242*		*0.037*		*0.235*	
2 & 3	LT affected [e]	0.88	85	0.70	113	0.40	122
2 & 3	LT not affected	0.96	45	0.86	14	0.60	5
	P-value [b]	*0.000*		*0.021*		*0.123*	
2 & 3	LT included + LT affected [e]	0.88	74	0.69	97	0.38	103
2 & 3	LT excluded + LT affected [e]	0.89	8	0.81	13	0.45	15
	P-value [b]	*0.866*		*0.091*		*0.387*	

[a] One-way ANOVA, [b] T-test for equality of means, [c] paired sample t-test

[d] The valuations of respondents that think the health state will influence income. The valuations of respondents who do not think the particular health state will affect income were removed.

[e] The valuations of respondents that think the health state will influence leisure time. The valuations of respondents who do not think the particular health state will affect leisure-time were removed.

In table 4.4, the average TTO valuations of the three health states of the respondents of the three different questionnaires are presented. Our respondents valued the health states higher than the corresponding MVH_A1 scores: 0.90 versus 0.88, 0.73 versus 0.59, and 0.41 versus -0.04 for health state 1 to 3, respectively. The respondents in our sample were reluctant to trade-off life years: 58%, 31%, and 12% of the respondents were unwilling to trade-off life years for health states 1 to 3, respectively. There were no differences between the versions in percentages of respondents who would not make the trade-off. Respondents who would not give up time were significantly older than those who would in all the three health states. Furthermore, the respondents who would not give up years in health states 1 and 2 had significantly higher incomes than those who did make the trade-off. This latter difference was not found in health state 3.

Instruction after silence

As mentioned, the respondents of version 1 who included income in their valuations were asked to value the same health states again, now assuming these health states would not affect their income. Although other studies [64,65,67] already show that a part of the respondents does not react to a change in instructions by changing their valuations, it was expected that overall, the respondents would value the same health states significantly higher the second time.

As seen in table 4.4, we found no significant differences between the first and second valuations of the respondents. These findings are similar to the findings of Brouwer et al. [67] and do not support the hypothesis (3) that respondents change their valuations if they are asked to ignore a negative aspect of ill health they had included before.

Explicit instructions

Table 4.4 shows that the respondents of version 2 (who were explicitly asked to exclude the effect of ill health on income), did not value the health states significantly higher than the respondents of version 3 (who are explicitly asked to include the effects), similar to the findings of Krol et al. [65]. The valuations of the respondents without instruction (version 1) moreover did not differ from those in the other two groups. These results, thus, seem to imply that instruction does not matter in the sense that it does not change the mean TTO valuations. Only considering respondents (in versions 2 and 3) who actually expected the presented health state to affect income in our analyses, regardless of the instruction given, increases the differences in mean valuations of the respondents in the expected direction (respondents of version 2 give higher values to the health states). These differences, however, are still insignificant, as shown in table 4.4.

Multivariate analysis

A linear regression model was used to estimate the association between the HSV of all respondents and the version administered and whether respondents expect a health state to affect income, regardless of whether income effects were included (interaction term included), corrected for age, sex, income, and self-reported health.

As shown in table 4.5, income had a significant positive association with the valuations of health states 1 and 2. In health state 1, the interaction term of the version of the questionnaire and whether respondents expect that the health state would actually decrease income turned out to be significant. Respondents who were handed version 1 or version 3 of the questionnaire and expected the relevant health state to decrease income therefore valued that health state lower than the respondents in version 2. These interaction terms were not significant for the valuation of health state 2; however, for health state 3 the interaction term of version 3 and expecting income to be affected was again significant. These results imply that the different instructions do seem to affect respondents' valuations, but only if respondents expect that the health states presented would actually decrease their income. Note that the linear regressions have little explanatory power.

Table 4.5 Results of regression analyses (n = 199)

	Valuation HS 1.		Valuation HS 2.		Valuation HS 3.	
	Beta	P-value	Beta	P-value	Beta	P-value
Constant	8.057	.000	6,403	.000	2.445	.174
Sex (male = 1, female = 0)	-.005	.544	.020	.202	.024	.207
Age	-.235	.240	-.661	.052	-.704	.080
Net income	.232	.004*	.374	.007*	.089	.578
Self-reported health	.557	.438	-.339	.779	1.533	.288
Version 1 vs. 2	.051	.854	-.457	.527	-1.489	.291
Version 3 vs. 2	.421	.153	.743	.387	2.552	.094
Does HS 1, HS 2 or HS 3 affect income (1 = yes, 0 = no)	.405	.255	-.378	.574	-.053	.964
Dummy version1*income affect HS 1, 2 or 3	-1.691	.003*	.093	.915	1.101	.461
Dummy version3*income affect HS 1, 2 or 3	-1.063	.043*	-1.256	.201	-3.594	.027*
Adjusted R^2	.094		.098		.094	

* $P < 0.05$

Next, we focused on the effects of explicit instructions (as given in versions 2 and 3 of the questionnaire). To that end, an LME model was used and respondents of version 1 of the questionnaire were excluded from the analysis. The LME model was used to estimate the association between the valuation of the health states and explicit instructions, stratified by whether respondents thought income would be affected by the respective health state, adjusted for age and sex. Respondents valuing the three health states were treated as random effects to account for within respondent variability and correlation using an unstructured correlation matrix.

Because the interaction term 'version*income-affected*health state valuations' in the LME model turned out to be significant (P = 0.048), the effect of explicit instruction on HSV was analyzed separately in each subgroup of respondents who expected income to be affected in, respectively, health states 1, 2, or 3. A linear regression model was used to separately test whether the valuations of the three health states were affected by the explicit instructions in versions 2 and 3. Covariates were added to the respective univariate model to improve the estimate and precision of the effect estimate in case the univariate model showed a P-value of 0.1 or less. Furthermore, in all models, a multivariate model was also used as a starting point and the model was then simplified to assess the effect of including/excluding covariates to the respective model on the main effect estimate of interest. Because the leisure time questions were only asked in versions 2 and 3, respectively, they are also included in the multivariate model. The multivariate model included the variables age, sex, consideration of income and leisure in HSV, and own VAS.

As shown in table 4.6, in the subgroup expecting the health state to affect their income, the explicit instruction variable was significantly associated with the valuation of health state 3, with lower scores for respondents including income. Also the variables sex, expecting the health state to affect leisure, and self-reported health were significantly associated with the valuation of health state 3.

The results of the subgroup analysis of versions 2 and 3 imply, similar to the analyses of the three versions together, that explicit instructions may have an effect on valuations of health states, however only in respondents who believe their income will be affected by this health state. In other cases, however, explicit instruction did not lead to changes in valuations. This means that the findings of Krol et al. [65] are largely but not completely reproduced in this study and that hypothesis (4) is for most situations, but not for all, confirmed. The picture is less uniform than in previous studies therefore.

Table 4.6 The effect of instruction on health state valuation stratified by whether respondent thinks health state affects income (TTO)

			Univariate model			Multivariate model*		
HS	Income affected	variable	Coefficient	SE	P-value	Coefficient	SE	P-value
1	Yes	Version 3 v 2	-.620	.447	.173			
1	No		.427	.287	.140			
2	Yes		-.387	.492	.433			
2	No		.394	.866	.652			
3	Yes	Version 3 v 2	-.975	.522	.064	-1.062	.500	.036**
		Sex				-1.479	.499	.004**
		Leisure affected				-6.507	2.606	.014**
		Self-reported health				4.116	2.052	.047**
3	No	Version 3 v 2	1.7381	1.9602	.3942			

* R^2 0.16, P-value of multivariate model = < 0.001
** P < 0.05
SE = Standard error

Leisure time

Eighty-eight percent of the respondents of versions 2 and 3 included the effects of ill health on leisure time while trading-off time versus health. Even though respondents are assumed to include leisure time [(2,68,91)], 12% of the respondents in our convenience sample did not. This supports hypothesis 1, that without instructions, some respondents will and some will not include these effects. The percentage of respondents who included leisure time is comparable with the 84% of the study of Krol et al. [(65)]. It is, however, higher than the percentage found in previous studies by Sendi and Brouwer [(64)] (75%) and Brouwer et al. [(67)] (61%).

The 88% of the respondents who included leisure time in their valuations had a lower income than the 12% who did not include leisure time (P = 0.03). As shown in table 4.4, respondents including leisure time gave lower TTO values to the three health states than respondents who had not included leisure time. The differences, however, were only significant for health state 2. These results can therefore not completely support the hypothesis (5) that respondents who include the effects of leisure time will value the health states lower than those who do not include these effects. These findings are similar to the findings of Brouwer et al. [(67)], who also found significant differences for only one health state (the worst health state in that case).

As mentioned, we asked the respondents whether they expected the three health states presented to affect their leisure time. Sixty-five percent, 89%, and 96% of the respondents believed that, respectively, health states 1 to 3 would affect their leisure time. We excluded the respondents who thought that the health states would not affect their leisure time. Next, we tested again whether including or excluding leisure time leads to differences in valuations to see if the differences would have increased. Opposite from what we expected, the differences between the valuations of respondents including and excluding income decreased after correcting for thinking a health state would or would not affect leisure time, as shown in table 4.4. The correction causes the difference earlier found in health state 2 to lose its significance. Interestingly, whether a respondent thinks the health state will affect leisure time is (regardless of whether a respondent included or excluded leisure time effects) of significant influence on respondents' valuations for health state 1 and health state 2, though not for health state 3. The latter result, also given that the difference between the groups grows with the severity of the health state, is probably due to lack of power given the low number of respondents thinking that leisure would not be affected. Indeed, in the study of Brouwer et al. [67], the fact that leisure was affected was most influential as well, although, after correction for other variables, it was only for the most severe health state that a significant difference was found between respondents who believed leisure to be affected and those who did not (in the expected direction).

4.5 Discussion and conclusions

This chapter has reported on the first study aimed at investigating the effects of explicit instructions regarding the inclusion or exclusion of income and leisure in HSV using TTO. In this study, five hypotheses were sought to be confirmed. In line with the results from previous studies, our study confirmed the first hypothesis that without instructions, respondents do not consistently include or exclude the effects of ill health on income. Given the high inclusion of income in this study relative to previous studies, the TTO method may be more sensitive to inclusion of income effects than the VAS. Scale compatibility may be an explanation for this [109]. On the other hand, Meltzer et al. [60] used TTO questions (and even triggered respondents to consider income effects by giving specific information about disability payments) but report lower percentages of inclusion of income by respondents. Differences in sample population, country, and methodology may have contributed to these differences.

Our second hypothesis was also confirmed. The respondents spontaneously including the effects of ill health on income did not value the three health states significantly different when compared to respondents not including these effects. Surprisingly, for all three health states, respondents including income gave higher valuations, although these differences were insignificant.

Hypothesis 3 was not confirmed. Respondents who, without instructions, included income effects did not value the three health states significantly higher the second time when asked explicitly to exclude these effects. Nevertheless, the valuations of the more severe health states 2 and 3, where potential effects of differences in including or excluding income effects would be strongest, did change in the expected direction.

Results regarding the fourth hypothesis were mixed. In most cases, explicit instructions on including or excluding income had no significant effect on HSV; although the group instructed to include income elicited lower valuations of health states 2 and 3. The effect of explicit instruction on the HSV was only significant in some cases where respondents actually expected an income effect.

For the final hypothesis, results were also mixed. Respondents spontaneously including leisure time effects gave lower HSV than respondents excluding leisure. The differences in valuation were however only significant for one of the three health states.

Before discussing the implications, some important limitations of this study need to be stressed. First of all, our study is based on a relatively small convenience sample of respondents. Some of the differences in HSV between groups may have been insignificant because of lack of statistical power, especially in cases where differences in the expected direction were observed but did not reach conventional levels of significance (e.g., for hypotheses 3, 4, and 5). Therefore, repeating this study in a larger sample remains important. Using a wider variation in valued health states in such studies may also be informative. Second, the format of our TTO questions might have leaded some respondents to believe that the minimum trade-off they could make was 1 year. The respondents indeed did only trade-off whole or half years. This may have caused respondents willing to trade-off some time, but less than a year, to choose to not give up years at all. The respondents in our sample indeed were reluctant to trade-off life years, although the reluctance to trade-off life years is common [(116)]. Potential bias caused by our TTO framing would be strongest in the mildest health state. It is however not expected to affect potential differences in valuations between the three health states and our focus was not on finding the 'right' TTO values, but on finding the effects of instructions on income in HSV. Third, the results of our study may be country-specific. In countries with a less comprehensive social security system than The Netherlands, respondents might be more inclined to consider income and ill health may affect income more strongly, leading to more respondents including the effects of ill health on income or perhaps even to significant differences in HSV between respondents including and excluding income effects. This could be investigated further. A fourth limitation of our study is that we did not ask the respondents afterwards whether they acted according to the instructions given. It could be that a part of the respondents ignored the instructions. Furthermore, it is possible that through the instruction to ignore possible income effects, respondents are actually triggered to do think of income ('the pink elephant effect' (68)). It is also possible that any explicit instruction places too much emphasis on income, causing respondents to overweigh possible income effects in the HSV. This forms a potential bias if we, in daily practice, consider instructing respondents to HSV on including or excluding income. Finally, we asked the respondents the extent to which they believed that the ill health states would affect their income. It, however, remains questionable whether respondents can make realistic estimations of income losses due to illness (67). This idea is strengthened by the high number of missing values when a specific amount of expected income loss is asked (missing values health state 1: 35%, health state 2: 33%, and health state 3: 40%). One way to circumvent this lack of knowledge is to include explicit information on reasonable estimates of income losses due to certain health states. A similar procedure was followed by Meltzer et al. (60).

In conclusion, our results are generally in line with earlier studies and indicate that silence regarding inclusion of income and leisure does not ensure consistency in this respect in HSV. Some respondents will include these effects while others exclude them. Moreover, the extent to which respondents expect health states to affect their leisure and income varies substantially across respondents. Still, as found in some previous studies, spontaneous inclusion or exclusion of income effects did not result in different valuations of health states. The influence of explicit instructions regarding inclusion of income effects appears to be limited, but still, in contrast to earlier findings does appear to matter. Because we do find some effects of including income after an explicit instruction on the matter, the TTO method may be more sensitive in picking up income effects than the VAS. Especially when respondents believe that income is indeed affected (which is more often the case in severe health states), explicit instruction to exclude income may lead to higher valuations. For leisure, the percentage of respondents including leisure time was relatively high in our study. Still, like in other studies, a substantial part of the respondents do not include leisure time in their valuations although they are normally expected to do so. We find no clear evidence that inclusion as such leads to differences in valuations. Expecting leisure to be affected seems more influential than the indicated inclusion of leisure in HSV, which may indicate simply a difference in perception of the severity of the health states between respondents.

What do these results imply then for productivity costs and leisure? First of all, it is important to realize that, given the limitations of our study and the previous studies conducted, it is premature to draw definite conclusions. There is a clear need for more research in this area using larger samples to further explore this important topic.

Notwithstanding this, it is clear that currently, most HSV methods include no instructions to the respondent regarding the inclusion or exclusion of income effects. In this case, respondents spontaneously include or exclude income. As we have seen, spontaneous inclusion or exclusion of income changes does not appear to result in noticeable differences in HSV even when such changes are expected to be large. If this finding is confirmed in future research, this casts empirical doubts on recommendations to include income changes in health-related quality of life, next to theoretical objections one may have against it (e.g. [(22)]). In that case, it would imply that it is best to include productivity costs on the cost side of the cost-effectiveness ratio, where they do have a measurable effect. If one actively wants to preclude income effects in HSV, explicit instructions to exclude income changes due to ill health may be considered appropriate. It must be noted, however, that such instructions may involve potential biases, like overweighting income, due to an active, although negative, emphasis on income considerations. Rather than telling respondents what not to consider, and thus risking that they will consider it, it may be better to tell them what to consider; i.e., they should consider that they are covered by full health insurance and salary continuation insurance [(113)]. More research here would clearly be useful. For now, silence on the matter may be considered the best solution, because it is unclear whether a move away from silence, and what kind of a move, is for the better or the worse.

Although in all the empirical studies on leisure time in HSV (including ours) small samples were used, it seems clear that most people include this in HSV. Not all respondents, however, automatically include these effects, and therefore, it may be necessary to explicitly instruct respondents to include these effects. But even then, more research in this area seems worthwhile because it remains unclear whether including leisure results in an adequate valuation of lost leisure. Silence may not be golden therefore, but how to break it best remains open for debate.

Appendix

4 A The TTO exercise

In this part of the questionnaire subsequently three health states are described. You are asked to value these health states. We are interested in your opinion and the choices you would make for yourself.
*Imagine that the health states remain unchanged for ten years en that life ends after this period. You are asked to state **how many years in full health for you would be equal to 10 years in the given health states**. You also have the choice to state you would rather live 10 years in the given (less than perfect) health state instead of given up life years for better health.*

Health state 1.
Some problems in walking
No problems with self-care
Some problems performing usual activities
No pain or discomfort
Not anxious or depressed

Mark and give your opinion on the dotted line

Ten years in **health state 1** is equal to...............years in full health

Or, if this is your preference, mark:

I rather live 10 years in **health state 1**, than giving up life years for full-health

5.

Does the EQ-5D reflect lost earnings?

Based on: Tilling C, Krol M, Tsuchiya A, Brazier J, Exel J, Brouwer W.

Does the EQ-5D reflect lost earnings?

Pharmacoeconomics 2012 Jan;30(1):47-61.

Summary

Background: An important methodological issue in economic evaluations of health care is how to include productivity costs (the costs related to reduced productivity due to illness, disability and premature death). Traditionally, they were included in the numerator of a cost-effectiveness analysis, through either the human capital or the friction cost method. It has been argued, however, that productivity costs are already included in the denominator (i.e. in the QALY measure) because respondents consider the effect a given health state will have on their income when valuing health states. If that is the case, many previous economic evaluations might have double counted productivity costs by including them in both the numerator and the denominator. Aim: The aim of this study was to determine whether respondents valuing EQ-5D health states using the time trade-off (TTO) method spontaneously consider income effects, whether this consideration influences subsequent valuations and whether explicit ex post instructions influence valuations. Methods: Through an online survey, we asked 321 members of the Dutch general population to value four EQ-5D health states through three different TTO exercises. The first exercise was a standard TTO question. Respondents were then asked whether they had included income effects. Depending on their answer, the second TTO exercise instructed them to either include or exclude income effects. The third TTO exercise provided explicit information regarding the income loss associated with the health state. Results: Data were available from 321 members of the Dutch general public. Of these respondents, 49% stated they had spontaneously included income effects. Twenty-five percent of the sample did not trade any time in any of the TTO exercises and these respondents were excluded from the analysis. Results of t-tests showed there were only weakly significant differences in valuations for one health state between those who spontaneously included income effects and those who did not. Explicit instruction led to some significant differences at the aggregate level, but the effect was inconsistent at the individual level. When explicit information on the amount of income loss was provided, all states were valued lower when associated with a larger income loss. Conclusions: This study offers further evidence indicating that income losses do not significantly affect health state valuations.

5.1 Background

Economic evaluation is increasingly used to inform the allocation of scarce health care resources. Many national decision-making bodies, such as the National Institute for Health and Clinical Excellence (NICE) in the UK [(72)] and the Dutch Health Insurance Board, [(74)] require or encourage these economic evaluations to take the form of a cost-utility analysis, with the QALY as the measure of health benefit. The QALY assigns a 'utility' value, representing average preferences, to a given health state. The values range from 0 to 1, where 1 represents 'perfect health' and 0 represents 'dead or equivalent'. Values less than zero can be used to represent states considered to be 'worse than dead'. The number of QALYs over a given period of time is then calculated by multiplying the utility values of the relevant states with the duration of these states. The quality adjustment in the QALY is commonly informed by national population value sets. Most of these value sets measure changes in health status on the EQ-5D[7] descriptive system and use the time trade-off (TTO)[8] method of preference elicitation (e.g. in the UK, [(108)] the Netherlands [(115)], the US [(117)] and Germany [(118)]).

A potentially important component in economic evaluations is productivity costs, defined as "*costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid*." [(8)] In some countries, guidelines recommend the exclusion of such costs from economic evaluations (e.g. the UK [(72)]), while in other countries the inclusion of these costs is advocated (e.g. the US [(2)] and the Netherlands [(74)]). This relates to the perspective encouraged in such guidelines (i.e. a broad societal perspective or a narrower health care perspective).

In a recent review, Tilling et al. [(92)] discussed the various different methods for including productivity costs in economic evaluations. These costs are normally included in the numerator of the cost-effectiveness ratio, using either the human capital approach [(88,119)] or the more recently developed friction cost

7 The EQ-5D descriptive system has five dimensions and three levels per dimension, giving a total of 243 health states. For example, 22322 describes the following state: some problems with walking about, some problems with washing and dressing, unable to perform usual activities, some pain or discomfort and moderate anxiety and depression.

8 The TTO method typically asks respondents to make a choice between living in a less-than-perfect state of health for 10 years and living in full health but for a shorter period of time. Commonly, a point of indifference is obtained through an iterative process. If an individual were indifferent between living for 10 years in poor health and for 6 years in full heath, the QALY adjustment for that health state would be 6/10 = 0.6.

method. [90] The latter commonly produces lower estimates of productivity costs by allowing for the replacement of sick individuals by previously unemployed individuals. The practice of valuing productivity costs in monetary terms was challenged in 1996 by the 'Washington Panel'. [2] In its influential guidelines, the panel recommended measuring and valuing most of the productivity costs in terms of quality of life, thus including them in the denominator of the cost-effectiveness ratio. Changes in income are then used as a proxy for productivity costs. In other words, the panel assumed that when people value health states, they take into account the effect of ill health on their ability to work and hence on their subsequent income (even when the question is silent on the issue). If this assumption holds, health state valuations, and consequently national value sets for measures such as the EQ-5D, already (to some extent) incorporate the impact of ill health on productivity. The panel, therefore, argued that to include changes in productivity in the numerator as well would constitute a form of double counting.

Regardless of whether an economic evaluator wishes to include productivity costs, and regardless of the method of inclusion, a key question must be answered to establish confidence in the accuracy of economic evaluations: do health state valuations, and therefore the published national value sets for the EQ-5D, reflect the impact of lost earnings due to ill health? If they do, previous studies including productivity costs in the numerator of the cost-effectiveness ratio might have double counted these costs. Alternatively, studies wishing to exclude these costs (e.g. in taking a strict health care perspective) might have implicitly included them through the use of 'polluted' value sets.

The review by Tilling et al. [92] outlined the existing empirical studies that have attempted to establish whether or not respondents in health state valuation exercises consider the potential impact of lost earnings. The review identified eight studies; [60-67] four of the studies used the EQ-5D [63,65-67] and three of these found that the majority of respondents did not spontaneously include income effects. [63,65,67] No firm conclusions can be drawn from the four studies on the effect of spontaneous inclusion on valuations or on the effect of explicit instructions, although the influence of considering income appears to be limited. The remaining four studies asked respondents to value blindness using TTO, [60] multiple sclerosis using a visual analogue scale (VAS), [64] carpal tunnel syndrome using standard gamble [61] and visual impairment using TTO. [62] Three of these studies found the majority of respondents did not spontaneously include income effects. [60,61,64] They also found evidence that spontaneous inclusion of income effects significantly affected valuations. However, these studies were hampered by small numbers of respondents, mostly from convenience samples. The differences in results among the studies make it difficult to draw firm conclusions regarding the inclusion of income effects and the potential effects this may have on health state valuations.

The study presented here adapted the study design used to generate population value sets for the EQ-5D, as first used in the Measurement and Valuation of Health (MVH) study [(108)], and carried out valuations of hypothetical EQ-5D states using TTO exercises through an online survey administered in the Netherlands. The study used a number of different TTO questions to explore the impact of losses in income on the valuation of hypothetical health states. Specifically, our objectives were to (i) examine whether EQ-5D health state values, obtained through online TTO exercises, reflected losses in income due to ill health; (ii) examine the impact on the health state values of including specific ex post instructions to consider, or not to consider, income changes when hypothetical EQ-5D states are valued; and (iii) examine how the impact on the health state values was distributed across the five different dimensions of the EQ-5D.

5.2 Methods

Sampling

Data were gathered through an online self-complete questionnaire, presented in Dutch, in the Netherlands. Invitations were sent out to potential respondents, within an existing Internet panel, in order to obtain a representative sample (in terms of age, sex and level of education) of 300 members of the Dutch general public between the ages of 18 and 65 years. This age range was chosen since health changes in people within this age range are most likely to affect their income. An online market research company (Survey Sampling International; www.surveysampling.com) performed the data collection.

Background, ranking and visual analogue scale

All respondents were asked a number of background questions regarding age, sex, education, marital status and occupation. In addition, the number of children, net personal income and net household income were included to help us understand the effect that dependents and personal income have upon the propensity to include income effects. Furthermore, to ensure representativeness, ethnic origins and religion were included to capture the diverse nature of

Dutch society. Following the questions regarding background characteristics, respondents were asked to describe their own health through the EQ-5D descriptive system. Respondents were next asked to rank four hypothetical EQ-5D health states (see The Health States section for details), in addition to the following states: (i) full health; (ii) dead; and (iii) 'your own health today'. They were then asked to place the same seven states on a standard EQ-5D VAS.

The time trade-off exercises

The TTO format used a 10-year time horizon. No iterative process was used. Respondents were simply asked how many years living in full health was equivalent to 10 years living in health state X. No visual aid was used. The main part of the survey consisted of a number of different TTO questions, as outlined in Appendix 5A.

Three versions of the questionnaire were used, with allocation of respondents being determined randomly. The versions differed only in terms of the levels of income loss (20%, 40% and 60%) used in the third TTO question, TTO3. Respondents first valued the four hypothetical EQ-5D health states through the first TTO question, TTO1 (the states were the same as those the respondents encountered in the VAS and ranking exercises). They were then asked a follow-up question on whether they had considered income effects in these valuations. Depending on their answer, in the second TTO question, TTO2, respondents were given instructions to either include or exclude income effects. In TTO3, respondents were given information about the specific level of income loss they would incur in the health state; after they valued each state in TTO3, they were asked if they thought the given health state would reduce their income and, if so, by how much.

Two further TTO exercises were included, but analysis of these questions is not included in this chapter; the first of these two questions asked respondents to trade-off length of life to avoid an income loss with health remaining constant at perfect health, while the second asked respondents to trade off length of life to achieve an income gain. Responses to these questions were only used, along with the other TTO responses, to inform the exclusion of 'extreme non-traders' (see Background Characteristics and Non-Traders section). Including these questions, each respondent had a total of 14 TTO exercises to complete (four health states were valued for each of the three main TTO exercises, and one income level was valued for each of the two further questions). This may be considered a large amount but it is not uncommon (e.g. both studies informing the Dutch [115] and Japanese [120] EQ-5D value sets asked each respondent to value 17 different states). Given a sample size of 300, we had 300 responses per state for TTO1 and TTO2, and 100 responses per questionnaire version.

The standard TTO question (TTO1) provided a baseline against which the later TTO questions could be compared. Directly following TTO1, respondents were asked a number of follow-up questions: they were asked if they had considered the effect the states would have on their ability to work, on their income, on their friends and relatives, and on their leisure time. The first null hypothesis was that the majority of respondents, when there is no mention of income, will not take income considerations into account; the second null hypothesis was that the valuations of those who do and do not spontaneously include income effects will not differ.

TTO2 was an ex post inclusion/exclusion question. Those respondents who stated they had not spontaneously included income effects in TTO1 were asked to repeat the exercise, this time including these effects ('ex post inclusion'). Those respondents who stated they had spontaneously included income effects in TTO1 were asked to repeat the exercise, this time excluding these effects ('ex post exclusion'). The ex post inclusion approach was used by Sendi and Brouwer [64], while Krol et al. [65,66] used the ex post exclusion approach; therefore, we were able to compare our results with these studies and further test the effect of explicit instructions. The third null hypothesis was that valuations will not change at the aggregate level following explicit instructions.

TTO3 specified the level of income losses that would be associated with the given health state. Meltzer et al. [60] also provided respondents with specific information, but they asked respondents to value blindness and back pain, so our results were not directly comparable with theirs. The fourth null hypothesis was that the valuations of the four health states in TTO3 will not differ depending on the level of income loss they are paired with.

The health states

Four EQ-5D health states were valued as follows: (i) 11112; (ii) 22211; (iii) 11222; and (iv) 22322 (see Appendix 5B for a description of the four states). The health states were presented in this order, from best to worst, to all respondents in the TTO exercises. In the VAS exercise, the ordering was randomized. We chose these particular health states in order to have variation in the severity of the health states, as well as in levels of impairment of the different domains. This may be particularly important for the 'usual activities' dimension (middle dimension) since it is perhaps most closely related to the ability to work.

Analysis

In order to convert data to utility scores, the number of years living in full health that a respondent gave as being equivalent to 10 years living in health state X was divided by ten. Therefore, if 6 years living in full health was deemed equivalent to 10 years living in health state X, then this response was coded as 0.6 (6/10). No protocol was included for states worse than dead, as we felt this would be too complicated for a self-complete questionnaire. A zero discount rate was assumed for this 10-year period, which is common, although results in a downward bias in values. [121] The smallest unit of trade was 1 month.

Some respondents (extreme non-traders) were unwilling to trade any time in any of the 14 TTO exercises. Chi-square tests were used to compare background characteristics between extreme non- traders and the rest of the sample. TTO1 valuations of respondents who did and did not spontaneously include income effects were compared through unpaired t-tests. Valuations before and after instructions were compared through paired t-tests.

Ordinary least squares (OLS) and random effects (RE) generalized least squares (GLS) regressions were used to show the effect of background characteristics on TTO1 and TTO3 valuations. A probit regression was used to show the effect of background characteristics on the propensity to spontaneously include income effects in TTO1. Data analysis was performed using Stata statistical software, Version 9.

5.3 Results

Background characteristics and non-traders

Data were available from 321 members of the Dutch general public who participated in the online survey. Preliminary data examination showed that 25% of respondents were unwilling to trade any time in any of the 14 TTO exercises. Table 5.1 shows the background characteristics, first for the entire sample and then for those who traded in at least one of the TTOs and for those who did not trade any time in any of the 14 TTOs (i.e. extreme non-traders). The full sample had slightly more males than females. Forty-two percent of the sample were not employed and these respondents are perhaps less likely to have considered income effects. More than half of the sample had children, and these respondents are more likely to have considered income effects as more people are dependent upon that income. Just under half of the sample were married and the mean VAS score for own health was 0.76.

Table 5.1 Background characteristics by traders and non-traders [a]

Variable	All	Traders	Extreme non-traders	Chi-square test (P-values), traders vs. non-traders
N	321	241	80	
Sex				
male	51.0	52.0	54.0	0.350
female	49.0	48.0	46.0	
Age (y)				
average (SD)	44.0 (13.1)	43.2 (13.2)	46.6 (12.4)	
18–35	29.0	32.0	21.0	0.148
36–50	32.0	31.0	33.0	
51–65	39.0	37.0	46.0	
Educated beyond min school-leaving age				
yes	67.0	66.0	70.0	0.507
no	33.0	34.0	30.0	
Educated to degree level				
yes	31.0	32.0	29.0	0.592
no	69.0	68.0	71.0	
Employment				
employed	52.5	53.5	50.0	0.874
self-employed	5.5	5.0	7.5	
housewife/house husband	13.0	12.5	15.0	
pensioner	6.5	7.0	5.0	
seeking work	3.0	3.0	2.5	
unable to work	11.5	10.0	16.0	
student	8.0	9.0	4.0	
Net personal monthly income (€)				
<1000	39.0	38.0	41.0	0.873
1000–1499	22.0	21.5	24.0	
1500–1999	18.0	19.0	16.0	
>2000	21.0	21.5	18.0	
Children				
yes	54.0	49.5	67.5	0.005
no	46.0	50.5	32.5	
Marital status				
married	46.5	42.5	59.0	0.118
single/never married	21.0	22.5	16.0	
divorced	10.0	12.0	4.0	
widowed	2.0	2.0	1.0	
living together	17.5	18.0	17.5	
other	3.0	3.0	2.5	
Mean self-reported health on the EQ-5D VAS [b]	0.76	0.75	0.80	0.073
Spontaneously included income in TTO1				
yes	49.0	42.5	70.0	<0.001
no	51.0	57.5	30.0	

[a] All values are presented as % unless otherwise indicated, [b] The sample sizes for this variable were as follows: all, n = 280; traders, n = 213; non-traders, n = 67. min = minimum; TTO1 = first time trade-off question; VAS = visual analogue scale.

Chapter 5

Two variables were highly significantly correlated with whether or not respondents were prepared to trade any time in any of the TTO exercises. Extreme non-traders were more likely to be parents than non-parents. Extreme non-traders were also more likely than traders to spontaneously consider income effects. Forty-nine percent of the whole sample spontaneously considered income effects, compared with 70% of the extreme non-traders alone. Self-reported health on the VAS was weakly correlated with whether or not respondents traded, with non-traders being in better health than traders.

The presence of more parents among the extreme non-traders does suggest that these may be meaningful preferences rather than strategic responses. However, the aim of our study was to compare changes in valuations depending upon income effects, not to generate health state valuations comparable with existing value sets. Responses of non-traders did not help us achieve this aim; instead, they may have diluted the more meaningful responses of traders. We chose to exclude these extreme non-traders from our analysis, which reduced the sample size from 321 to 241. Furthermore, 41 respondents gave negative VAS valuations of own health (13 of whom were extreme non-traders). It seems unlikely that someone in a state of health worse than dead would be able to complete an online questionnaire. Examination of these responses suggested that they were not meaningful and were predominantly caused by very high valuations of dead as a given health state. Comparison with their EQ-5D valuations showed that these respondents were generally in good health. These respondents were excluded from analysis involving VAS of own health (reducing the sample size to 213), but were included in all other analyses.

Results of the chi-square tests showed that the background characteristics did not differ significantly across the three versions of the questionnaire. Only employment was weakly significantly different across the versions, with a smaller proportion of respondents in the 40% income-loss version being in employment than in the other two versions. Further subgroup analysis was performed to study the propensity to spontaneously include income effects by background characteristics. Respondents who had been educated beyond the minimum school-leaving age (significant at 10%) and who had a degree (significant at 5%) were less likely to spontaneously include income effects.

Spontaneous inclusion/exclusion

Table 5.2 shows the results for the standard TTO1 (as used in the MVH study) and for TTO2 (ex post inclusion/exclusion question), first for the main sample (n = 241) and then by response to the follow-up question of whether or not

Table 5.2 Time trade-off (TTO) results from the first and second TTO questions (TTO1 and TTO2) both including and excluding income effects

Question and subgroup	Health state [mean (SD) median]			
	11112	22211	11222	22322
TTO1 (MVH based)				
All (n = 241)	0.92 (0.18) 1.00	0.86 (0.21) 0.97	0.82 (0.22) 0.90	0.68 (0.28) 0.73
(1) Spontaneously included income effects (n = 102)	0.93 (0.16) 1.00	0.85 (0.22) 0.95	0.84 (0.21) 0.90	0.72 (0.26) 0.80
(2) Spontaneously excluded income effects (n = 139)	0.91 (0.19) 1.00	0.86 (0.21) 0.98	0.81 (0.23) 0.90	0.65 (0.29) 0.70
TTO2 (ex post instruction)				
All (n = 241)	0.92 (0.17) 1.00	0.85 (0.21) 0.91	0.81 (0.22) 0.90	0.67 (0.28) 0.70
(3) Explicitly instructed to exclude income effects (n = 102)	0.95 (0.12) 1.00	0.89 (0.17) 0.94	0.85 (0.17) 0.90	0.73 (0.25) 0.80
(4) Explicitly instructed to include income effects (n = 139)	0.90 (0.20) 1.00	0.83 (0.23) 0.90	0.78 (0.24) 0.87	0.63 (0.29) 0.66
T-test P-values				
Including vs. excluding [a] (1 vs. 2)	0.270	0.698	0.289	0.051
Ex post exclusion instruction [b] (1 vs. 3)	0.056	0.029	0.598	0.618
Ex post inclusion instruction [b] (2 vs. 4)	0.242	0.004	0.037	0.340

[a] Unpaired t-tests, [b] Paired t-tests.
MVH = Measurement and Valuation of Health.

Chapter 5

respondents spontaneously included income effects. Two-sided t-tests directly compared the mean results of those who did and did not spontaneously include income effects. The first observation was that respondents consistently valued state 22211 higher than state 11222, which suggests they considered pain and depression to be worse than problems with mobility and self-care. Values for spontaneous inclusion were lower than those for spontaneous exclusion [(1) vs. (2) in table 5.2] for only one of the four states. The t-tests suggested that the differences in valuations were only weakly significant for the most severe state (22322) and, in this case, spontaneous inclusion gave a higher result, contrary to expectations.

Ex post instructions

When those who spontaneously included income effects were asked to exclude these effects [(1) vs. (3) in table 5.2], the valuations of all four health states went up at the aggregate level. These changes were only significant for the first two states, but the statistical significance was weak and the magnitude of the change was small. When those who did not spontaneously include income effects were instructed to include these effects [(2) vs. (4) in table 5.2], the valuations of all four health states went down at the aggregate level. These changes were statistically significant for states 11222 and 22211, but again the magnitude of the change was limited.

Table 5.3 shows the number of respondents who increased and decreased their valuations for each of the four health states in TTO2, following explicit instructions to either include or exclude income effects. Among the group who were in structed to exclude income effects, in three of the four states more respondents increased their valuations than decreased them. Mean increases were larger than mean reductions for all four states, resulting in positive overall mean changes in valuations. More respondents changed their valuations as the health state became more severe (e.g. state 11222 elicited more changes than 22211). We expected the valuations made by the group instructed to include income effects to be revised downwards. More respondents revised their valuations downwards for all four health states, but a relatively large number of respondents revised their valuations upwards. Mean reductions were larger than mean increases for three of the four health states, but mean overall changes were negative for all four states. More respondents changed their valuations as the health state became more severe and, once again, state 11222 elicited more changes than state 22211.

Background characteristics were compared against whether or not respondents increased, decreased or did not change their valuation of state 22322,

Table 5.3 Number of respondents who changed their valuations following explicit instructions to either include or exclude income effects

Variable	Health state			
	11112	22211	11222	22322
Spontaneously included income effects (instructed to exclude in TTO2) [n = 102]				
No. who increased valuation	17	27	28	31
Mean increase	0.172	0.206	0.169	0.163
No. who reduced valuation	12	22	31	29
Mean reduction	–0.061	–0.094	–0.126	–0.147
Mean change [mean (SD) median] [a]	0.076 (0.20) 0.025	0.071 (0.01) 0.221	0.014 (0.21) –0.008	0.013 (0.20) 0.025
Spontaneously excluded income effects (instructed to include in TTO2) [n = 139]				
No. who increased valuation	25	25	30	35
Mean increase	0.118	0.114	0.161	0.199
No. who reduced valuation	29	44	48	55
Mean reduction	–0.160	–0.188	–0.195	–0.169
Mean change [mean (SD) median] [a]	–0.031 (0.25) 0.196	–0.079 (0.10) 0.216	–0.058 (0.24) –0.075	–0.026 (0.26) –0.05

[a] Excluding those who did not change their valuation.
TTO2 = second time trade-off question.

following explicit instructions both to include and to exclude income effects. The only significant (5%) characteristic was age for those instructed to include income effects. Of those who reduced their valuations, 49.1% were aged 18–35 years, compared with 37.1% for those who increased their valuations and 20.4% for those who did not change their valuations. Also, more people in the oldest age category (51–65 years) counter-intuitively increased their valuations following instructions to include income effects.

Regression analysis

Table 5.4 shows the results of multivariate regression analysis. The four OLS columns show results from OLS models in which the dependent variables are the valuations of the four health states through the standard TTO1. The explanatory variables are background characteristics, whether or not respondents spontaneously included income effects, whether or not they thought the given health state would reduce their income, and whether or not they thought about their family and friends when answering the questions. An interaction term indicating whether respondents had children and thought about their family was also included. The results suggest that having children significantly increased valuations for three of the four states; that is, having children resulted in fewer life years sacrificed in order to regain full health. Being married significantly reduced valuations for the most severe health state. Spontaneous inclusion of income led to significantly higher valuations for the most severe state only. Uncertainty about income insurance also led to significantly higher valuations for the most severe state. The only variable that was significant for all four states was whether respondents thought the given health state would reduce their income. Respondents who thought the state would reduce their income gave significantly lower valuations.

The RE GLS column of table 5.4 shows the results of a pooled (long dataset) RE GLS regression that includes valuations of all four health states. There were a number of significant variables in this model, suggesting that (i) older people give slightly higher valuations; (ii) married people give lower valuations; (iii) parents give higher valuations; and (iv) people who think the state will reduce their income give lower valuations. Three dummy variables were included to show the effect of health state severity on valuations. As expected, states 22211, 11222 and 22322 all elicited significantly lower valuations when compared with state 11112. The sizes of the coefficients were consistent with the raw data presented in table 5.2 (i.e. state 11222 elicited lower valuations than 22211, and 22322 elicited the lowest valuations).

Table 5.4 Ordinary least squares (OLS), random effects (RE) generalized least squares (GLS) and probit regressions to show the effect of background characteristics on valuations in the first time trade-off (TTO) question, TTO1, and on the propensity to spontaneously include income effects [a]

Variable	OLS (n = 213)				RE GLS (long dataset) (n = 852) All four TTO valuations (R2 = 0.254)	Probit (n = 213) Included income (pseudo R2 = 0.157)
	11112 (R2 = 0.109)	22211 (R2 = 0.149)	11222 (R2 = 0.137)	22322 (R2 = 0.150)		
Intercept	0.841***	0.694***	0.721***	0.545***	0.773***	
Sex: male = 1, female = 0	–0.011	–0.009	–0.036	–0.033	–0.02	0.052
Age	0.001	0.002	0.002	0.003	0.002*	0.003
Married = 1, other = 0	–0.041	–0.037	–0.037	–0.089**	–0.053**	–0.048
Educated beyond min school-leaving age	0.026	0.019	0.037	0.052	0.033	–0.079
Have a degree	–0.011	0.007	–0.037	–0.058	–0.026	–0.123
Working = 1, not working = 0	0.040*	0.024	0.014	0.033	0.021	–0.037
Have children	0.072*	0.126**	0.045	0.153**	0.101***	–0.071
Thought about family [b]	0.007	0.041	0.004	–0.004	0.011	0.308***
Have children, thought about family [b]	–0.070	–0.074	–0.043	–0.083	–0.068	0.087
VAS own health	–0.019	0.099	0.095	0.006	0.044	–0.196
Income over €999 per month	0.014	0.008	–0.005	–0.030	–0.005	0.117
Included income effects	0.037	–0.012	0.033	0.078**	0.032	
Thought given health state would decrease income	–0.070**	–0.077***	–0.120***	–0.103**	–0.067***	0.116
Income insurance: no/don't know = 0, yes = 1	0.015	–0.019	–0.007	0.001	0.011	0.189**
Income insurance: no/yes = 0, don't know = 1	–0.004	–0.037	0.032	0.098**	0.022	0.019
State 22211 vs. state 11112					–0.043***	
State 11222 vs. state 11112					–0.082***	
State 22322 vs. state 11112					–0.226***	

[a] Values presented are coefficients. All regressions were performed using robust standard errors.
[b] Thought about their family and friends when answering the questions.
min = minimum; VAS = visual analogue scale; * P ≤ 0.1, ** P ≤ 0.05, *** P ≤ 0.01.

Table 5.5 Valuations of the four health states combined with the three levels of income loss (from the third time trade-off [TTO] question, TTO3), and comparisons with the first TTO question, TTO1

Health state	Income loss [mean (SD) median]				T-test P-values					
	20% (n = 78)	40% (n = 80)	60% (n = 83)	TTO1 (n = 241)	20% vs. 40%	20% vs. 60%	40% vs. 60%	TTO1 vs. TTO3, 20%	TTO1 vs. TTO3, 40%	TTO1 vs. TTO3, 60%
11112	0.89 (0.19) 1.00	0.81 (0.29) 0.98	0.78 (0.27) 0.90	0.92 (0.18) 1.00	0.052	0.006	0.529	0.032	0.001	0.000
22211	0.82 (0.21) 0.90	0.78 (0.28) 0.89	0.70 (0.27) 0.70	0.86 (0.21) 0.97	0.283	0.002	0.068	0.263	0.002	0.000
11222	0.77 (0.23) 0.82	0.74 (0.29) 0.83	0.70 (0.29) 0.75	0.82 (0.22) 0.90	0.469	0.084	0.366	0.034	0.000	0.001
22322	0.67 (0.27) 0.70	0.63 (0.30) 0.60	0.61 (0.30) 0.60	0.68 (0.28) 0.73	0.330	0.178	0.722	0.915	0.054	0.065

The final column of table 5.4 shows the results of a probit model in which the dependent variable was whether or not respondents spontaneously included income effects. Those with income insurance were more likely to spontaneously include income effects, as were those who thought about their family. For the four health states 11112, 22211, 11222 and 22322, the percentage of respondents who thought their income would fall was 13%, 42%, 39% and 54%, respectively. It is interesting to note that, although state 11222 was valued lower than state 22211, more respondents thought 22211 would affect their income. For all health states except 11112 (the mildest state), age had a significant negative impact on the likelihood of thinking the states would reduce income. Again, for all but the mildest state, being in employment highly significantly increased the likelihood of thinking a state would reduce income. This is unsurprising, given that the incomes of those not in work would not be affected if ill health hindered their ability to work. Having income insurance significantly reduced the likelihood of thinking two of the four states would reduce income.

An explicit level of income loss

Table 5.5 shows the valuations of the four health states combined with the three different levels of income loss (20%, 40% and 60%) that were given to respondents in TTO3 depending on which version of the questionnaire they received. The valuations of the four health states in all three versions of the questionnaire were in order from best to worst (the same order as in TTO1): 11112, 22211, 11222 and 22322. This holds in all but one case: for 60% income loss, states 11222 and 22211 had the same mean values. A key finding was that all states were valued lower as the amount of income loss increased. These differences were significant once (for state 11112) when 20% loss was compared with 40% loss, once (for state 22211) when 40% loss was compared with 60% loss, and for three out of the four states (for all states except 22322) when 20% loss was compared with 60% loss. Comparison with TTO1 valuations showed that all TTO3 valuations were lower. These differences were significant in all but two cases (20% loss for states 22211 and 22322).

Table 5.6 outlines results, by questionnaire version, from RE GLS regressions showing the effect of background characteristics on TTO3 valuations. For all three versions of the questionnaire, thinking a given health state reduces income led to significantly lower valuations. For 20% and 40% income loss, if a respondent had a partner with an income over 999 euros per month, they were likely to give a higher TTO valuation. For 20% income loss, having spontaneously included income in TTO1 led to higher valuations in TTO3;

Table 5.6 Random effects (RE) generalized least squares (GLS) regressions showing the effect of background characteristics on valuations from the third time trade-off (TTO) question (TTO3) by questionnaire version [a]

Variable	RE GLS (dependent variable =TTO3 valuations)		
	20% income loss (n = 276) R2 = 0.332	40% income loss (n = 276) R2 = 0.198	60% income loss (n = 300) R2 = 0.216
Intercept	0.651***	0.396*	0.670***
Sex: male = 1, female = 0	–0.048	0.070	0.046
Age	0.000	0.003	0.001
Married = 1, other = 0	–0.018	–0.054	0.020
Educated beyond min school-leaving age	0.000	–0.119	–0.030
Have a degree	0.008	0.090	0.009
Working = 1, not working = 0	0.096	0.030	0.091
Have children	0.214**	–0.024	0.119
Thought about family [b]	0.044	–0.003	0.076
Have children, thought about family [b]	–0.151	0.017	–0.178
VAS own health	–0.080	0.167	–0.088
Income over €999 per month	0.039	–0.013	–0.076
Partner's income over €999 per month	0.125*	0.244*	0.050
Included income effects in TTO1	0.093*	–0.078	0.077
Think given state will reduce income	–0.107***	–0.084***	–0.110***
Income insurance: no/don't know = 0, yes = 1	–0.010	0.032	0.055
Income insurance: no/yes = 0, don't know = 1	0.068	0.024	–0.129
State 22211 vs. state 11112	–0.041	–0.024	–0.049**
State 11222 vs. state 11112	–0.089***	–0.066***	–0.048**
State 22322 vs. state 11112	–0.170***	–0.165***	–0.143***

[a] Values presented are coefficients. [b] Thought about their family and friends when answering the questions. min = minimum, TTO1 = first time trade-off question, VAS = visual analogue scale, . P ≤ 0.1, ** p ≤ 0.05, *** P ≤ 0.01.

having children also led to higher valuations for 20% income loss. The health state dummies showed that states 11222 and 22322 elicited significantly lower valuations than state 11112 for all three income-loss levels. Health state 22211 elicited significantly lower valuations than state 11112 only for the largest income-loss level (60%).

5.4 Discussion

The results from this study can help to shed light on a number of questions regarding what people consider in TTO valuations.

First, do people consider income losses when completing a standard MVH study TTO exercise? Our results show that, for the whole sample, 49% of respondents claimed to have spontaneously included income effects. This is lower than one of the two studies using TTO valuation of EQ-5D health states [66], which produced a value of 64%; however, it is higher than the other study [63], which found that only 6% of respondents spontaneously included income effects. It is possible that respondents might have considered these effects for some states but not others. However, we could only ask respondents whether they had taken income effects into account after valuing all four health states, in order to avoid contaminating the exercise.

Second, does spontaneous inclusion of income effects affect valuations? Our findings support the findings of three existing studies valuing EQ-5D states [65-67] that spontaneous inclusion of income effects does not significantly affect health state valuations at the aggregate level. We had expected values for spontaneous inclusion to be lower than those for spontaneous exclusion. However, this was only the case for one of the four states, and in this case the t-test was not significant. Comparisons in table 5.2 show only weakly significant differences in valuations in the case of state 22322, and in this case the value for spontaneous exclusion was lower. Regression analysis in table 5.4 confirmed this finding. This suggests that previous cost-effectiveness studies using either the human capital or the friction cost method to value productivity costs in the numerator of the cost-effectiveness ratio have not double counted these costs. Similarly, from the current NICE perspective, the results suggest that economic evaluations explicitly not including productivity costs have not included these costs implicitly through the health state valuation exercise either.

Third, does explicit instruction on the inclusion of income lead to statistically significant differences in valuations? The results contradict the findings of one study by Krol et al., [66] but support the findings of another study by Krol et al. [65] by finding that explicit instruction led to statistically significant differences in valuations in some cases. Significant differences were found for two states each when comparing spontaneous exclusion with explicit inclusion, and when comparing spontaneous inclusion with explicit exclusion. However, table 5.3 shows that many respondents did not change their valuations, and even among those who did, many changed them in a counter-intuitive fashion. Following explicit instructions to exclude income effects, for one of the four health states, more respondents decreased their valuations than increased them. This suggests that explicit instructions may not always have the 'desired' effect at the individual level.

In the case of explicitly stating a specific level of income loss, table 5.5 shows that this led to significant differences when compared with TTO1 for 10 of 12 cases. A comparison among the different income-loss levels shows there were generally only significant differences in valuations if the differences in income-loss levels were large (e.g. 20% vs. 60%).

Study limitations

Some weaknesses of this study need to be noted. The large number of extreme non-traders is not ideal. For some respondents this may have been a genuine representation of preferences but we suspect that many of these respondents strategically chose not to trade. Respondents were selected from a database of individuals who had signed up to complete online surveys of this nature. Therefore, they may have deduced that the quickest way to complete the exercise was by choosing not to trade. The sooner respondents completed the exercise, the sooner they were awarded a given amount of money to be donated to a charity of their choice and the chance to win a prize themselves. Van Nooten et al. (122) also found numerous respondents opted not to trade any time in TTO exercises in their online questionnaire.

The use of an online self-complete survey may not be appropriate for a large number of different TTOs, as suggested by the number of non-traders in this study. Ideally, this study would be replicated using an interview method of administration (as used to generate commonly used value sets), which would allow continual guidance and explanation and would also enable qualitative feedback from respondents to be gathered, which may enable researchers to further understand respondents' thought processes. An interview-based study would also allow an iterative TTO procedure with a visual aid to be used, which may improve the accuracy of valuations.

Our results show that employed people were more likely to think a given health state would reduce income. Therefore, since only 58% of our respondents were employed, we cannot rule out the possibility that spontaneous inclusion of income effects may have caused significant differences in valuations if our sample had contained a greater number of employed persons. This finding also casts some doubts as to the generalizability of the results of some previous studies using, for instance, student samples (61,63) or samples consisting of employed individuals only. (64) The high level of unemployment may have been caused by the sample selection method. An existing Internet-based panel of respondents was used, and this panel may attract people who have relatively more time to complete surveys than the employed.

This study may be slightly underpowered. Assuming a standard deviation in TTO valuations of 0.16 (the lowest standard deviation generated by TTO1) and an alpha of 0.05, we can detect a difference of 0.1 with power 0.998. However, assuming a standard deviation of 0.29 (the highest standard deviation in table 5.2), we can only detect a difference of 0.1 with power 0.753. This is just below conventional thresholds for power. Future studies need to be appropriately powered as well. This may be difficult if the interview method of administration is used, due to the time-intensive and costly nature of this approach.

This study did not have a protocol for states worse than dead. We felt that since respondents completed the tasks independently and without guidance, it may become too complicated and time-consuming to include a protocol for states worse than dead. Given that the worst health state (22322) has a value on the Dutch tariff of 0.092 [115], we were concerned that a significant proportion of respondents might have valued this state as worse than dead. In fact, in TTO3 with the highest income-loss level of 60% (which should elicit the lowest values), only seven, four, five and seven responses (out of 83) were zero for the four health states (11112, 22211, 11222 and 22322, respectively). However, if this study were to be repeated, it may be worth including a protocol for states worse than dead.

One potential problem was that we paired all health states with all levels of income loss in TTO3, and some respondents might have considered it unrealistic for state 11112 to cause a 60% loss in income. Excluding the extreme non-traders, a total of 30 of 83 respondents did not trade any time in this exercise (compared with 13 of 83 when state 22322 was paired with a 60% income loss).

Results from this study cannot be generalized to other countries. The social security system in the Netherlands may be more generous than in some other countries, which will obviously influence the effect a given health state might have on income. In order to inform the appropriate method of economic evaluation within a specific country, work of this nature needs to be carried out in that jurisdiction.

5.5 Conclusions

It appears that income losses are not consistently included in health state valuations and, when they are included, they have a negligible impact on results. Within the debate on whether to include productivity costs in the numerator or the denominator of a cost-effectiveness analysis, one should consider that even if income losses are included in health state valuation exercises they may not be an accurate proxy for productivity costs. For example, income insurance may reduce the loss to the individual in the event of illness, but it does not reduce the loss to society. Explicit instruction may potentially be able to change valuations, but whether these amended valuations accurately reflect productivity costs is quite dubious in light of theoretical concerns and also given the results presented here. This chapter therefore offers further evidence to challenge the controversial recommendations of the Washington Panel. We argue that, for economic evaluators wishing to adopt the societal perspective in their evaluation, inclusion of productivity costs in the numerator of the cost-effectiveness ratio may represent the most suitable option.

Acknowledgements

The authors would like to thank Lisa Gold, who discussed a version of this chapter at the Health Economists' Study Group meeting in Manchester in January 2009. They would also like to thank Allan Wailoo, who refereed a version of this chapter for inclusion in the Health Economics and Decision Science discussion paper series.

Appendices

5 A The TTO exercises

TTO			Version		
			A	B	C
1	Standard MVH TTO question	*"You can live for 10 years in health state X or a shorter period of time in full health."* *Follow Up: "Did you consider the effect these states might have on your income?"*	4 states **	4 states	4 states
2	Repeat of TTO1 with instruction to include or exclude income effects*	*"You can live for 10 years in health state X or a shorter period of time in full health. Please include/exclude any effect the state might have on your income."*	4 states	4 states	4 states
3	Respondents explicitly told how much income they will lose in the given health state	*"You can live for 10 years in health state X or you can live for a shorter period of time in full health. In state X your ability to work will be impaired and your current income will fall by 20% [or 40% or 60%]."*	4 states, 20% income loss	4 states, 40% income loss	4 states, 60% incomes loss

* Determined by follow-up to TTO 1

** The four EQ-5D states valued in all versions of TTO1, TTO2 and TTO3 were: 11112, 22211, 11222, 22322

5 B Descriptions of the four EQ-5D health states used in the study

11112
I have **no** problems in walking about
I have **no** problems with self-care
I have **no** problems performing my usual activities
I have **no** pain or discomfort
I am **moderately** anxious or depressed

22211
I have **some** problems with walking about
I have **some** problems with self-care
I have **some** problems with performing my usual activities
I have **no** pain or discomfort
I am **not** anxious or depressed

11222
I have **no** problems in walking about
I have **no** problems with self-care
I have **some** problems with performing my usual activities
I have **moderate** pain or discomfort
I am **moderately** anxious or depressed

22322
I have **some** problems with walking about
I have **some** problems with self-care
I am **unable** to perform my usual activities
I have **moderate** pain or discomfort
I am **moderately** anxious or depressed

6.

Do productivity costs matter?
The impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders

based on: Krol M, Papenburg J, Koopmanschap M, Brouwer W.

Do productivity costs matter? The impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders.

Pharmacoeconomics 2011 Jul 1;29(7):601-619.

Summary

Background: When guidelines for health economic evaluations prescribe that a societal perspective should be adopted, productivity costs should be included. However, previous research suggests that, in practice, productivity costs are often neglected. This may considerably bias the results of cost-effectiveness studies, particularly those regarding treatments targeted at diseases with a high incidence rate in the working population, such as depressive disorders. Objectives: This study aimed to, first, investigate whether economic evaluations of treatments for depressive disorders include productivity costs and, if so, how. Second, to investigate how the inclusion or exclusion of productivity costs affects incremental costs. Methods: A systematic literature review was performed. Included articles were reviewed to determine (i) whether productivity costs had been included and (ii) whether the studies adhered to national health economic guidelines about the inclusion or exclusion of these costs. For those studies that did include productivity costs, we calculated what proportion of total costs were productivity costs. Subsequently, the incremental costs, excluding productivity costs, were calculated and compared with the incremental costs presented in the original article, to analyze the impact of productivity costs on final results. Regression analyses were used to investigate the relationship between the level of productivity costs and the type of depressive disorder, the type of treatment and study characteristics such as time horizon used and productivity cost valuation method. Results: A total of 81 unique economic evaluations of treatments for adults with depressive disorders were identified, 24 of which included productivity costs in the numerator and one in the denominator. Approximately 69% of the economic evaluations ignored productivity costs. Three-quarters of the studies complied with national guidelines regarding the inclusion of productivity costs. For the studies that included productivity costs, these costs reflected an average of 60% of total costs per treatment arm. The inclusion or exclusion of productivity costs substantially affected incremental costs in a number of studies. Regression analyses showed that the level of productivity costs was significantly associated with study characteristics such as average age, the methods of data collection regarding work time lost, the values attached to lost work time, the type of depressive disorder, the type of treatment provided and the level of direct costs. Conclusions: Studies that do not include productivity costs may, in many cases, poorly reflect full societal costs (or savings) of an intervention. Furthermore, when comparing total costs reported in studies that include productivity costs, it should be noted that study characteristics such as the methods used to assess productivity costs may affect their level.

6.1 Introduction

Scarce resources in the health care sector have caused economic evaluations to become increasingly important in reimbursement decisions. These evaluations inform decision makers on the amount of health an intervention generates and at what costs. In many countries, including the Netherlands, Belgium and France, it is mandatory (at least in some instances) to present information on the cost-effectiveness of treatments to be considered for reimbursement. (123) Common types of economic evaluations used to help define basic benefit packages are cost-effectiveness analysis (where benefits are expressed in natural or clinically relevant effect measures) and cost-utility analysis (where effects are expressed as QALYs). In both types, the incremental costs and effects or QALYs of an intervention compared with some relevant alternative(s) are calculated.

Although several important handbooks on the methodology of cost-effectiveness studies have been written to give guidance in conducting these studies, (2,7) some methodological issues remain unsolved. One important area of debate is the inclusion of productivity costs. These costs can be seen as "*Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid.*" (8) The inclusion of productivity costs remains controversial, as does the question of how to best include them. (11)

Although little research has been conducted on this topic, and the percentage of studies that actually include productivity costs remains largely unknown, it has been suggested that less than 10% of economic evaluations include them. (10) Given that these costs may sometimes account for a large proportion of total costs (sometimes even more than 50% (11)), their exclusion could significantly misrepresent total societal costs. The extent to which this is the case is unclear, yet misrepresenting societal costs may ultimately lead to suboptimal decisions.

This review investigated the impact of productivity costs on incremental costs, focusing on economic evaluations of treatments for depression. It is expected that productivity costs are particularly relevant in this disease. The aims of this review were to (i) observe how often economic evaluations of treatments for depression actually include productivity costs (investigated via a systematic literature review); (ii) assess the impact of productivity costs on incremental costs[9] (assessed through the selection and investigation of studies

9 The effect of in- or exclusion of productivity costs on incremental costs obviously translates in an effect on incremental cost-effectiveness ratios. The precise influence depends on the incremental effects and the outcome measures used.

that included productivity costs, highlighting the impact on final results); and (iii) to assess the extent to which disease, treatment and study characteristics (e.g. methods of valuing productivity loss) affect the level of per-patient productivity costs.

6.2 Background

The debate on inclusion of productivity costs in economic evaluation revolves around two main questions: (i) whether productivity costs should be included in economic evaluations; and (ii) conditional on an affirmative answer to the first question, how these costs could best be included. There is no current consensus regarding these two questions.

Should productivity costs be included?

Debates regarding the first question have focused on two aspects: (i) the appropriate perspective to take in economic evaluations; and (ii) ethical concerns regarding the inclusion of productivity costs. It is clear that the inclusion of productivity costs in economic evaluations may lead to the favoring of treatments targeted at the (paid and unpaid) working population, potentially at the expense of other groups such as the elderly. [13] On the other hand, excluding these costs implies that actual societal costs are ignored, which may lead to welfare-damaging decisions. [12,13,41,100] Inclusion of productivity costs immediately relates to the former issue, i.e. that of the appropriate perspective. While leading health economic text-books advocate adopting a broad (societal) perspective [2,7] (consistent with the welfare theoretical roots of economic evaluations, e.g. Brouwer and Koopmanschap [104]), in practice, many economic evaluations deliberately take a narrower, most notably a health care, perspective. Indeed, this perspective is prescribed in many national health economic guidelines. Justification for this narrower perspective is commonly that economic evaluations need to aid health care decision makers in spending a fixed health care budget in line with the goal set for these decision makers (which may be something like maximizing or optimizing health). Therefore, costs and savings that fall outside the narrower scope relevant for this decision maker may be deemed irrelevant and ignored. Brouwer et al. [41] suggested an intermediate position, using a two-perspective approach as a standard in health economic guidelines. Differences in position in these matters have led to a large variety in the content of guidelines throughout Europe, decreasing the transferability of research beyond national borders. To illustrate, of the 21 European guidelines gathered on the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) website, [124] only nine (among them, guidelines for the Netherlands, Austria and Finland) stipulate a societal perspective. The Italian guidelines state that economic evaluations should take the societal and the health perspective. Five guidelines (i.e. the UK, Belgium and the Baltic states) prescribe the health care payer perspective and seven guidelines (including Denmark, Hungary and Switzerland) do not clearly specify the perspective to be adopted. Although very little research has been conducted on the extent to which productivity costs are actually included in economic evaluations, Stone et al. [10] found that they were presented in only 8% of the 228 cost-utility analyses they examined. To some extent, this lack of inclusion may also reflect the lack of consensus on how productivity costs should be included in economic evaluations.

How can productivity costs be included?

Three valuation approaches

Consensus is lacking regarding the question of how productivity costs should be included. Three important approaches for valuing productivity costs can be distinguished. The first two, the human capital approach (HCA) and the friction cost approach (FCA), value productivity costs in monetary terms so that they may be included in the numerator of the cost-effectiveness ratio. Simplified, the HCA values 'all working hours lost due to health problems and related treatments times the gross hourly wage', [88] irrespective of the period of absence. The FCA is based on the idea that, in long-term absenteeism, an ill worker can be replaced by a previously unemployed individual, so that the initial production level will be restored after replacement. Therefore, with the FCA, productivity costs are only included for the duration of the 'friction period' (i.e. the time it takes to hire and train a new worker). [47,56,90] As a result, productivity costs calculated according to the HCA are usually higher than those calculated according to the FCA, especially when considering long-term absence. The third approach, the so-called Washington Panel approach (WPA), values productivity costs in terms of quality of life (QALYs), so that they can be included in the denominator of the cost-effectiveness ratio. [2] The Washington Panel assumed that, during the valuation process, respondents in health state valuations will already account for negative income effects due to ill health and therefore productivity costs will be an integral part of QALY weights attached to health states. Consequently, also valuing these costs on the cost side would result in double counting. All three approaches, especially the WPA, have generated criticism in the literature (for a review on the approaches and the theoretical debate see, for example, Tranmer et al. [11]). The theoretical differences of opinion are reflected in national (European) guidelines: 15 of 21 guidelines at least mention the measurement of productivity costs (five prescribe the HCA, two the FCA and eight do not give clear direction on the methodology to be used). [124]

The impact of including productivity costs at the cost or effect side

Over the last decade, the theoretical debate has moved on to a more empirical discussion. One important question resulting from the debate regarding the WPA was whether respondents in health state valuations do in fact (noticeably) include income losses in health state valuations. Several attempts have been made to investigate whether and how respondents take income effects into account. [31,61-66] It seems clear that respondents do not consistently include or exclude these effects. The percentage of respondents including productivity costs ranged from 6% [63] to 64% [66] across studies and, although the empirical evidence is inconclusive, it generally seems that the effects of (alleged) inclusion of income losses in health state valuations do not (significantly) alter health state valuations. This suggests that health state valuations are rather insensitive to capturing income effects and that 'inclusion' of productivity costs at the effect side by simply stating that respondents will have considered income will not significantly affect outcomes. On the other hand, it is clear that including productivity costs on the cost side will have a noticeable impact on total costs and on the cost-effectiveness ratio. However, little is known about the exact impact on incremental costs and subsequently on final outcomes. Obviously, this will depend on many factors, such as treatment, disease and patient population. To our knowledge, only two studies have focused on quantifying the effect of including productivity costs on final outcomes of an economic evaluation. [9,42]

Koopmanschap and Rutten [9] investigated the impact of productivity costs on the cost-effectiveness outcomes of eight health care programs targeted at different diseases. Productivity costs were restricted to paid labor, and costs were calculated using the FCA (using a friction period of approximately 3 months). The total incremental costs of the different programs changed with the inclusion of productivity costs: from a decrease of 18% to an increase of 52%. One program moved from 'positive costs' to 'savings'. Given the diversity in impact of the inclusion of productivity costs, Koopmanschap and Rutten

[9] concluded that, in general, a significant effort should be made to calculate productivity costs when (i) treatments produce health effects in the short run; (ii) there is a strong impact on ability to work in the short run; and (iii) a large proportion of the study population is in paid work when the health effects occur.

Lindholm et al. [42] investigated, by means of a case study of different interventions aimed at preventing cardiovascular diseases (CVD), the impact of taking either the health care or the societal perspective for reimbursement decisions with a given budget constraint. The cost-effectiveness of treatments for several target groups was calculated from each perspective. The outcomes were put in two QALY rank tables and, fictively, the health care budget for CVD prevention was spent, starting with treatments for the groups with the lowest cost-effectiveness ratio. Adopting either perspective, 10% of the target groups would not receive preventive care. When taking the health care perspective, this 10% not receiving CVD prevention were mainly middle-aged men, while, when taking the societal perspective, it was mainly elderly women. This highlights the distributional consequences of adopting different perspectives.

Productivity costs in economic evaluations of treatments for depressive disorders

Even though Koopmanschap and Rutten [9] and Lindholm et al. [42] make clear that the inclusion or exclusion of productivity costs will affect cost-effectiveness outcomes and allocation of resources across groups, it is also clear that the importance of including productivity costs depends on the disease under study. For instance, treatments mainly targeting diseases affecting the elderly are not likely to generate much productivity savings or costs, at least not those related to paid work (on which we focus here), while interventions that have a strong effect on the productivity of the working population may produce productivity costs that reflect a large part of total costs.

The importance of including productivity costs in economic evaluations of treatments for depression was emphasized over 10 years ago, given the high incidence of depression among people of a working age. [125] In total, 5–12% of men and 10–25% of women have a major depressive episode during their lifetime. [126] In 2002, the 12-month prevalence of major depressive disorders was estimated as 6.6% in the US. According to the WHO, depression will be the highest ranking cause of burden of disease in the Western world by the year 2020. According to Bostwick and Pankratz, [127] around 2–9% of patients with depressive disorders eventually commit suicide, compared with <0.5% of those without affective disorders. The incidence of depression is highest in middle-aged individuals, which may indicate that it strongly affects society's productivity, especially in light of the recurrent nature of the disease. (128) Kessler et al. (129) found that employees with depression had 1.5–3.2 more short-term disability days per month than those without. In 2000, the total economic burden of depression in the US was estimated to be $US 51.5 billion, of which 62% was reflected by workplace costs. (130) Such figures highlight the need to further investigate the influence of productivity costs on outcomes of economic evaluations regarding treatments for depressive disorders.

6.3 Methods

Literature review

A systematic review was performed to identify original economic evaluations of treatments for depression that included productivity costs, in order to investigate the impact of such costs on incremental costs. By recalculating incremental costs after excluding productivity costs, the effect of including them in economic evaluations on outcomes is illustrated. The systematic review was performed using the Cochrane Library and PubMed databases. A publication date limit of January 1997 to May 2008 was chosen because standardization of cost-effectiveness studies became increasingly common from the late 1990s. Some publications, such as the influential US guidelines (2) and the textbook by Drummond et al. (7) made important contributions to this development.

The queries used for the database search were 'depression' AND 'cost' OR 'costs' AND 'effectiveness'. To identify relevant economic evaluations, the following inclusion and exclusion criteria were used. Only unique scientific articles in peer-reviewed journals using the English language were considered; articles needed to focus on the estimation of incremental cost-effectiveness of therapeutic interventions for depression; abstracts were excluded and articles had to be available in the Netherlands or in the British Library. Although reviews were excluded, the references used in reviews were hand searched to identify any missing relevant economic evaluations. Only articles with at least a part of the patient population aged between 18 and 65 years were included, since productivity costs related to depressive disorders are most relevant in this age category. Cost items had to be reported separately in order to (re)calculate incremental costs including and excluding productivity costs. All searches were undertaken independently by two reviewers (M. Krol and J. Papenburg).

Analysis

After the identification of relevant publications, the articles were scanned for the inclusion of productivity costs. We examined whether studies followed their national health economic guidelines (according to the study's country of origin) regarding the inclusion of such costs. This was done to investigate whether guidelines can potentially explain their inclusion or exclusion. Next, of the articles that included productivity costs, the percentage of total cost reflected by these costs was calculated. The results are reported separately for the HCA and FCA. The incremental costs excluding productivity costs were then calculated and compared with those presented in the original article, in order to analyze the impact of productivity costs on incremental costs.

Finally, meta-regression analyses were used to investigate whether the level of (monthly) productivity costs could be explained by type of depressive disorder, treatment and study characteristics (e.g. time horizon applied, methodology used to value productivity losses). For this purpose, a database was created of 27 of the 30 economic evaluations including productivity costs in the numerator. Three studies [131-133] were not included in the database due to a lack of information on the absolute amount of productivity costs. The 27 remaining studies accounted for a total of 87 treatment arms useful for the statistical analyses. Data were extracted for these treatment arms on the per-patient average monthly direct costs and productivity costs, methodology regarding measurement of lost work time, the valuation approach (i.e. FCA or HCA) and the values attached to lost work time. The database also included the study time horizon, nature of the depression, type of intervention, mean age of patients under study and percentage of females. For modelling studies, data on mean age and percentage of females were extracted from the original clinical studies where possible; however, this proved impossible for 30 treatment arms. In these cases, missing data were imputed using the respective mean values of the other treatment arms included in the database.

The study characteristics examined were the applied time horizon, total monthly direct costs per patient, sex distribution and average age of the patient population. Dummy variables were created to distinguish between productivity cost approach used (FCA or HCA), method of measurement of lost work time (patient questionnaires, literature estimates, estimates by medical professionals and other/not specified) and values attached to lost work time (national average wages, values based on GDP per capita, sickness insurance fund payments, patients' gross wages, other/not specified). Moreover, dummies were created for the type of depression of the study population (mild, major or severe depression, dysthymic disorder or a mixed population) and the type of treatment (combination) provided (preventive treatment, psychotherapy, drugs, drugs and psychotherapy, drugs administered in a hospital setting). Univariate correlations were first used to explore the relationship between study characteristics and the dependent variable 'average monthly productivity costs per patient'. The dummy variables were included in the univariate models as single dichotomous variables.

A multivariate model was then constructed by entering all variables in the model and stepwise eliminating non-significant variables ($P > 0.5$). Standard econometric tests were applied to the multivariate model to test for the Normality of the residuals (skewness, kurtosis, Kolmogorov-Smirnov). Furthermore, residual plots were used to investigate the appropriateness of the model. The Ramsey RESET test [134] was applied to test the model specification.

The statistical analyses were performed with use of SPSS® Statistical Software Package 17.0 (Chicago, IL, USA).

6.4 Results

Systematic review of inclusion of productivity costs

The literature search identified 732 articles (PubMed: 507, Cochrane Library: 225). As can be seen in figure 6.1, 169 were duplicates, four were published before 1997 and four were not in the English language. Of the remaining 555 articles, 164 were reviews, 178 were not exclusively concerned with depressive disorders, 93 were not classified as economic evaluations, 26 studied adolescents or elderly people, 6 articles were not available in the Netherlands or the British library and 7 were (congress) abstracts. This search finally identified 81 economic evaluations of treatments for adults with depression. Of the 81 articles, 25 (approximately 31%) included productivity costs, of which 24 included these costs in the numerator (i.e. as costs) and one included them in the denominator (i.e. using the WPA, in effects). The references of the 164 reviews were checked in order to detect additional economic evaluations meeting the inclusion criteria and additionally including productivity cost. This procedure resulted in the inclusion of six additional economic evaluations with productivity costs included in the numerator of the cost-effectiveness ratio. As a result, a total of 87 studies were included in our study, of which 30 articles were suitable for the analysis of the impact of productivity costs on incremental costs (i.e. all identified studies except the one using the WPA).

Productivity costs and health economic guidelines

For the 87 economic evaluations, we examined whether the inclusion or exclusion of productivity costs was in line with the relevant national guidelines. As shown in table 6.1, 64 studies (74%) complied with the relevant guidelines. Productivity costs were excluded (or reported separately from the base case) in 46 of these studies, while the remaining 18 studies included productivity costs. Of the 18 studies including productivity costs in line with their national

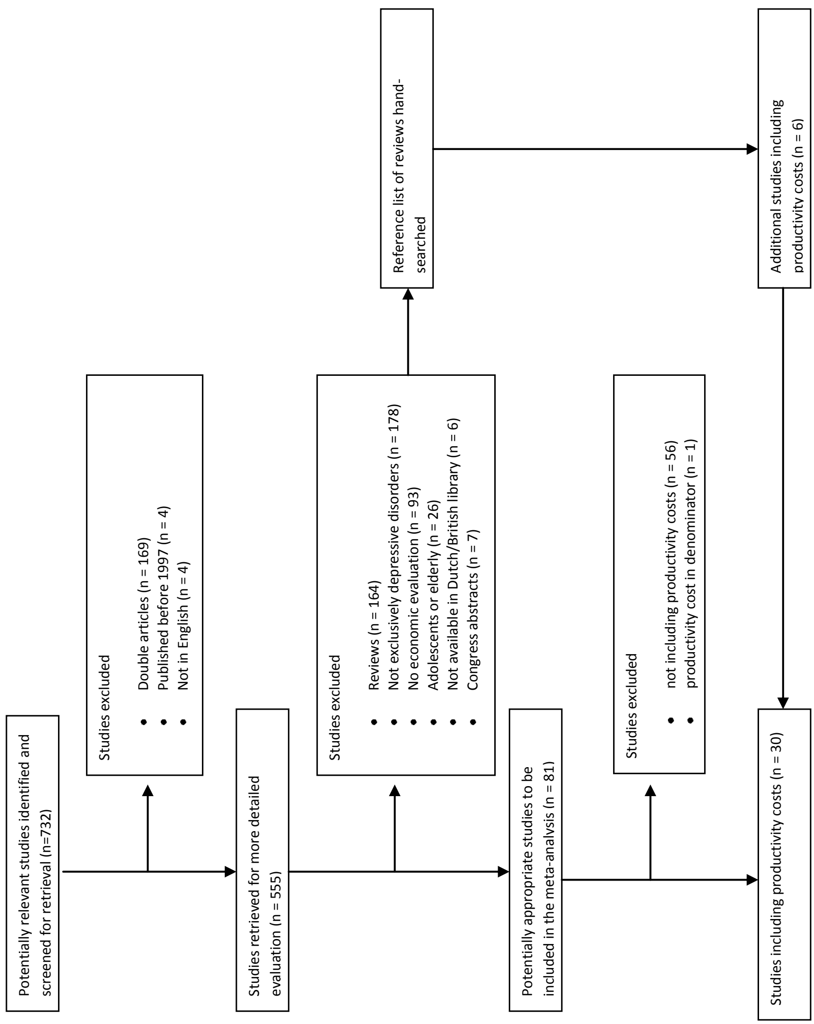

Figure 6.1 Flow diagram of the systematic literature review

guidelines, five used the HCA to calculate productivity costs (despite the national guidelines recommending the FCA). A total of 21 studies deviated from the national guidelines: 9 did not include productivity costs, although national guidelines prescribed a societal perspective. However, 12 evaluations included productivity costs, even though the national guidelines advocated a health care perspective. Two studies were conducted in countries without guidelines, or with guidelines in development.

Table 6.1 Country guidelines on perspective and productivity costs

Country	Guidelines: perspective	Guidelines: productivity costs approach	Studies with PC			Studies without PC
			FCA	HCA	WPA	
Australia	Health care (for base case) and societal	Not specified				2 (135,136)
Austria	Societal	HCA		3 (137-139) a		
Belgium	Health care	HCA	1 (140) a			
Brazil	Societal or health care payer	HCA				1 (141)
Canada	Health care (for reference case)	FCA		1 (142)		1 (143)
Chile	Unknown	-				1 (144)
Czech	Guidelines in development	-				1 (145)
Denmark	Societal	Not specified		1 (146) a		
France	Societal	Preferable FCA (research needed before FCA can be applied in France)		3 (147-149) a		3 (150-152)
Germany	Societal	FCA				1 (153)
Italy	Health care payer and societal perspective	HCA				1 (154)
Norway	Societal and health care payer	FCA or HCA		1 (155) a		
Spain	Societal	Not specified		1 (156)		1 (157)
Sweden	Societal	HCA		2 (158) a, (131)		
The Netherlands	Societal	FCA	5 (159) a, (160-163)	1 (164) a		1 (165)
United Kingdom	Health care payer	Exclude PC		8 (155,166-171) a, (172,173)		16 (174-189)
United States of America	Health care payer (base case)	Not specified		3 (132) a, (133,190)	1 (191)	27 (192) b, (193-218)
Total			*6*	*24*	*1*	*56*

[a] These studies performed economic evaluations from both the health care payer and the societal perspectives, in line with the recommendation of Brouwer et.al. (104)

[b] One USA article (192) included patients of 10 countries in the study-design. Although in some of the corresponding country guidelines a societal perspective is prescribed, in the study a health care perspective was taken (in line with the USA guidelines).

FCA = Friction cost approach HCA = Human capital approach WPA = Washington panel approach PC = productivity costs

Table 6.2 Productivity costs as percentage of total costs

Productivity costs approach	Number of articles	Number of arms	Average percentage PC of TC	Range
Friction costs approach	6	20	56%	19 – 78 %
Human capital approach	24 [a]	73	61%	3 – 92 %
Total	*30*	*93*	*60%*	*3 – 92 %*

[a] In 3 articles [131-133] only incremental costs are presented, which makes it impossible to recalculate percentage PC of TC

Productivity costs as proportion of total costs

In total, the costs and effects of 103 treatment arms were presented within the 30 articles suitable for studying the impact of inclusion of productivity costs (background information of these studies is available in Appendix 6A). In three of these articles, [131-133] in which ten treatment arms were presented, it was not possible to extract the productivity costs per treatment arm due to a limited level of detail for the cost items. As can be seen in table 6.2, productivity costs reflected an average of 60% of total costs per treatment arm. In the 20 treatment arms in which productivity costs were calculated with the FCA, 56% of the total costs per arm consisted of productivity costs; productivity costs calculated using the HCA reflected 61% of the total costs. These figures were somewhat influenced by one study, [142] in which the HCA was used and productivity costs represented just 3% of total costs. How units of lost work time were valued in that study was not specified. Excluding this outlier increased the average percentage of productivity costs in total costs in the HCA group to 65%, and the percentage of the total group to 64%. The range of percentages in the HCA group then changed from 3–92% to 24–92%.

Diversity in productivity costs measurement

The level of productivity costs in individual studies depended not only on the valuation approach used (HCM or FCA), but also on how lost work time was measured and which values were attached per unit of time (e.g. average national income or patients' gross income). The level of detail presented on productivity cost measurement and valuation differed across studies. To summarize: almost all studies (28 of 30 studies) only collected data on absenteeism related to paid work. One study also included productivity losses related to presenteeism, [132] while another study included absenteeism, presenteeism and unpaid labor. [162] Clearly, the focus in productivity cost valuation is still on absence from paid work. Ten studies [133,142,156,161-163,167-169,172] used questionnaires to collect data on absenteeism, ten [131,132,137,140,149,155,158,159,166,167] used literature estimates, three [146-148] used estimates provided by medical professionals, one [161] used cost diaries,

Chapter 6

one [173] used data on doctors' certificates for absenteeism and five [138,139,164,170,171] did not specify the methods used to collect data on lost work time. Regarding the valuation of time lost from work, 13 studies [138,139,158,160-162,164,168-170,173,190] applied average (age/sex dependent) wage and employment rates, five [147-149,159,166] used values based on GDP per capita, one [156] based the values on national minimum wages and four [137,142,146,167] used sick fund/insurance payments made to patients. Only two studies [133,172] specifically mentioned basing the values on patients' wages and five [131,132,140,155,163] did not specify the source of the values used.

The impact of productivity costs on incremental costs

The 30 economic evaluations that included productivity costs used more than ten different effectiveness outcomes (e.g. QALYs based on EQ-5D-scores, the Quality of Life in Depression Scale (QLDS), [219] incidence rates of depressive disorders, time without depression and the Beck depression inventory). [220] However, the majority used the increase in proportion of successfully treated patients as the main outcome measure. The indication of the severity of depression and successful treatment can be based on instruments such as the Montgomery Asberg Depression Rating Scale (MADRS), [221] the Hamilton Rating Scale for Depression (HAM-D) [222] or the Hopkins Symptom Checklist 90 (SCL-90). [223] In these three measures, lower scores indicate better functioning. Obviously, the use of different outcome measures hampers direct comparisons of studies as well as more general observations on the influence of productivity costs on incremental cost-effectiveness ratios (ICERs). Hence, we focus here on the influence on incremental costs.

The 30 economic evaluations (103 treatment arms) presented a total of 61 incremental cost comparisons of treatment arms. In some studies, more than two therapeutic options were compared and, in seven of the studies, ICERs were calculated with multiple effects measures. ICERs were calculated for 74 cases. In five cost-effectiveness comparisons, the incremental effectiveness was zero, implying that an ICER could not be calculated.

The differences in incremental costs including or excluding productivity costs are illustrated in figure 6.2. All costs were converted to euros, year 2007 values. Incremental costs excluding productivity costs are expressed on the vertical axis and those including productivity costs are on the horizontal axis. The five comparisons of the study of Valenstein et al. [132] were excluded from this figure due to insufficient details presented on incremental costs to allow recalculations as required here. As can be seen by the points in the north-west

Table 6.3 Regression analyses

	Monthly productivity costs per patient (N=87)					Monthly productivity costs (n= 78), extreme values excluded				
	Univariate			Multivariate		Univariate			Multivariate	
	Beta	P-value	Adjusted R^2	Beta	P-value	Beta	P-value	Adjusted R^2	Beta	P-value
Age	-0.094	0.388	-0.003	-0.154	*0.018*	-0.076	0.510	-0.007	-0.273	*0.002*
Direct costs p/m	0.728	*0.000*	0.525	0.501	*0.000*	0.670	*0.000*	0.441	0.344	*0.000*
Time horizon	-0.176	0.103	0.020	-0.252	*0.000*	-0.107	*0.352*	-0.002		
FCA (vs. HCA)	-0.261	*0.015*	0.057			-0.290	*0.010*	0.072		
Percentage females	-0.312	*0.003*	0.087			-0.036	0.755	-0.012		
Major depression [a]	-0.157	0.146	0.013	-0.222	*0.000*	-0.026	0.820	-0.012	-0.238	*0.002*
Severe depression [a]	-0.106	0.327	0.000	-0.244	*0.000*	-0.118	0.304	0.001	-0.496	*0.000*
Dysthymic disorder [a]	-0.198	0.067	0.028			-0.291	*0.010*	0.072	-0.540	*0.000*
Mild depression [a]	-0.012	0.992	-0.012			0.078	0.500	-0.007		
Mixed population [a]	0.302	*0.005*	0.080			0.210	0.065	0.031		
Prevention [b]	-0.091	0.408	-0-004			-0.098	0.394	-0.003		
Psychotherapy [b]	-0.169	0.122	0.017			-0.146	0.201	0.009		
Drug treatment [b]	0.193	0.073	0.026			0.084	0.467	-0.006		
Drugs in hospital [b]	-0.212	*0.051*	0.033			-0.294	*0.009*	0.075		
Psych. ther.+drugs [b]	0.118	0.282	0.002			0.348	*0.002*	0.110	0.228	*0.001*
Questionnaires [c]	-0.265	*0.013*	0.059			-0.279	*0.013*	0.066	-0.322	*0.001*
Datacollection NS [c]	0.003	0.979	-0.012	0.162	*0.012*	0.217	0.056	0.035	0.293	*0.000*
Interview doctors [c]	0.676	*0.000*	0.451	0.444	*0.000*		*Not applicable*		*Not applicable*	
Literature [c]	-0.139	0.200	0.008			0.119	0.299	0.001		
Income patients [d]	-0.148	0.171	0.010	-0.171	*0.005*	-0.182	0.111	0.020	-0.243	*0.003*
GDP [d]	0.253	*0.018*	0.053	0.105	*0.050*	-0.230	*0.043*	0.040	-0.248	*0.000*
Insurance payments [d]	-0.034	0.756	-0.011	-0.164	*0.003*	0.025	0.831	-0.013	0.306	*0.000*
Mean nat. income [d]	0.013	0.904	-0.012			0.376	*0.001*	0.130		
Other values or NS [d]	-0.169	0.117	0.017			-0.162	0.157	0.013		
Adjusted R^2				*0.822*					*0.784*	
Skewness				*0.045*					*-0.022*	
Kurtosis				*0.901*					*-0.303*	
Kolmogorov-Smirnov					*0.093*					*0.200*
Ramsey RESET				*F = 23.005*	*0.000*				*F =2.287*	*0.087*

[a] Dummy variables for depression types, [b] Dummy variables for treatment modalities, [c] Dummy variables for data collection methods of lost working time, [d] Dummy variables for the values attached to lost working time

FCA = friction cost approach HCA = human capital approach NS = not specified GDP = gross domestic product

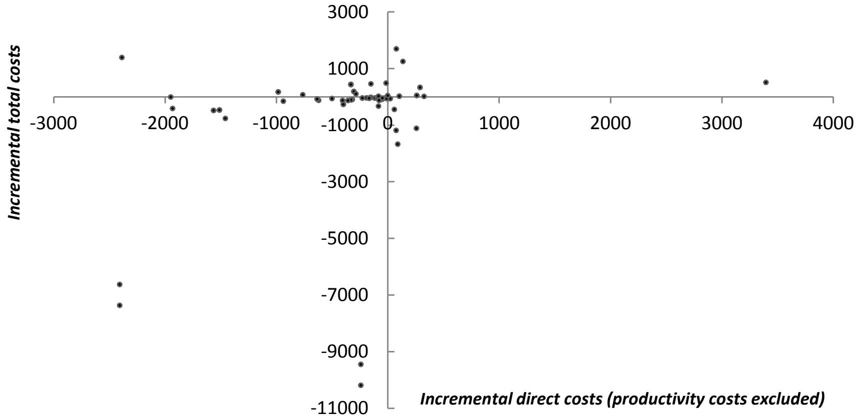

Figure 6.2 Incremental direct costs versus incremental total costs

quadrant, in some cases, including productivity costs caused the incremental costs to change from positive to negative, therefore turning the new treatment into a cost-saving intervention. For other studies (those located in the south-east quadrant) the opposite was true; including productivity costs made an otherwise cost-saving intervention cost. Next to this relative impact, the absolute impact of productivity costs sometimes proved substantial, with occasional differences between incremental costs with and without productivity costs of over 2000 euros (e.g. Kendrick et al. [168] and Antonuccio et al. [190]).

One might expect that 'new interventions' (if indeed more effective than the comparator) would result in reduced productivity costs (i.e. relative savings). However, incremental costs changed in both directions after the inclusion of productivity costs: decreasing in 43 cases, increasing in 16 cases and remaining equal in two cases. It is possible that some new treatments, while perhaps being more effective than the comparator, could also be more intensive or time consuming (e.g. certain types of psychotherapy compared with drug therapy), thus causing the treatment itself to result in increased productivity costs.

The relationship between productivity costs and study characteristics

Univariate and multivariate linear regression models were used to tentatively explore the relationship between per-patient monthly productivity costs and study characteristics in those studies that included productivity costs. Although the residuals of the multivariate model passed the tests for Normality, skewness and kurtosis, the QQ-plot showed that a few residuals had extreme values. Moreover, the multivariate model failed the Ramsey test (see table 6.3).

Given the non-Normal distribution of productivity costs and direct costs, the effects of several transformations (such as log-transformations and quadratic terms) were explored. Furthermore, we tested applying other link functions and the use of interaction terms. None of these measures improved the model. Excluding the nine observations (derived from four studies [146-148,168]) with high leverage and large residuals according to Cook's distance improved the model specification substantially (Ramsey RESET P = 0.087). The extreme values may be partially explained by the use of medical professional estimates of absenteeism, since all eight observations from studies using such estimates were identified as extreme values. [146-148] Only one 'outlier' of nine could not be explained by having used estimates from medical professionals. Nevertheless, the exclusion of extreme values should be handled with caution, especially since cost data are known to be positively skewed. [224] Given the explorative nature of the meta-regression, the results of the multivariate analyses are presented both including and excluding the extreme values (see table 6.3).

Monthly direct costs

Both multivariate models showed similar patterns and explained a high proportion of variance (adjusted R-square 0.822 and 0.784). The models demonstrated a significant positive association between per-patient monthly productivity costs and monthly direct costs ($P < 0.001$). In other words, the higher the direct costs, the higher the productivity costs. This association may be explained by the fact that higher direct costs may be associated with more severely ill patients who are consequently also less productive.

Age

Monthly productivity costs were negatively associated with the average age of patients in the studies. Monthly productivity costs were lower for studies with a higher average patient age. Since our review excluded studies that mainly focused on elderly patients, this association may signal that younger people in these samples may be more frequently involved in paid work and older individuals may more frequently be retired. It would also be possible that younger employees have more health-related absence days than older employees (after controlling for health status) or that younger employees might have more work-related mental health problems.

Time-horizon

Monthly productivity costs were significantly lower for studies with longer time horizons in the full-data model. Since patients generally recover from a depressive episode within a few months, most productivity costs are likely to occur in a short time period. If follow-up continues after recovery, monthly productivity costs will, on average, decrease. However, since time horizon loses significance in the second model, the significant relation seemed to be especially induced by the few extreme values.

Valuation approach

Using the FCA instead of the HCA was significantly associated with lower monthly productivity costs in the univariate models; however, this significant association was not confirmed in the multivariate models, possibly due to correlation with other methodological choices.

Type of depression

Patient populations with a dysthymic disorder seem to be related to lower monthly productivity costs than those with other depressive disorders. This negative relationship might be explained by the fact that a dysthymic disorder is a relatively mild but chronic condition. Unexpectedly, patient populations with major and severe depressive disorders were associated with lower monthly productivity costs than those with other depressive disorders. A possible explanation for this finding might be that a priori unemployment rates in these groups could be relatively high.

Type of treatment

The only significant variable regarding the type of treatment provided was the combination of drugs and psychotherapy, although only in the multivariate model without the extreme values. Psychotherapy combined with drugs led to higher monthly productivity costs, which may be related to the intensity of treatment.

Method of data collection

In the full-data model, lost work time based on estimates from medical doctors was related to higher monthly productivity costs than that based on literature estimates and patient questionnaires. The observations from the three studies using doctors' estimates [146-148] were labeled as extreme values according to Cook's distance. Medical professionals might overestimate patients' lost work time due to depressive disorders and subsequent treatments. Furthermore, not specifying the data collection method for lost work time was associated with significantly higher monthly productivity costs. Obviously, it is difficult to explain this. Excluding extreme values, the use of patient questionnaires to collect absenteeism data was associated with lower productivity costs per month than other methods. The effect of data collection methods on the level of productivity costs raises important questions regarding the accuracy of the individual methods.

Values attached to lost working time

Valuations of lost work time based on GDP per capita was significantly associated with the level of monthly productivity costs in both multivariate models compared with using mean national incomes or unspecified methods. Intriguingly, this association was positive in the full-data model but negative in the model that excluded extreme values. This is explained by the fact that four of the nine excluded extreme values (which were mostly caused by high quantities of lost work time) related to estimates in which GDP was used to attach values to lost work (these four extreme values were derived from two studies [147,148]). In the model without these extremes, using GDP per capita resulted in relatively low monthly productivity costs. Since the calculation of GDP per capita usually does not correct for employment rates, resulting in relatively low estimates of added value per worker (national output is divided by the total population: working and non-working), this result seems plausible. Values based on patients' incomes were related to lower productivity costs, suggesting that patients' incomes in the studies may have been relatively low compared with mean national income. Values based on sick fund/insurance payments to patients led to lower productivity costs in the full-data model and to higher costs in the model without the extreme values. Again, the difference between the models is explained by some of the excluded observations using values based on sick fund payments. It is not clear why this variable is positively associated in the model without the extreme values, since one might expect sick fund payments to result in lower estimates, since these payments might be lower than average national income.

6.5 Discussion

Our study revealed that 69% of the economic evaluations of interventions for depressive disorder do not include productivity costs. While this may be seen as an important omission of real societal costs, it should be noted that almost three-quarters of the studies followed national guidelines regarding the inclusion or exclusion of productivity costs. In cases where studies did not follow national guidelines, it is possible that the decision to include or exclude indirect costs in the analysis could be endogenous, i.e. could be subject to a kind of perspective-selection bias. Endogenous factors around the exclusion of productivity costs could be the time and effort related to measurement of such costs. However, their inclusion or exclusion may possibly also be induced by the (expected) positive or negative effects of productivity cost inclusion on cost-effectiveness outcomes, emphasizing the need for standardization.

It is clear that differences in inclusion of productivity costs between studies are important, given the high average proportion of productivity costs of total costs and given the strong impact of productivity costs on incremental costs. Our results showed that productivity costs, on average, reflect more than half of the total costs for treatments for depressive disorders. Moreover, moving from excluding to including productivity costs in many cases substantially affected the incremental costs.

These findings indicate that the choice of perspective prescribed in reimbursement guidelines may influence cost-effectiveness outcomes (and conceivably subsequent decision making) to a great extent in the area of depressive disorders. Furthermore, our findings illustrate that comparing cost-effectiveness outcomes between studies should involve a thorough examination of the separate cost (and effect) items included in the analyses. Even when comparing cost-effectiveness outcomes between studies that include productivity costs, it remains important to examine other elements of study design. Indeed, our results indicated that the amount of productivity costs (and consequently the impact of productivity costs on incremental cost-effectiveness) is associated with study characteristics such as methods used to value productivity losses. Several guidelines allow investigators the freedom to deviate from the preferred perspective if such a deviation is properly motivated. Although this freedom may be useful given the aim and audience of individual studies, it limits comparability of studies and, as such, may complicate decision making. Moreover, it can lead to a perspective-selection bias.

Most of the studies that included productivity costs used the HCA to calculate them. Note that we assumed that studies that did not clearly specify the method used, applied the HCA, which was in line with the articles' description of productivity cost calculations. One study used the WPA and thus assumed the effects of disease and treatment on productivity to be fully valued at the effect side of the cost-effectiveness ratio. This study did not specify the (assumed) effects of productivity changes on quality of life outcomes, making it impossible to indicate the impact on the ICER. For studies including the HCA or the FCA, it is usually expected that the HCA generates higher productivity costs since, in the FCA, the time in which productivity costs are assumed to occur is limited. However, in this review, the approach used was only significantly associated with the amount of monthly productivity costs in a univariate model. The small differences found between the HCA and the FCA in our findings are likely to be explained by the relatively short time horizon of the studied economic evaluations. Most studies used a time horizon of 6 months to 2 years. In the first few months, the differences between the HCA and the FCA are expected to be negligible, considering a friction period of approximately 5 months. Furthermore, the acute phase of most depressive disorders will last only a few months (maybe not exceeding the friction period). If patients return to work after treatment and fall ill again, a new friction period will start, and productivity costs between the HCA and the FCA will be quite similar. Whereas differences in productivity cost approaches did not seem to affect outcomes much, other aspects regarding the methodology of productivity costs seemed to matter more. Differences in data collection methods and the values attached to work time lost were significantly related to the level of productivity costs.

Although our study was primarily concerned with productivity cost measurement, the regression analyses showed that the level of productivity costs was strongly associated with the level of direct costs. Although not investigated in this study, the diversity in direct cost measurement may be equal to the diversity in productivity cost measurement. It would be worthwhile to investigate such differences in direct cost measurement and the effect of these differences on cost-effectiveness outcomes, especially given the association between direct costs and productivity costs found in this study.

Limitations

The productivity cost analyses in this study were based on a limited number of studies including productivity costs. Moreover, our study solely focused on depressive disorders; therefore, the results cannot be generalized. The impact of productivity costs on total costs could be explored further by using larger and more diverse databases (covering other diseases and treatments). Ideally, this might facilitate a predictive model that may provide a rough indication of productivity costs in cases where these costs are omitted.

6.6 Conclusions

Productivity costs in economic evaluations of treatments for depressive disorders reflect a large part of total costs, which implies that, in studies that do not include productivity costs, a large part of actual societal costs are ignored. This is apparently the case in the majority of economic evaluations in this area. In many cases, the exclusion of productivity costs is explained by not having taken a societal perspective, often in line with national guidelines. The influence of the inclusion or exclusion of productivity costs on incremental costs, and consequently on cost-effectiveness, in the area of depression underlines the importance of the discussion on the appropriate perspective in health economic evaluations.

Acknowledgements

The authors would like to thank the anonymous reviewers for their useful comments on an earlier draft of this chapter. We would also like to thank Maarten van Gils for his useful help on the statistical analyses

Appendix

6A Economic evaluations including productivity costs

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per monthb (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Bosmans et.al. The Netherlands [161]	UCnoAD	12	(i) MADRS (ii) QALYs	Cost diaries. Absenteeism. Mean age sex-specific income. Unpaid work with shadow price method	FCA	354	487	73			
	UCAD					272	420	65			
	(UCAD vs. UCnoAD)[a]								174	-810	-983[d]
Stant et.al. The Netherlands [163]	TAU	18	SCL-90	Questionnaires. Absenteeism, Dutch standard prices not further specified	FCA	77	283	27			
	CST					92	236	39			
	(CST vs. TAU)[a]								-1,111	-853	258
Sørensen et.al. Denmark [146]	Citalopram	6	Success rate (MADRS)	Survey among medical doctors. Absenteeism. Sick fund payments	HCA	853	1,208	71			
	Escitalopram					750	1,084	69			
	(Escitalopram vs. Citalopram)[a]								-122	-743	-620
	Venlafaxine-XR					656	941	70			
	Escitalopram					647	917	71			
	(Escitalopram vs. Venlafaxine-XR)[a]								-92	-145	-53
Bosmans et.al. The Netherlands [160]	UC	6	SCL-90	Questionnaires. Absenteeism. Average age+sex-dependent income	FCA	405	533	76			
	Pharmacist coaching					459	589	78			
	(Pharmacists coaching vs. UC)[a]								16	339	323
Sobocki et.al. Sweden [131]	UC	60	QALYs	Literature. Absenteeism	HCA	NS	NS	-			
	Hypothetical AD therapy					NS	NS	-			
	(Hypothetical AD therapy vs. UC)[a]								-762	-2,220	-1458
Kendrick et.al. United Kingdom [168]	Usual GP care	6	QALYs	Questionnaires. Absenteeism. Average national gender dependent earnings	HCA	1,024	1,109	92			
	Generic nurse care					999	1,161	86			
	(Generic nurse care vs. Usual GP care)[a]								459	308	-151
	PST					1,590	1,761	90			
	(PST vs. Usual GP care)[a]								511	3,907	3,396

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Smit et.al. The Netherlands [(162)]	UC	12	IRDD	Questionnaire: TIC-P. Absenteeism, presenteeism and unpaid work. Average age and sex-dependent earnings. Unpaid work by shadow price	FCA	571	759	75			
	CBT					409	596	69			
	(CBT vs. UC)[a]								-5	-1,954	-1,948
Serrano-Blanco et.al. Spain [(156)]	FLU	0	(i) MADRS (ii) CGI	Questionnaires. Authorized absenteeism. Minimum wage	HCA			63			
	IMI							59			
	(IMI vs. FLU)[a]								-42	-158	-115
	FLU	1				302	402	75			
	IMI					150	232	65			
	(IMI vs. FLU)[a]								-18	-170	-152
	FLU	3				218	292	75			
	IMI					117	182	64			
	(IMI vs. FLU)[a]								189	-329	-518[d]
	FLU	6				186	247	75			
	IMI					80	129	62			
	(IMI vs. FLU)[a]								-79	-713	-634
Van Baardewijk et.al. The Netherlands [(164)]	Venlafaxine-XR	6	(i) Success rate (HAM-D or MADRS) (ii) SFDs	Absenteeism. Average national hourly wages	HCA	1,412	2,100	67			
	Duloxine					1,460	2,204	66			
	(Duloxine vs. Venlafaxine-XR)[a]								334	624	290
Wade et.al. United Kingdom [(170)]	Citalopram	6	Success rate (MADRS)	Absenteeism. Average market wages	HCA	78	200	39			
	Escitalopram					71	186	38			
	Escitalopram vs. Citalopram[a]								-52	-89	-37

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Wade et.al. United Kingdom (171)	Citalopram	6	Success rate (MADRS)	Absenteeism. Average market wages	HCA	566	713	79			
	Escitalopram					498	624	80			
	(Escitalopram vs. Citalopram)[a]								-128	-534	-406
	Venlafaxine-XR					386	498	77			
	Escitalopram					390	491	79			
	(Escitalopram vs. Venlafaxine-XR)[a]								-63	-41	23
Demyttenaere et.al. Belgium (140)	Citalopram	6	Success rate (MADRS)	Literature. Absenteeism	FCA	133	225	59			
	Escitalopram					117	205	57			
	(Escitalopram vs. Citalopram)[a]								-25	-121	-95
	Venlafaxine-XR					104	183	57			
	Escitalopram					102	177	58			
	(Escitalopram vs. Venlafaxine-XR)[a]								-21	-36	-15
Fernandez et.al. United Kingdom (167)	Venlafaxine-XR	2	QALYs	Questionnaires. Absenteeism. Average national wage based daily benefits paid by sickness insurance for period off work.	HCA	376	461	82			
	Escitalopram					346	404	86			
	(Escitalopram vs. Venlafaxine-XR)[a]								-54	-114	-60
François et.al. Norway (155)	Citalopram	6	Success rate (MADRS)	Literature. Absenteeism. Norway statistics	HCA	441	566	78			
	Fluoxetine					443	570	78			
	Venlafaxine-XR					401	530	76			
	Escitalopram					388	497	78			
	(Escitalopram vs. Citalopram)[a]								-94	-412	-318
	(Escitalopram vs. Fluoxetine)[a]								-110	-439	-329
	(Escitalopram vs. Venlafaxine-XR)[a]								-123	-201	-78

Chapter 6

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
McCrone et.al. United Kingdom (173)	TAU	6	(i) QALYs (ii) BDI	Absenteeism with doctors certificate. Age+gender dependent national average daily wages.	HCA	177	251	71			
	BtB					94	157	60			
	(BtB vs. TAU)[a]								-60	-560	-500
	TAU	8				127	211	60			
	BtB					32	125	26			
	(BtB vs. TAU)[a]								75	-688	-763[d]
Hemels et.al. Austria (138)	Citalopram	6	Success rate (MADRS)	Absenteeism. Market wage rates.	HCA	313	390	80			
	Escitalopram					282	352	80			
	(Escitalopram vs. Citalopram)[a]								-37	-227	-190
Hemels et.al. Austria (139)	Citalopram		Success rate (MADRS)	Absenteeism. Market wage rates.	HCA	278	611	45			
	Escitalopram					264	542	49			
	(Escitalopram vs. Citalopram)[a]								-328	-411	-83
Romeo et.al. United Kingdom (169)	Paroxetine	5.5	(i) Success rate(17-HAMD) (ii) Change in QLDS	Questionnaires. Absenteeism. National average wage rates	HCA	213	717	30			
	Mirtazapine					140	596	24			
	(Mirtazapine vs. Paroxetine)[a]								-266	-664	-399
Browne et.al. Canada (142)	IPT	12	MADRS	Questionnaires. Absenteeism. Unemployment insurance payments	HCA	10	236	4			
	SSRI					9	275	3			
	SSRI + IPT					16	143	11			
	(SSRI vs. IPT)[a]								486	472	-14
	(SSRI + IPT vs. IPT)[a]								-1,184	-1,109	75
	(SSRI + IPT vs. SSRI)[a]								-1,671	-1,581	90
	IPT	24				9	189	5			
	SSRI					12	263	5			
	SSRI + IPT					15	247	6			
	(SSRI vs. IPT)[a]								1,695	1,772	77
	(SSRI + IPT vs. IPT)[a]								1,251	1,387	136
	(SSRI + IPT vs. SSRI)[a]								-445	-385	59

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Nuijten et.al. The Netherlands (159)	Discontinuation of TCA for 9 months + termination	9	(i) TWD (ii) QALYs	Literature. Absenteeism. Value of days lost based on GDP per capita.	FCA	55	84	66			
	Discontinuation of TCA for 9 months + prolongation					18	96	19			
	(Discontinuation of TCA for 9 months + prolongation vs. discontinuation of TCA for 9 months + termination)[a]								437	107	-330
	Continuation of TCA for 9 months + termination					55	84	66			
	Continuation of TCA for 9 months + prolongation					18	96	19			
	(Continuation of TCA for 9 months + prolongation vs. Continuation of TCA for 9 months + termination)[a]								439	107	-331
	Discontinuation of TCA for 9 months + termination 21					111	233	48			
	Discontinuation of TCA for 9 months + prolongation					39	139	28			
	(Discontinuation of TCA for 9 months + prolongation vs. discontinuation of TCA for 9 months + termination)[a]								-464	-1,975	-1511
	Continuation of TCA for 9 months + termination					111	233	48			
	Continuation of TCA for 9 months + prolongation					37	135	27			
	(Continuation of TCA for a 9 months + prolongation vs. Continuation of TCA for 9 months + termination)[a]								-480	-2,044	-1564
Valenstein et.al. United States of America (132)	No screening followed by UC	Lifetime	QALYs	Literature. Absenteeism and presenteeism.	HCA	NS	NS	-			
	Annual screening					NS	NS	-			
	(Annual screening vs. no screening followed by UC)[a]								209, 959[c]	245, 987[c]	36,028
	Periodic screening every 3 years					NS	NS	-			
	(Periodic screening every 3 years vs. no screening followed by UC)[a]								89,120[c]	126, 481[c]	37,361
	Periodic screening every 5 years					NS	NS	-			

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
	(Periodic screening every 5 years vs. no screening followed by UC)[a]								55, 629[c]	93, 477[c]	37,848
	One-time screening					NS	NS	-			
	(One-time screening vs. no screening followed by UC)[a]								34, 970[c]	49, 421[c]	14,450
	Opportunistic screening					NS	NS	-			
	(Opportunistic screening vs. no screening followed by UC)[a]								339, 205[c]	327, 184[c]	-12,022
King et.al. United Kingdom (172)	Routine GP care	4	(i) BDI	Questionnaires.	HCA	164	268	61			
	CBT		(ii) QALYs	Absenteeism. Patients		122	214	57			
	NDC			gross income.		190	300	63			
	(CBT vs. GP care)[a]								-49	-216	-167
	(NDC vs. GP care)[a]								23	127	104
	Routine GP care	12				106	174	61			
	CBT					87	151	58			
	NDC					128	199	64			
	(CBT vs. GP care)[a]								-41	-268	-228
	(NDC vs. GP care)[a]								49	310	261
Brown et.al. Austria (137)	Amitriptyline	6.5	Success rate (17-HAMD)	Literature. Absenteeism. Sick fund payments.	HCA	886	1,320	67			
	Mirtazapine					841	1,292	65			
	(Mirtazapine vs. Amitriptyline)[a]								103	-182	-285[d]
	Fluoxetine	6				867.14	1,320	66			
	Mirtazapine					867.14	1,327	65			
	(Mirtazapine vs. Fluoxetine)[a]								38	38	0
Nuijten et.al. France (149)	TCAs	18	TWD	Literature. Absenteeism.	HCA	288	565	51			
	Fluvoxamine			Value of days lost based on		180	435	41			
	(Fluvoxamine vs. TCAs)[a]			GDP per capita.					-410	-2,342	-1,932

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Antonuccio et.al. United States (190)	Fluoxetine	24	Treatment success, NS	Literature. Absenteeism. Average net wage rate.	HCA	912	1,644	55			
	Combination of both protocols					822	1,671	49			
	Individual CBT					812	1,267	64			
	Group CBT					812	1,237	66			
	(Individual CBT vs. Fluoxetine)[a]								-6,626	-9,033	-2,407
	(Group CBT vs. Fluoxetine)[a]								-7,366	-9,772	-2,407
	(Individual CBT vs. Combination both protocols)[a]								-9,449	-9,690	-241
	(Group CBT vs. Combination both protocols)[a]								-10,188	-10,430	-241
Löthgren et.al. Sweden (158)	Citalopram	6	Success rate (MADRS)	Literature. Absenteeism. Average national wage rates	HCA	442	710	62			
	Escitalopram					382	620	61			
	(Escitalopram vs. Citalopram)[a]								-129	-486	-357
	Venlafaxine-XR					329	551	60			
	Escitalopram					321	536	60			
	(Escitalopram vs. Venlafaxine-XR)[a]								-43	-92	-48
Borghi et.al. United Kingdom (166)	Amitriptyline	6.5	Success rate (17-HAMD)	Literature. Absenteeism. Value of days lost based on GDP per capita.	HCA	400	517	77			
	Mirtazapine					398	507	79			
	(Mirtazapine vs. Amitriptyline)[a]								-60	-70	-10
	Fluoxetine	6				388	500	78			
	Mirtazapine					388	507	76			
	(Mirtazapine vs. Fluoxetine)[a]								44	44	0
Brown et.al. Austria (137)	Amitriptyline	6.5	Success rate (17-HAMD)	Retrospective interviews with medical doctors. Absenteeism. Sick fund payments. Value of days lost based on GDP per capita.	HCA	3,036	3,728	81			
	Mirtazapine					2,892	3,560	81			
	(Mirtazapine vs. Amitriptyline)[a]								-153	-1,091	-938

Chapter 6

Study	Comparators (comparisons)	Months follow-up	Effect measure	PC methodology (data collection, data collected, values attached)	PC approach	PC per month[b] (€)	TC per month[b] (€)	% PC of TC	IC -PC[b] (€)	IC +PC[b] (€)	IC change[b] (€)
Brown et.al. Austria (147)	Fluoxetine	6	Success rate (17-HAMD)	Retrospective interviews with medical doctors. Absenteeism. Value of days lost based on GDP per capita.	HCA	3,224	3,960	81			
	Mirtazapine					3,210	3,950	81			
	(Mirtazapine vs. Fluoxetine)[a]								23	-60	-83[d]
Zhang et.al. United States (133)	GP	12	NS	Questionnaires baseline. Absenteeism. Patients' wages.	HCA	NS	NS	-			
	MHS					NS	NS	-			
	(MHS vs. GP)[a]								1,391	-996	-2,387[d]

[a] New intervention versus comparator, [b] in euros 2007 per patient, [c] Costs per QALY gained, [d] Shift in cost-saving/costing, PC = productivity costs, TC = Total costs (direct costs + PC), IC = Incremental cost, FCA = Friction cost approach, HCA = Human capital approach, Vs. = versus, NS = not specified, UC = Usual care, AD = Antidepressant UC, no AD = Usual care without antidepressants, UCAD = Usual care with antidepressants, TAU = Treatment as usual, CST = Cognitive Self-Therapy, XR = extended-release, GP = General Practioner, PST = Nurse problem-solving treatment, FLU = Fluoxetine, IMI = Imipramine, BtB = Beating the Blues, computerized therapy program of a 15 min introductory video followed by 8 50 min sessions of cognitive-behavioral therapy, IPT = Interpersonal Psychotherapy, SSRI = Sertraline, a selective serotonin reuptake inhibiting, TCA = Tricyclic antidepressant, CBT = Cognitive-behavior therapy, NDC = Non-directive counseling, MHS = Mental Health Specialist, MADRS = Montgomery Asberg Depression Rating Scale, QALY = Quality Adjusted Life Year, SCL-90 = Hopkins Symptom Checklist 90, IRDD = Incidence rate of depressive disorder, HAM-D = Hamilton Rating Scale for Depression, SFD = symptom-free days, CGI = Checklist Clinical Global Impression, BDI = Beck Depression Inventory, QLDS = Quality of Life in Depression Scale, TWD = Time without depression

7.

A noticeable difference?
Productivity costs related to paid and unpaid work in economic evaluations on expensive drugs

Based on: Krol M, Papenburg J, Tan S, Brouwer W, Hakkaart L.

A Noticeable Difference? Productivity costs related to paid and unpaid work in economic evaluations on expensive drugs.

Submitted paper.

Summary

In economic evaluations of very expensive intramural drugs it could be hypothesized that productivity costs are not significant in terms of their impact on cost-effectiveness outcomes given high direct costs. If so, productivity costs could be excluded from the analysis. This study aimed to: (i) determine whether economic evaluations on expensive drugs commonly include productivity costs related to paid and unpaid work; (ii) investigate the effect of productivity costs on the cost-effectiveness of expensive drugs, and (iii) examine the relationship between in- or excluding productivity costs and countries' health economic guidelines. To meet these aims, we conducted a systematic literature review to identify economic evaluations of 33 expensive intramural drugs. We then investigated the impact of productivity costs on cost-effectiveness outcomes in studies that included productivity costs. Finally, we analysed whether evaluations included productivity costs and whether in- or exclusion could be explained by the study populations' age, health and the relevant national health economic guidelines. Of the 249 identified economic evaluations of expensive drugs 22 (9%) included productivity costs related to paid work. Only one study included unpaid productivity. Productivity cost inclusion sometimes had a strong impact on incremental cost-effectiveness. Productivity costs were more often included if the economic evaluation originated from countries whose guidelines prescribed a societal perspective. The results of this study indicate that productivity costs in economic evaluations of expensive intramural drugs are ignored in the majority of studies. Given their potential impact, caution is warranted when interpreting the cost-effectiveness of expensive intramural drugs based on evaluations that a priori exclude productivity costs.

7.1 Introduction

The development of new and expensive health care technologies has increased the pressure on national health care budgets as well as hospital budgets, leading to difficult questions of affordability of new medicines at the national and hospital level. The Netherlands is no exception in this context. Traditionally, Dutch hospitals received a fixed budget from which they should finance intramurally dispensed pharmaceuticals. With the increase in new, expensive intramural drugs, this leads to difficulties, since the budgets did not increase at the same pace as the expenses did. While these fixed budgets controlled costs at the national level, they appeared to lead to differences in patients' access to expensive intramural drugs, as some hospital specialists were reluctant to prescribe very expensive drugs. [225]

In 2006 the Netherlands therefore introduced a new policy for reimbursement and financing of expensive intramural drugs, [226] which aimed to avoid unequal access and prevent delays in treatment. The policy implies that hospitals, normally subject to a yearly fixed budget, receive additional financing for drugs placed on the 'expensive drug list'. [226] A drug is eligible for placement on the list and thus additional financing if it is administered within the hospital and its total expenses exceed a certain threshold. Placement on the list is held for four years, within which the drug is subject to a mandatory real-life cost-effectiveness study. The outcome of the study partially determines the drug's eligibility for re-enlistment in the next period.

Given the methodological challenges involving real-life studies, the Dutch Health Insurance Board (CVZ) prescribes a pragmatic approach in conducting the studies, implying that data collection should be focused on the most relevant and influential costs and effects in terms of their impact on final cost-effectiveness outcomes. [227] Such an approach in some ways resembles 'the rule of reason' recommended in the influential health economics textbook of Gold et al. [2] which states that if costs "…*are trivially small or do not differ across regimens, their inclusion will have little effect on the final results of an analysis, and they may therefore be omitted at the analyst's discretion.*"

Inter alia, costs that may be expected to be less relevant in cost-effectiveness studies of expensive intramural drugs are productivity costs. Productivity costs can be defined as *"costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid."* [8] Productivity costs (at least related to paid work) may not be expected to be very influential in expensive intramural drugs studies, given the typically older age of patients and the absolute height of direct costs. Furthermore, most drugs placed on the Dutch expensive drugs list [226] are prescribed to patients in very poor health, such as those suffering from severe rheumatoid arthritis or metastatic cancer, who are unlikely to return to paid work even with treatment. If productivity costs only marginally affect cost-effectiveness outcomes in studies of expensive intramural drugs, the collection of productivity data is in itself not a cost-effective exercise and is unnecessarily burdensome to patients. Currently, little is known regarding productivity costs in studies of intramural expensive drugs and how they affect cost-effectiveness outcomes. It has been shown that productivity costs are often ignored in cost-effectiveness studies [10,43,44], but it is unclear why and what the impact of ignoring these costs on final outcomes has had. However, ignoring costs following the 'rule of reason' must be based on the assertion that their impact on final outcomes is negligible. Obviously, another plausible reason for exclusion of productivity costs is that a country's health economic reimbursement submission guidelines disallow or do not prescribe their inclusion, since a health care perspective is prescribed. Whether this is the dominant reason for current exclusion is unknown, however.

Our research therefore aims to determine whether cost-effectiveness studies of expensive intramural drugs normally include productivity costs related to paid and unpaid work and whether inclusion of these costs in expensive drugs studies is relevant, in other words, whether they have a substantial impact on cost-effectiveness outcomes. A secondary objective is to determine whether including productivity costs is related to factors that may indicate the application of the rule of reason (patients' age and health status) as well as countries' health economic reimbursement submission guidelines.

To meet the study objectives, we conducted an extensive systematic review of economic evaluations of 33 drugs on the Dutch 'expensive drug list'. [226] The effect of including productivity costs on the cost-effectiveness outcomes was assessed by investigating the studies that included productivity costs and evaluating their impact on final results.

7.2 Background

Productivity costs can be influential. For instance, in economic evaluations of treatments for depression such costs, on average, reflect more than half of total costs, often strongly influencing incremental costs and, in turn, cost-effectiveness [(44)]. The inclusion of productivity costs thus can influence the allocation of scarce health care resources across diseases and patients when determined – at least to some extent – by incremental cost-effectiveness.

Although productivity costs and savings can be substantial, previous studies suggest that, depending on the types of interventions and economic evaluations studied, not more than 8 to 31 per cent of economic evaluations include productivity costs related to paid work. [(10,43,44)] To our knowledge the extent to which economic evaluations include production loss related to unpaid labour has only been investigated once, [(43)] suggesting that unpaid labour is seldom included in economic evaluations.

If productivity costs (related to both paid and unpaid labour) are indeed often ignored, it is important to understand why. Several factors may contribute to neglecting productivity costs in economic evaluations, such as the principles on which the economic evaluation is based or pragmatic considerations throughout the execution of the economic evaluation. Economic evaluations adopting a societal perspective aim to include all relevant effects and costs, regardless of who bears the costs and who receives the benefits. Influential health economics text books such as the book of Drummond et al. [(7)] and Gold et al. [(2)] promote a societal perspective consistent with the theoretical welfare foundations of economic evaluations. [(104)] Only by including all relevant costs and effects is it possible to make optimal (welfare improving) decisions. Moreover, if the underlying objective of economic evaluations is to inform decision makers about costs and consequences of alternative actions, taking a broad perspective seems reasonable. [(41)] Although adopting a societal perspective is often advocated, [(2,7,34,35,40)] it is certainly not an undisputed choice. When economic evaluations are seen as a tool to aid decision makers in spending a fixed health care budget in line with a goal (such as maximising or optimising health), some have argued for evaluations with a narrower health care perspective that includes costs falling on the health care budget only (e.g. [(33)]). Productivity costs should be excluded in such cases since their effects are external to the health care budget and thus irrelevant to the allocation problem addressed.

The lack of consensus regarding the appropriate perspective seems also to be reflected in national health economic guidelines, which stipulate how to conduct economic evaluations for reimbursement of health interventions. National guidelines are likely (and intended) to influence how economic evaluations are conducted in practice. Approximately half of the national guidelines on the ISPOR 'pharmacoeconomic guidelines around the world' website [124] prescribe taking a health care perspective (at least for the base case scenario) and the other half a societal perspective or a health care *and* societal perspective.

Prescribing a health care perspective and excluding productivity costs may also be related to important equity concerns [12,13] because their inclusion may lead to favouring reimbursement of health interventions targeted at the working population. If such interventions produce substantial societal savings by improving productivity levels, including productivity costs may result in more favourable cost-effectiveness outcomes than when similar interventions are used in less productive populations (e.g. very young or elderly). Equity concerns potentially favouring the employed are explicitly mentioned in some guidelines (e.g., Australia [228] and New Zealand [229]) as an argument for adopting a health care rather than societal perspective.

We could question whether ignoring the existence of costs and savings outside the health care sector is the proper solution to such ethical concerns. The approach denies decision makers the opportunity to make well-informed decisions and balance potential savings with the equity implications of their decisions, which might be why most countries (except the United Kingdom [72] and New Zealand [229]) prescribing a health care perspective allow presenting additional costs-effectiveness scenarios that include broader societal costs.

Pragmatic factors that contribute to ignoring productivity costs in economic evaluations could be a lack of time, data, or research experience. [230] It could also be related to the lack of scientific consensus regarding appropriate methods to measure and value such costs. Numerous instruments, for example, can measure productivity costs (mainly related to paid work) but which instrument provides the most valid estimate is currently unknown. Estimates that vary substantially [82,83] by using different instruments can result in a lack of confidence in the trustworthiness of productivity cost estimates.

Next to measurement difficulties, the valuation of productivity costs (related to paid work) has been fiercely debated. [22,23,47,58,59,231] The suitability of three valuation approaches has dominated the debates: the human capital approach, [88] the friction cost approach, [47] and the Washington panel approach. [2] With the human capital approach a monetary value is placed on 'all working hours lost due to health problems and related treatments' [88] regardless of the period of absence. The friction cost approach limits productivity cost calculations to the duration of time it takes to hire and train a replacement. The idea behind it is that a sick worker being replaced by, for example, a previously unemployed worker restores societal production levels. The duration of the friction period depends on countries' unemployment rates and vacancy durations. In contrast to the first two, the Washington Panel approach values productivity costs in terms of quality of life effects related to income changes. Strong criticism of the Washington panel approach [8,22,22,31] initiated a number of empirical studies to establish whether quality of life instruments could capture the effects of health changes on income. [61-67,232] Outcomes indicated that quality of life instruments are rather insensitive to such effects, invalidating the approach. The Washington panel approach thus receives little theoretical and practical support, but lack of consensus on whether to apply the friction cost or the human capital approach translates to the use of both in practice. [43,44,233] How to measure and value productivity costs related to unpaid work has received little attention in the scientific literature.

Following 'the rule of reason' [2], productivity costs need only be included in economic evaluations when deemed relevant. If productivity is not affected by some treatment, including productivity related costs would be superfluous, such as with some treatments aimed at very mild conditions or treatments aimed at very severe conditions in which patients are fully impaired without expectation of returning to paid or unpaid work. In most cases, however, whether health interventions will affect patients' productivity is difficult to predict.

According to the rule of reason, productivity cost inclusion would also be irrelevant if productivity is unchanged relative to the comparator (i.e. if productivity changes are similar in both the intervention and comparator groups), since productivity costs will not change incremental cost-effectiveness ratios. For most interventions determining up front whether productivity costs in both study arms are equal is difficult and excluding productivity costs is thus debatable.

In practice, productivity costs may also be regularly omitted based on patients' age. For instance, if most of the patients receiving an intervention are above retirement age productivity losses related to paid work may be negligible. Note, however, that this is not the case for losses related to unpaid work. Treatments targeted at the very young patients could affect future productivity, rendering productivity costs an important factor. Whether such effects are considered important to economic evaluations depends on the valuation method chosen: related costs or savings should be included if using the human capital approach [88] but not the friction cost approach [47].

A final rationale for ignoring productivity costs in line with the rule of reason might be the expectation that these costs are not very influential if direct costs are relatively high. This might be the case in economic evaluations of (very) expensive drugs, especially when these drugs are administered in a (costly) inpatient setting. If, as a rule, productivity costs have little effect on cost-effectiveness outcomes in economic evaluations of expensive intramural drugs, they might be excluded a priori, limiting burden to patients and saving time and other resources. If the costs are influential (as they are in some areas [9,42,44]), ignoring productivity costs upfront could lead to sub-optimal decision making.

7.3 Methods

Review of economic evaluations of intramural expensive drugs.

We performed an extensive systematic review of all 33 drugs on the Dutch 'expensive drug list' in June 2009 (table 7.1) to identify any economic evaluations. We used the Cochrane Library and PubMed databases with a publication date limit of January 1998 to June 2009.

Queries used for the database search were the drugs' names and "cost" or "costs." Inclusion criteria were (i) unique scientific articles in peer-reviewed journals in English and (ii) titles with terms or phrases such as cost(s), budget, economic, financial, price, money, dollar, economy, expenditure, pay, expense, fund, resource, reimbursement, consumption, or expensive. After excluding any reviews, we read abstracts of the remaining articles to determine if they were indeed economic evaluations. Finally, the full texts of the remaining articles were examined (with the exception of those unavailable in the Netherlands or British Library). Title and abstract searches were independently undertaken by two researchers. Full text examinations were carried out by two people in close collaboration.

Table 7.1 Pharmaceuticals included in the review (Dutch expensive intramural drug list June 2009) [226]

Docetaxel, Irinotecan, Gemcitabine, Oxaliplatin, Paclitaxel, Rituximab, Infliximab, Immunoglobulin, Trastuzumab, Botulin toxin, Verteporfin, Doxorubicin liposomal, Vinorelbine, Bevacizumab, Pemetrexed, Bortezomib, Omalizumab, Ibritumomab, Pegaptanib, Alemtuzumab, Palifermin, Drotrecogin-alfa, Natalizumab, Cetuximab, Ranibizumab, Abatacept, Voriconazole, methyl aminolevulinate, Panitumumab, Anidulafungin, Caspofungin, Temsirolimus, Temoporfin

Inclusion of productivity costs and the impact on costs and cost-effectiveness.

After identifying economic evaluations of expensive drugs, we investigated whether they included productivity costs related to paid or unpaid work. Next, (where possible) we extracted the average or median age and the severity of illness of the patients in the study populations. To explore whether age may explain the inclusion of productivity costs related to paid work, we assumed that at least a considerable part of a study population would be in working age if the average age of the study population was between 18 and 70. Subsequently, it was investigated whether inclusion or exclusion of productivity costs could be explained by the (estimated) health related ability to work, or the likelihood that at least a part of the patient populations in individual studies would lose or regain ability to work (in comparison to the control group). These estimations were based on a medical doctor's expert opinion regarding the severity of disease of the patient populations (e.g., in case of metastatic cancer it was assumed that (most) patients would not be able to perform paid work regardless of treatment).We also examined whether studies aligned with their national health economic guidelines regarding productivity costs. National and regional health economic guidelines were retrieved through the ISPOR 'pharmacoeconomic guidelines around the world' website. [124] Where possible the original guideline documents were studied. If the guidelines were not available in English we followed the ISPOR 'Key Features' pages. If guidelines were non-existent or not listed, they were labelled 'unknown' and productivity cost exclusion could not be related to a recommended perspective.

For the studies that included productivity costs, the percentage of total cost accounted for by productivity costs was calculated and reported. We then excluded productivity costs from the study's incremental cost-effectiveness ratio(s) to analyse the impact of productivity costs on cost-effectiveness outcomes. We described whether inclusion of productivity costs led to change in incremental cost-effectiveness and the magnitude of the change. For cost-minimization studies (where incremental effects are insignificant) we examined only changes in incremental costs. All prices were adjusted to 2009 euros using the European Union Harmonized Indices of Consumer Prices published by Eurostat [234].

7.4 Results

Review

The database search resulted in 2,157 articles in Pubmed and 422 in the Cochrane Library, 834 of which were doubles (figure 7.1). The number of doubles was quite high because many economic evaluations included more than one expensive drug and, as a consequence, appeared in several of the database searches of the individual drugs. Ten articles were not in English; 1,120 did not include an economic term or phrase in the title; 15 were not available in the Netherlands or British library; 52 were not scientific research articles; 89 were congress abstracts; three did not evaluate one of the 33 drugs; and 95 could not be qualified as economic evaluations. This resulted in 249 studies that qualified as economic evaluations of drugs from the Dutch expensive intramural drug list. (226) Of these, 22 (about 9%) included productivity costs related to paid work and only one included productivity costs related to unpaid work. (235) (See appendix 7A for details of the studies.) Of the 22 economic evaluations including productivity costs, three were identified as cost minimization analyses (CMA); (236-238) the remaining 19 were cost-utility analyses (CUA), where effects are expressed in quality adjusted life years (QALYs).

Productivity costs methodology

Three (239-241) of the 22 studies including productivity costs provided no details on how productivity costs were measured and valued. Ten (235-238,242-247) used average (age- and gender-dependent) wage and employment rates to value lost working time. One study (248) assumed that productivity costs would be either one or three times the direct costs and another (249) used employers' annual labour costs. The remaining ten studies (239-241,250-256) did not specify the values used.

Two economic evaluations (236,250) explicitly assumed that productivity would be relatively unaffected in all study arms; i.e., productivity costs were assumed not to vary. Only seven studies (238,244-246,250-252,254) described actual data collection on productivity among the patient population, but none of them specified the productivity costs measurement instrument used. Three studies that applied the friction cost approach (235,236,255) also applied the human capital approach. Only four of the remaining studies (245,246,253,256) explicitly mentioned applying the human capital approach. Most other studies appeared to have used the human capital approach (in some form) without explication. The one study that included productivity costs related to unpaid work based the cost estimates on changes in household work and volunteer work. (235) The time spent on unpaid

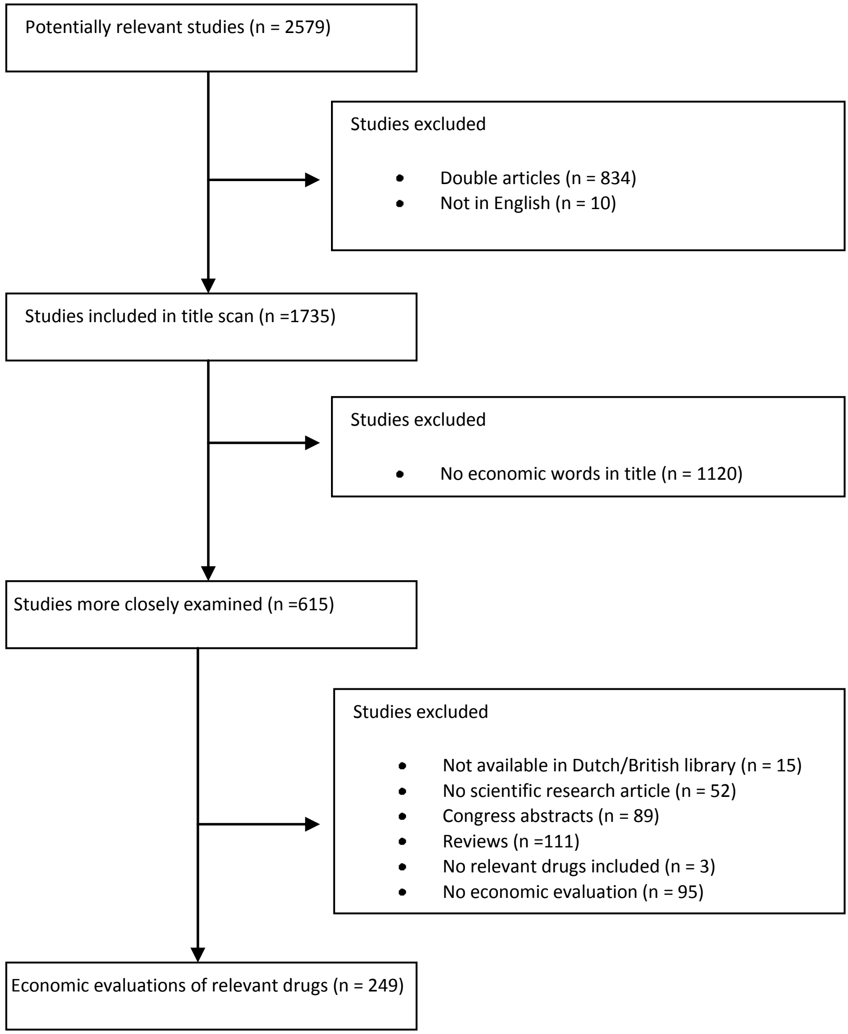

Figure 7.1 Flow diagram of the systematic literature review

work was valued at the same rate as informal care (correcting for the costs already included for household help and informal care).

Productivity costs inclusion: the proportion of total costs

We could not determine the proportion of total costs accounted for by productivity costs for 11 of the 22 articles due to a limited level of detail in the presented cost items. Four of these 11 (240-242,244) did not provide specifications regarding the amount of productivity costs and seven only provided information on incremental productivity costs. (245,246,249-253)

The remaining 11 articles provided information on the absolute amount of productivity costs and accounted for 37 study arms; i.e., 37 estimates of productivity costs as a proportion of total costs related to an intervention or comparator were available. Twenty-four of these estimates valued productivity costs according to the human capital approach; 13 applied both the human capital and the friction costs approach. As expected, studies applying the human capital approach generated higher productivity costs, since in the friction cost approach the duration of inclusion of productivity loss is shorter. For studies applying the human capital approach, productivity costs on average comprised 45% (range: -106% to 83%) of total costs per treatment arm (table 7.2). Note that we assumed that studies not clearly specifying method applied the human capital approach, which was indeed in line with their descriptions of productivity cost calculations. Productivity costs for the 13 study arms for which both the methods were applied were on average 24% (range: 4% to 38%) of the total with the friction cost approach and 44% (range: -106% to 80%) with the human capital approach. For the study on unpaid labour, (235) costs were on average 0.7% of the total, ranging from -25% (i.e., led to savings) to 15%, depending on treatment arm.

The impact of productivity costs on cost-effectiveness outcomes

To study the impact of productivity costs on cost-effectiveness outcomes the incremental cost-effectiveness ratios (ICERs) including productivity costs were recalculated after excluding these costs in the analyses. Because only one study considered unpaid labour, we limited our calculations to paid labour only. Recalculation of cost-effectiveness outcomes was only possible for the 15 CUAs and three CMAs (representing a total of 36 ICERs and eight incremental cost calculations) specifying the amount of productivity costs, or the effect of productivity costs on the ICER. The change in incremental cost-effectiveness after excluding productivity costs is shown in figure 7.2. For all ICERs two bars are included. The first reflects the incremental cost-effectiveness in 2009 euros per QALY *excluding*

Table 7.2 Productivity costs inclusion

	Economic evaluations	% PC of TC (range)
Including PC (paid)	22 of which 19 HCA and 3 FCA + HCA	45 (-106 - 83) HCA 24 (4 - 38) FCA
Including PC (unpaid)	1	0.7 (-18 - 15)
Excluding PC	227	
Total	*249*	

PC = productivity costs, FCA = friction cost approach, HCA = human capital approach

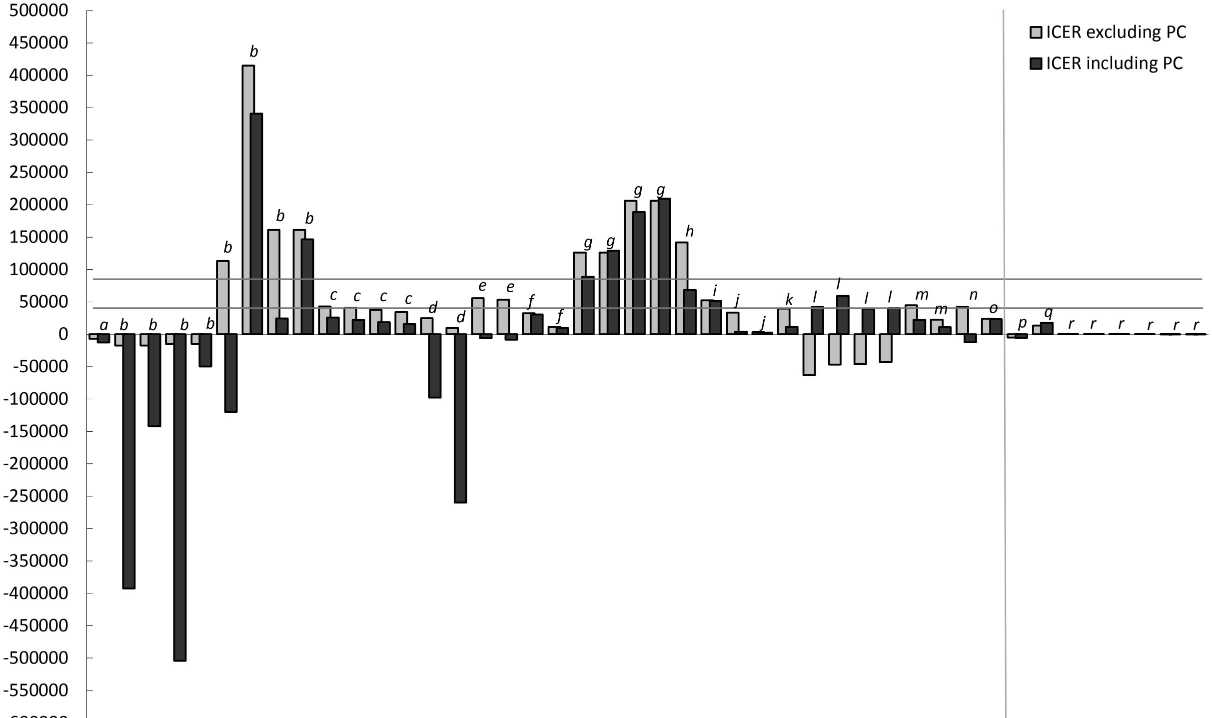

Figure 7.2 ICERS with and without productivity costs

ICERS from: a) Lindgren et al. (239) b) van den Hout et al. (235) c) Davies et al. (243) d) Kobelt, Sobocki et al. (251) e) Kobelt et al. (252) f) Kobelt et al. (253) g) Boonen et al. (255) h) Kobelt, Andlin-Sobocki et al. (245) i) Kobelt, Eberhardt et al. (246) j) Kobelt et al. (254) k) Wong et al. (248) l) Lidgren et al. (256) m) Norum et al. (249) n) Kobelt et al. (250) o) Manns et al. (247) p) Maniadakis et al. (238) q) Walsh et al. (237) r) Norum & Holtmon (236) productivity costs; the second reflects the incremental cost-effectiveness *including* productivity costs. The incremental costs of the three CMAs are presented on the right-hand side of the vertical line in figure 7.2.

We drew two fictive ICER thresholds in the figure for illustrative purposes: one at approximately €40,000 per incremental QALY in the region of the upper limit of the NICE threshold (257) and one at €80,000, a suggested threshold for diseases with a very high burden of disease by the Dutch Council for Public Health and Health Care. (258) These threshold lines illustrate how productivity cost inclusion (exclusion) potentially affects decision making for treatments where the ICER with productivity cost falls below (above) the threshold and the ICER without productivity costs above (below) the threshold. Note that we did not include uncertainty around the ICERS and the thresholds.

A comparison of ICERs with the in- and exclusion of productivity costs shows that ICERs increase due to inclusion of productivity costs in six out of 36 cases. In four of these cases the new treatment changed from cost-saving to cost-spending. Including productivity costs led to a decrease of the IC(ER) in 30 cases. In six of these 30 cases the decrease caused the incremental costs to change from positive to negative, therefore turning the new treatment into a cost-saving intervention.

Taking into account the fictive threshold of €40,000 per QALY, eight ICERs exceed the threshold when excluding productivity costs, while none exceed the threshold when including productivity costs. The other way around, three ICERs lie below the €40,000 threshold without productivity costs and exceed this threshold after including productivity costs. In other words, in 11 of the 36 ICERs (31%) including or excluding productivity costs would alter decision making (if exclusively based on an ICER decision rule), based on a fixed €40,000 threshold. If we raise the threshold to €80,000, three ICERs exceed the threshold excluding productivity costs and none exceed the threshold including productivity costs.

Table 7.3 Patients' age and health related work-ability

	Studies including PC (n = 22)	Studies excluding PC (n = 227)
Productive age (18-65)		
Yes (mean age 18-70)	19 (86 %)	163 (72 %)
No (mean age > 70)	-	21 (9 %)
Unknown	3 (14 %)	42 (19 %)
Work ability based on severity of illness a		
Likely to be able to work	21 (95 %)	146 (64 %)
Doubtful	1 (5 %)	8 (4 %)
Unlikely to be able to work	-	73 (32 %)

[a] These estimations were based on a medical doctor's expert opinion regarding the severity of disease of the patient populations.
PC = productivity costs

Table 7.4 Productivity cost inclusion and national health economic guidelines

Perspectives in HE guidelines	Economic evaluations	Studies including PC	% inclusion
Health care for base case: PC not allowed in any scenario	56	4	7
Health care for base case: PC allowed in additional scenarios	74	4	5
Societal	41	7	17
Societal *and* health care	33	4	12
Societal *or* health care	32	2	6
Unknown	13	1	8
Total	*249*	*22*	*9*

HE = health economic, PC = productivity costs

With the €80,000 threshold, none of the ICERs is above the threshold including productivity costs and below the threshold after excluding these costs. With an €80,000 euro threshold decision making could alter in three of 36 cases (8%).

Inclusion of productivity costs inclusion and health economic guidelines

Inclusion (or exclusion) of productivity costs of all 249 economic evaluations was compared with the study-populations median or average age, the estimated work ability and the health economic guidelines of the respective countries. As can be seen in table 7.3, most study-populations had an average age below 70 in both the studies including and excluding productivity costs. Based on doctor's opinion, the study populations of approximately one-third of the studies excluding productivity costs were expected not to be able to perform paid work regardless of treatment.

Based on national health economic guidelines, for 56 of the 249 economic evaluations productivity cost inclusion was not allowed but four of the 56 included productivity costs anyway. In all four studies the ICERs decreased due to including productivity costs. Fourteen economic evaluations originated from countries for which no guidelines were available, only one of which included productivity costs. [238] In 74 economic evaluations inclusion of productivity costs was mandatory according to the relevant health economic guidelines but only 11 studies followed the rule. Despite this low level of inclusion, table 7.4 shows that productivity costs were (as expected) more often included in evaluations originating from countries with guidelines prescribing inclusion. Only six of the 106 studies from countries in which productivity cost inclusion is allowed but not required included the costs.

7. 5 Discussion

The first aim of this study was to determine whether economic evaluations of expensive intramural drugs usually included productivity costs related to paid and unpaid work, and whether such costs had a substantial impact on cost-effectiveness outcomes. Only one of the 249 identified economic evaluations of expensive intramural drugs included productivity costs related to both paid and unpaid work and 21 (8%) included productivity costs related to paid work. Such results indicate that productivity costs related to unpaid work rarely seem to play a role in cost-effectiveness calculations of expensive intramural drugs and productivity costs related to paid work are ignored in the vast majority of studies. That is not to say differences in outcomes are unimportant; including productivity costs can clearly affect cost(-effectiveness) outcomes. Despite the high direct costs related to expensive drug treatment, productivity costs reflect a relatively high proportion of total costs and can strongly affect incremental cost-effectiveness ratios. With a fixed €40,000 threshold, for example, decisions regarding reimbursement of expensive drugs could alter in almost one-third of the cases by including or excluding productivity costs in the cost-effectiveness analyses. Given these potential strong effects on final outcomes, productivity costs cannot be simply excluded in economic evaluations of expensive intramural drugs based on 'the rule of reason' introduced by Gold et al. [(2)] Consequently, for the recommended pragmatic approach of real-life cost-effectiveness studies of expensive intramural drugs in the Netherlands the results imply that productivity costs should be included when possible and are ethically acceptable. If direct measurement of productivity costs is not possible because of severity of illness it might be possible to predict productivity losses based on quality of life data. [(55)]

A secondary objective of our study was to determine the extent to which productivity costs in- or exclusion is related to patients' age, severity of illness and countries' health economic reimbursement submission guidelines. If assuming that with an average study-population age below 70 a considerable part of the patients is in working age, age does not seem to explain the exclusion of productivity costs in economic evaluations of expensive intramural drugs. Health status may be of influence, however. In approximately one-third of the studies, the severity of illness of the patients could have been a reason to exclude productivity costs (related to paid work) upfront. Moreover, our results indicate that health economic guidelines influence productivity cost inclusion or exclusion, but most guidelines leave room for judging *when* to include productivity costs and *how* to do so. Moreover, in most of the studies we examined, clear information on how productivity costs were derived was lacking. It has been suggested that the decision to include productivity costs may be driven by strategic considerations regarding the expected influence on final outcomes, resulting in a selection-perspective bias. [44] Although too few studies in our review included productivity costs to be able to confirm this suggestion, the ICERs of every economic evaluation that included productivity costs against the relevant health economic guidelines decreased. The existence of a selection-perspective bias emphasizes the importance of standardizing economic evaluations. Given the large impact of productivity costs, transparency in measurement and valuation methods is paramount. Studies' comparability, completeness and transferability would be served by consistent and preferably uniform inclusion of productivity costs, perhaps presented as a separate item. Inclusion also raises decision makers' awareness of societal costs (or savings). If these costs are for any reason not included it is important to justify their exclusion.

A limitation of our review is that by necessity we assessed the impact of productivity costs on cost-effectiveness outcomes based on the studies that actually included them. The amount of productivity costs in these studies may poorly reflect productivity costs in studies excluding these costs, especially if based on strategic considerations. Moreover, we were unable to determine how inclusion or exclusion of productivity costs related to unpaid work affects cost-effectiveness outcomes, since only one economic evaluation in our review considered unpaid work.

7.6 Conclusions

Productivity costs lead to noticeable differences in cost-effectiveness outcomes of economic evaluations of treatments with expensive intramural drugs. Despite the high direct costs related to the drugs, productivity costs reflect a non-negligible part of total costs when included and, therefore, a priori exclusion of productivity costs is not easy to defend when adopting a societal perspective. Ignoring productivity costs in economic evaluations of expensive intramural drugs without clear motive could imply ignoring important societal costs. That notwithstanding, productivity costs related to paid work are omitted in the majority of cases and productivity costs related to unpaid work are seldom included. The neglect of productivity costs is to some extent explained by the relevant national health economic guidelines prescribing a health care perspective, but the rationale for the majority of studies is unclear. We would argue that if productivity costs are ignored, motivation should be clear. Moreover, excluding productivity costs simply to comply with national guidelines does not render them less relevant to society or welfare improvement. An intermediate solution for economic evaluations in countries prescribing a health care perspective could be to measure and value productivity costs whenever relevant and report them separately (as was done by some studies in this review). A two-perspective approach, in which ICERs are presented from both a societal and health care perspective, may be advisable. [41]

Acknowledgements

This study was funded by the Netherlands organisation for health research and development (ZonMW Grant number 152002018). The authors have no conflicts of interest to declare.

Appendix

7 A Studies including productivity costs (all costs in 2009 euros)

Study	country	Type	PC inclusion	Perspective country	disease	age	Comparators	% PC/TC HCA	% PC/TC FCA	ICER	ICER-PC (or IC-PC)	ICER+PC (or IC+PC)	ICER change
Maniadakis et al. 2009 (238)	Greece	CMA	Days lost due treatment and AE, average daily earnings	Unknown	high risk colorectal cancer	66	XELOX	0.3		FOLFOX-XELOX	-5138	-5210	-70
							FOLFOX 6	0.6					
Bristow et al. 2007 (242)	USA	CUA	Losses based on work-type adjusted hourly wages for full treatment period	Health care	stage III ovarian cancer	NS	IV/IV			IP/IV- IV/IV	NS	30308	
							IP/IV						
Lindgren et al. 2009 (239) (239)	Sweden	CUA	NS	Societal	RA	56	RTX	45.0		RTX-TNF-inhibitor	-6875	-12500	-5625
							TNF-inhibitor	45.0					
van den Hout et al. 2009 (235)	NL	CUA	Average working hours minus worked hours, Age + sex-dependent hourly costs, FCA+HCA	Societal	recent-onset RA	54	Group 1 sequential monotherapy	45.0	38.41	2-1 HCA	-17422	-392772	-375350
										2-1 FCA	-17422	-142225	-124803
							Group 2 combination therapy	22.2	31.57	3-1 HCA	-14967	-504189	-489222
										3-1 FCA	-14967	-49845	-34878
							Group 3 combi+prednisone	-26.1	36.25	4-1 HCA	113156	-119830	-232986
										4-1 FCA	414904	340931	-73973
							Group 4 combi+IFX	-106.4	23.60	4-3 HCA	161202	24305	-136897
										4-3 FCA	161202	146541	-14661
Davies et al. 2009 (243)	USA	CUA	PC based on proportion average earnings lost associated with lower HAQ score	Health care	early RA	NS	DMARD	69.6		IFX+MTX-DMARD	43078	25971	-17107
							IFX+MTX	58.4		ETN-DMARD	40653	22201	-18453
							ETN	57.4		ADA+MTX-DMARD	37895	18785	-19110
							ADA+MTX	55.4		ADA+MTX/ETN-DMARD	34342	15804	-18538
							ADA+MTX/ETN	47.2					

Chapter 7

Study	country	Type	PC inclusion	Perspective country	disease	age	Comparators	% PC/TC HCA	% PC/TC FCA	ICER	ICER-PC (or IC-PC)	ICER+PC (or IC+PC)	ICER change
Kobelt, Sobocki et al. 2008 [251]	Spain	CUA	Survey on work capacity, not further specified	Health care + societal	ankylosing spondylitis	57	IFX in double blind trial			IFX-no IFX	24955	-97618	-122574
							IFX in Spanisch open trial			IFX-no IFX	9823	-259993	-269817
Kobelt et al. 2007 [252]	UK	CUA	Survey on work capacity, not further specified	Health care	ankylosing spondylitis	50	IFX in Braun trial			IFX-no IFX	55625	-6184	-61809
							IFX in ASSERT trial			IFX-no IFX	53528	-8141	-61669
Walsh et al. 2007 [237]	Ireland	CMA	Time spent on treatment, average industrial wage	Health care	RA	55	IFX	6.5		IFX-ADA	13383	17962	4579
							ADA	8.9					
Spalding and Hay 2006 [244]	USA	CUA	Average employment rates compared with RA employment rage, age-adjusted income, productivity data of patients + or – ETN	Health care	RA	55-60	MTX			ADA-MTX	NS	53107	
							ADA			ETN-MTX		74763	
							ETN			ADA+MTX-MTX		1620556	
							ADA + MTX			IFX+MTX-MTX		341055	
							IFX + MTX						
Kobelt et al. 2006 [253]	Canada	CUA	Lost work capacity. Incl. early retirement, HCA	Health care	ankylosing spondylitis	NS	No IFX			IFX-no IFX	32562	30626	-1936
							IFX						
Boonen et al. 2006 [255]	NL	CUA	Sick leave in friction period, FCA + HCA	Societal	active ankylosing spondylitis	42	TAU	72.2	8.64	ETN-TAU	126100	88730	-37370
							ETN	45.7	4.95	FCA	126100	129222	3122
							IFX			IFX-TAU	206220	188982	-17238
								44.2	3.99	FCA	206220	209492	3272
Kobelt, Andlin-Sobocki et al. 2004 [245]	UK	CUA	Days absent, early retirement, sex-dependent hourly wages, HCA	Health care	ankylosins spondylitis	57	PI			IFX-PI	141762	68585	-73177
							IFX						

Study	country	Type	PC inclusion	Perspective country	disease	age	Comparators	% PC/TC HCA	% PC/TC FCA	ICER	ICER-PC (or IC-PC)	ICER+PC (or IC+PC)	ICER change
Kobelt, Eberhardt et al. 2004 (246)	Sweden	CUA	Days absent, average annual gross income, HCA	Societal	RA	57	ETN			IFX-ETN	52355	51216	-1139
							IFX						
Kobelt et al. 2003 (254)	UK+ Sweden	CUA	Days absent + early retirement	UK: health care, Sweden: health care +societal	RA	54	Sweden 1 yr IFX + MTX	77.2		IFX+MTX-MTX	33665	4098	-29567
							Sweden 1 yr MTX	82.9					
							UK 1 yr IFX + MTX	53.1		IFX+MTX-MTX	3254	2735	-519
							UK 1 yr MTX	65.6					
Wong et al. 2002 (248)	USA	CUA	Indirect cost estimates first year trial based, for further years PC were estimated to be one or three times DC	Health care	RA	53	MTX	62.7		IFX+MTX-MTX	39658	11457	-28202
							IFX + MTX	60.1					
Lidgren et al. 2008 (256)	Sweden	CUA	HCA, not further specified	Societal	early breast cancer	55	TAU	45.3		Test 1-TAU	-63214	42342	105556
							IHC test 1	57.5		Test 2-TAU	-46804	59233	106037
							IHC test 2	59.7		Test 3-TAU	-46270	39837	86107
							IHC test 3	58.2		Fish test-TAU	-42740	40387	83127
							Fish test	58.6					
Norum et al. 2007 (249)	Norway	CUA	Average female in workforce, Employers annual labor costs during 17 visits	Health care + societal	breast cancer adjuvant	50	NoTRAS 10			TRAS-NoTRAS 10	44907	22081	-22826
							TRAS 10						
							NoTRAS 20			TRAS-NoTRAS 20	22453	11040	-11413
							TRAS 20						

Chapter 7

Study	country	Type	PC inclusion	Perspective country	disease	age	Comparators	% PC/TC HCA	% PC/TC FCA	ICER	ICER-PC (or IC-PC)	ICER+PC (or IC+PC)	ICER change
Norum & Holtmon 2005 [259]	Norway	CMA	Estimated employment rate patients estimated by MDs and compared with average age-sex employment and income, FCA+HCA	Health care + societal	early breast cancer	50	FEC DI100 FEC DI90 FEC DI 80 CMF DI100, DI90, DI80	79.3 80.2 79.4	27.67 28.81 30.05	FEC100-CMF100 FEC90-CMF90 FEC80-CMF80	328 320 -60	328 320 -60	0 0 0
Dewilde et al. 2006 [240]	Sweden	CUA	NS	Societal	severe peristent IgE-mediated (allergic) asthma	43	TAU OM			OM-TAU	NS	63568	
Manns et al. 2002 [247]	Canada	CUA	Estimated employment rates of ICU patients, average gross annual salary	Health care	severe sepsis	61	TAU Activated protein C	1.6 1.6		Act. P.C-TAU	24134	23267	866
Kobelt et al. 2008 [250]	Sweden	CUA	Short-time absence and early retirement	Societal	MS	29	TAU NAT			NAT-TAU	42397	-12481	-54879
Gani et al. 2008 [241]	UK	CUA	NS	Health care	highly active RRMS	36	NAT Inferon β Glatiramer acetate BSC			NAT-Inferon β NAT-GA NAT-BSC	NS	3722 3361 13645	

UK = United Kingdom, USA = United states of America, NL = the Netherlands, NS = not specified, FCA = friction cost approach, HCA = human capital approach, RA = Rheumatoid arthritis, MS = multiple sclerose. RRMS = Relapsing-Remitting multiple sclerose, HAQ = health assessment questionnaire, TAU = treatment as usual, IFX = infliximab, FEC = fluorouracil, epirubicin, cyclophosphamide, CMF = cyclophosphamide, methotrexate, fluorouracil, ADA = adalimumab, MTX = methotrexate, RTX = rituximab, ETN = etanercept, DMARD = disease-modifying antirheumatic drugs, IHC = immunohistochemical, XELOX = capecitabine/oxaliplatin, FOLFOX = 5-fluorouracil/leucovorin/oxaliplatin, IV/IV = outpatient intravenous paclitaxel and carboplatin, IV/IP = inpatient intravenous paclitaxel and intraperitoneal cisplatin plus outpatient intraperitoneal paclitaxel, TRAS = trastuzumab, OM = omalizumab, NAT = natalizumab, Act.P.C = activated protein C, PI = placebo, DI = dose intensity, GA = glatiramer acetate, BSC = best supportive care, MDs = medical doctors, DC =direct costs

8.

Productivity cost calculations in health economic evaluations:

Correcting for compensation mechanisms and multiplier effects

Based on: Krol M, Brouwer WB, Severens JL, Kaper J, Evers SM.

Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects.

Social Science & Medicine. 2012 Dec;75(11):1981-1988.

Summary

Productivity costs related to paid work are commonly calculated in economic evaluations of health technologies by multiplying the relevant number of work days lost with a wage rate estimate. It has been argued that actual productivity costs may either be lower or higher than current estimates due to compensation mechanisms and/or multiplier effects (related to team dependency and problems with finding good substitutes in cases of absenteeism). Empirical evidence on such mechanisms and their impact on productivity costs is scarce, however. This study aims to increase knowledge on how diminished productivity is compensated within firms. Moreover, it aims to explore how compensation and multiplier effects potentially affect productivity cost estimates. Absenteeism and compensation mechanisms were measured in a randomized trial among Dutch citizens examining the cost-effectiveness of reimbursement for smoking cessation treatment. Multiplier effects were extracted from published literature. Productivity costs were calculated applying the friction cost approach. Regular estimates were subsequently adjusted for (i) compensation during regular working hours, (ii) job dependent multipliers and (iii) both compensation and multiplier effects. A total of 187 respondents included in the trial were useful for inclusion in this study, based on being in paid employment, having experienced absenteeism in the preceding six months and completing the questionnaire on absenteeism and compensation mechanisms. Over half of these respondents stated that their absenteeism was compensated during normal working hours by themselves or colleagues. Only counting productivity costs not compensated in regular working hours reduced the traditional estimate by 57%. Correcting for multiplier effects increased regular estimates by a quarter. Combining both impacts decreased traditional estimates by 29%. To conclude, large amounts of lost production are compensated in normal hours. Productivity costs estimates are strongly influenced by adjustment for compensation mechanisms and multiplier effects. The validity of such adjustments needs further examination, however.

8.1 Introduction

Economic evaluations are increasingly used to inform decision-makers on cost-effectiveness of health interventions. Such economic evaluations can be performed from different perspectives. Most national health economic guidelines prescribe either a health care perspective or a societal perspective [233]. In economic evaluations performed from a health care perspective only costs falling on the health care budget are included. In analyses from the societal perspective all relevant costs and effects are included regardless of who bears the costs and who receives the benefits [7]. Although no consensus seems to exist on which perspective to use for decisions regarding reimbursement of health interventions, there are strong arguments to perform economic evaluations from a societal perspective [2,7,34,40,41]. The most compelling argument is that costs falling outside the health care budget are equally real as costs that fall inside this budget and thus affect overall welfare of a society. Subsequently, ignoring these costs could lead to suboptimal decisions for the society as a whole, especially if these costs strongly affect cost-effectiveness outcomes, which is often the case for productivity costs [9,44]. Productivity costs can be seen as *"Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid"* [8]. In this chapter we focus on productivity costs in relation to paid work.

It is important to note that the inclusion of productivity costs in economic evaluations is not without dispute. This relates to two broad issues: (i) which perspective to adopt in economic evaluations and (ii) the equity implications of including productivity costs. Regarding the first issue, it is clear that not all (national) guidelines for economic evaluations recommend adopting a societal perspective. Many prescribe a narrower perspective, i.e. a health care perspective. Then, only costs falling on the health care budget, and thus are the responsibility of the health care policy maker economic evaluations seek to advice, are deemed relevant [33,41]. Second, inclusion of productivity costs (like many other cost categories, including medical costs) can have equity implications. Their inclusion may sometimes lead to favoring interventions targeted at the working age population [12,13], although it may also work in the opposite direction [44]. The equity implications may be particularly pronounced when only costs of paid labor are included while costs of unpaid labor are ignored. To provide policy makers with information on the distributional consequences of including productivity costs, it has been recommended that final cost-effectiveness ratios may be best presented both with and without productivity costs [41,43]. This chapter will not further address the issue of *whether* to include productivity costs. It rather focuses on the question *how* to do so. This is by no means a trivial question.

Productivity costs related to paid work can reflect over half of total costs [11,44] and because of their substantial potential impact on cost-effectiveness outcomes estimates should be accurate. Most commonly, productivity costs related to paid work are estimated through measuring absenteeism and/or diminished productivity while at work. Productivity loss is estimated by calculating the number of days individuals are absent from work due to illness and treatment and/or the amount of diminished productivity at work. These estimates are then multiplied with a relevant value of production per time unit (e.g. hour), commonly some wage rate estimate. Such estimations implicitly assume a direct linear relation between diminished input (through absenteeism and diminished productivity while being at work) and diminished output (lost production). However, a linear translation from diminished productive input to diminished economic output may not be accurate.

It has been suggested that diminished productivity may partly be compensated during normal working hours by the ill worker itself or its colleagues. Koopmanschap & Rutten [76] describe that sometimes absenteeism may not change production levels and lead to extra costs. "*This situation may occur if work can be made up for by the sick employee on his return to work or if an internal labor reserve exists, allowing work to be taken over by colleagues without extra costs.*" However, they note that "*the existence of permanent internal labor reserves raises labor costs, which may have medium term macroeconomic implications*". The existence of such compensation mechanisms was explored further by Severens et al. [17] and Jacob-Tacken et al. [18]. It seemed that missed work was indeed partly compensated during normal working hours. Moreover, compensation mechanisms seemed related to occupational and job characteristics. Severens et al. [17] and Jacob-Tacken et al. [18] argued that, given these compensation mechanisms, traditional productivity cost calculations might overestimate true productivity losses.

Others [14], on the contrary, suggested that common productivity estimates are, in fact, underestimations of 'true' productivity costs. They concluded that "*when there is a team production and substantial team-specific human capital, the value of lost output to the firm from an absence will exceed the wage per day of the absent worker.*" In other words, a person's absenteeism or diminished productivity at work can negatively affect the output of coworkers in cases of team dependencies. Moreover, substitution of ill workers can be difficult and imperfect. To incorporate effects of substitution difficulties and effects on team output it is suggested to multiply the individual's wage with a job-dependent multiplier [15,16]. The multiplier is determined by considering the level of team dependency in realizing the teams' output, the time sensitivity of the output and the ease of substitution of the ill employee.

Traditional productivity cost estimates can be adjusted for such 'multiplier effects' and 'compensation during normal working hours', if this is deemed desirable. To date, however, only a limited amount of empirical research is available regarding the existence of multiplier effects or compensation mechanisms [15-18], and there is no empirical study investigating both concepts integrally. Recently, a new productivity cost questionnaire was developed which is the first to simultaneously include questions on compensation mechanisms and the effects on coworkers [85]. The validity of this questionnaire needs to be further investigated, however.

In this chapter we present an empirical study on compensation mechanisms of reduced productivity among a large group of smokers in a smoking-cessation program. Moreover, based on this information and estimated effects of ill-health on coworkers, we calculated productivity costs for several scenarios. To our knowledge, this is the first study to simultaneously explore multiplier effects and compensation effects. The aim of this study is twofold. First, it aims to increase knowledge on how diminished productivity is actually compensated within firms and to further investigate the association between compensation mechanisms and occupational and job characteristics. Second, the study aims to explore how compensation and multiplier effects potentially affect productivity costs. In doing so, this study seeks to contribute to the discussion on how to increase accuracy of productivity cost estimates. This chapter starts with presenting the methods of our empirical study on compensation mechanisms and describing the productivity costs scenarios. Subsequently, the results are presented and then discussed. Finally, some suggestions are made regarding future research and the estimation of productivity costs.

8.2 Methods

Study population

In order to meet our first aim to increase knowledge on compensation of diminished productivity within firms, absenteeism and compensation mechanisms were measured in 2002 in a large randomized trial conducted among smokers living in the North of the Netherlands examining the cost-effectiveness of including smoking cessation treatments in the basic benefit package. Participants in the smoking cessation study had to be daily smokers, aged over 18 years. A detailed description of the randomized trial was reported elsewhere [(260)]. For this study, we only used the data of respondents who had experienced absenteeism.

Questionnaire

Information on absenteeism and compensation mechanisms was collected cross-sectional among the participants in paid employment at six months follow up through a mail-based questionnaire. Besides the smoking cessation questions, questions were included regarding age, gender, educational level and labor participation. Next, respondents' occupational and job characteristics were evaluated as an indicator for the work situation. Respondents were requested to provide information on job characteristics. These questions were based on the Karasek Job Demand-Control model [(261)] and involved questions on the autonomy to plan own work, the possibility to postpone work, the requirement of special knowledge/skills, the frequency of complex problems and the number of colleagues in similar professions. Subsequently, respondents were asked to state their occupational sector (e.g. industry, health care, education etc.). Next, respondents were asked to state their contracted number of working hours, the number of hours actively worked and over how many days per week these hours were divided. Respondents were also asked whether they were working in shifts or irregular hours and whether they were working in an executive function. Subsequently, respondents were questioned on absenteeism and compensation mechanisms. These questions were obtained from the PROductivity and DISease Questionnaire (PRODISQ) [(262)]. Respondents were asked whether they had experienced absenteeism from work during the past six months. If so, they were asked about the cause of absenteeism, the incidence of periods of absenteeism and the duration of absenteeism. Next, the respondents were asked whether their missed work time was compensated and in what manner. Choice options to this question were: colleagues took over work in normal hours, colleagues took over in extra time, hired workers took over, the respondent took over work later in normal hours, the respondent took over in extra hours, the work was not compensated, and compensation is unknown. Subjects not responding to the mailed questionnaire were contacted by phone. Subsequently, a limited telephone interview was undertaken including questions on absenteeism, compensation mechanisms and baseline characteristics. The questions regarding occupational and job characteristics were not included in the telephone interview. The randomized smoking cessation trial was approved by the Medical Ethics Committee of the Trimbos Institute in Utrecht, the Netherlands.

Analysis of compensation mechanisms, occupational and job characteristics

The association between compensating mechanisms and job characteristics, occupational characteristics and personal characteristics was examined through chi-square tests and binary logistic regression analysis. For the logistic regression, the measure of association was reported by the odds ratio and the 95% confidence interval.

In line with previous research on compensation mechanisms, to serve as dependent variable in the logistic regression compensation mechanisms were dichotomized in (1) compensation mechanisms leading to productivity costs and (0) compensation mechanisms suggested not to result in productivity costs. The first category included absenteeism compensated in extra time and uncompensated absenteeism (assumed to lead to productivity costs through extra investments needed to maintain productive output or by a decrease in output). The second category included absenteeism compensated in normal working hours by the ill worker or his or her colleagues. Unknown compensation was not included in the logistic regression. Independent demographic variables included in the model were income level, educational level, age and gender. Occupational variables included were the possibility to postpone work, being in an executive function, dealing with complex problems, needing special skills, weekly working hours, having over three colleagues in the same profession, autonomy in planning work and having a mental or physical demanding profession. First, the correlation between the independent variables was examined to prevent problems with multicollinearity. Next, back and forward stepwise procedures were used to construct the optimal model fit. Variables with a P-value of ≤ 0.05 were retained in the model. Categorical variables with more than two levels were included as a block if one of the levels had a P-value of ≤ 0.05 as compared to the excluded level. We additionally tested the effects of including all variables with a P-value of ≤ 0.10. Subsequently, the Hosmer & Lemeshow test was performed to test the goodness of fit of the model.

Multiplier estimates

While information on absenteeism, compensation mechanisms, age, sex and occupation was extracted from the smoking cessation trial [260], multiplier estimates were obtained from the paper by Nicholson et al. [15]. Nicholson et al. define the multiplier *'as the cost of an absence as a proportion (often greater than one) of the absent worker's daily wage'*. The multipliers vary between jobs based on ease of substitution, the extent to which productive output is team dependent and the time sensitivity of a team's output. The paper of Nicholson et al. provides a list with multiplier estimates for a number of distinct professions. The respondents' reported professions in the smoking cessation study were compared with the list of professions in the paper of Nicholson and colleagues. If the professions were a (near) match the published profession specific estimate was applied and, if not, the median multiplier of 1.28 was applied. The applied multipliers ranged from 1.00 (waiter) to 2.26 (medical assistant). The two week multiplier estimates reported in the paper were used in cases were absenteeism exceeded one week and in other cases the 3-day multiplier estimates were used.

The potential impact of compensation mechanisms and multiplier effects

In order to meet our second research aim to explore the potential impact of compensation mechanisms and multiplier effects on productivity cost estimates, productivity costs were calculated both including and excluding compensation and multiplier effects. Productivity costs were calculated applying the friction cost approach following the Dutch pharmacoeconomic guidelines [263] and the corresponding Dutch Costing Manual [264]. The friction cost approach assumes that productivity losses only occur during the 'friction period'; this period reflects the time needed to replace (and train) a new worker [47,56,90]. If a person returns to work after a period of absenteeism shorter than the friction period and then falls ill again (and subsequently this illness leads to absenteeism), a new friction period starts. The friction period for the Netherlands was estimated to be 160 days (5.3 months) in 2008. If a person is absent for a longer period of time, productivity is assumed to be restored and, therefore, productivity costs are no longer assumed to occur. Productivity cost estimates were based on Dutch age and sex dependent national hourly wage rates in 2009 multiplied with 0.8 to adjust for the elasticity between working time and productivity, as prescribed by the costing manual [264].

In the regular scenario, scenario 1, productivity costs were calculated as is commonly done; i.e. these calculations do not take into account multiplier effects or compensation mechanisms. Productivity costs were calculated by multiplying the total lost work time of the individuals in the smoking cessation study during the last six months by the age and sex dependent average wages and an elasticity factor of 0.8. This calculation is illustrated by equation 8.1.

Equation 8.1

$$PC = \sum_{i=1}^{N} \sum_{t=1}^{T} w_i \cdot 0.8 \cdot (1 - p_{it})$$

In equation 8.1, T represents the 6 months' time horizon given the way data were gathered in this specific trial. And w_i the average age and sex dependent Dutch gross wage over 6 months based on the individuals' contracted hours, age and sex (age was assumed to be constant over the total six months). T could also represent smaller time periods, such as days. In that case w_i would reflect daily wages. The individuals' productivity is expressed by p. If an individual was absent from work (and the friction period would be irrelevant) the total 6 month p equals zero and if a respondent had not experienced any absenteeism p equals one.

In the compensation mechanisms scenario, scenario 2, productivity costs were only calculated in cases were lost work was not compensated, or compensated through extra working hours by the respondent, by colleagues or by extra hired employees. I.e., in cases were lost work was compensated during normal hours it was assumed that absenteeism did not lead to extra costs and/or lost productive output. Productivity cost calculation in scenario 2 is expressed by equation 8.2 where c is the fraction of diminished productivity that is compensated during normal working hours and $0 \leq c \leq 1$.

Equation 8.2

$$PC = \sum_{i=1}^{N} \sum_{t=1}^{T} w_i \cdot 0.8 \cdot (1 - c_{it})\ (1 - p_{it})$$

In the multiplier scenario, scenario 3, regular productivity costs estimates were adjusted by adding costs to the regular scenario related to effects of absenteeism on coworkers' and team productivity. These additional costs were calculated as a proportion of the absent worker's daily wage by multiplying absenteeism (1- *p)* with the corresponding wage rate (w_i) and the job dependent effects of absenteeism on coworkers' and team productivity (1- m_i). Consequently, productivity costs were estimated as illustrated by equation 8.3 where m_i is the job-dependent multiplier and $m_i \geq 1$. We assumed that the multipliers used already capture the elasticity of 0.8 normally used in the friction cost approach. We therefore did not correct for that elasticity in calculating multiplier costs.

Equation 8.3

$$PC = \sum_{i=1}^{N} \sum_{t=1}^{T} w_i \cdot 0.8 \cdot (1 - p_{it}) + w_i \cdot (1 - p_{it}) (m_i - 1)$$

In scenario 4, compensation mechanisms and multiplier effects were simultaneously accounted for. This is illustrated by equation 8.4 including both c and m. In this scenario productivity costs were calculated if absenteeism was compensated in extra time, or not compensated. Subsequently, costs were adjusted by adding the costs related to team effects based on professions specific multipliers.

Equation 8.4

$$PC = \sum_{i=1}^{N} \sum_{t=1}^{T} w_i \cdot 0.8 \cdot (1 - c_{it}) (1 - p_{it}) + w_i \cdot (1 - p_{it}) (m_i - 1)$$

In the final scenario (5), which may be viewed as a sensitivity analysis, again compensation mechanisms were simultaneously accounted for. Moreover, the 0.8 elasticity factor was removed under the assumption that including this elasticity together with correcting for compensation mechanisms would result in 'double correction', since the elasticity factor may already represent internal labor slack and compensation of lost work during normal working hours. Besides, it was assumed that compensation in normal working hours is not costless, but, per hour, involves half of the costs of compensation in extra hours. The productivity costs calculation in scenario 5 is expressed by equation 8.5.

Equation 8.5

$$PC = \sum_{i=1}^{N} \sum_{t=1}^{T} w_i \cdot (1 - 0.5 \cdot c_{it})\,(1 - p_{it}) + w_i \cdot (1 - p_{it})\,(m_i - 1)$$

The statistical analyses were performed with use of SPSS Statistical Software Package 17.0 (Chicago, Illinois).

8.3 Results

In total, 1,266 persons were included in the smoking cessation trial (see [260]). The response on the absenteeism and compensation questionnaire was 66% (n = 836) of the initial response. In total, 501 respondents were in paid employment. Of these, 187 respondents had experienced absenteeism in the past 6 months and, as a result, completed the questions on compensation mechanisms. Baseline characteristics of these 187 respondents are shown in table 8.1.

Compensation mechanisms

As can be seen in table 8.2, as much as 64% of the respondents stated that lost work was compensated during normal working hours by either colleagues or the employee themselves and 23% stated that their lost work was compensated in extra hours. For absenteeism of more than two week, fewer respondents stated missed work was compensated in normal hours. However, the differences in compensation mechanisms between the periods of absence were not significant (Chi-square, P = 0.214). Even for absenteeism longer than two weeks in 52% of the cases missed work was compensated in normal hours.

Differences in income and educational level of respondents in our sample did not lead to significant differences in compensation mechanisms (Chi-square, P = 0.290 and 0.109). Some significant differences in compensation mechanisms were associated with work related factors. Respondents in jobs involving both physical and mental tasks more often reported that diminished productivity was compensated in extra hours compared with respondents in jobs with either physical or mental tasks. Absenteeism occurring in occupations that require 'special skills' led to a larger percentage of missed work being compensated in extra hours compared with absenteeism in occupations that did not require such skills (Chi-square, P = 0.008). To note, only 22 of the 187 respondents stated not to need special skills in their occupation. A higher

Table 8.1 Personal and job characteristics of respondents with absenteeism (n = 187)

Baseline characteristics	
Age (SD)	37.8 (9.5)
Females (%)	96 (51.3)
Income level [a]	
Low (&)	47 (25.1)
Intermediate (%)	120 (64.2)
High (%)	10 (5.3)
Missing (%)	10 (5.3)
Educational level [b]	
Low (%)	38 (20.3)
Intermediate (%)	96 (51.3)
High (%)	53 (28.3)
Occupational and job characteristics	
Mental (%)	73 (39)
Physical (%)	64 (34.2)
Mixed (%)	50 (26.7)
Autonomy in work (%)	91 (48.7)
Possibility to postpone (%)	9 (4.8)
Special skills needed (%)	165 (88.2)
Complex problems (%)	86 (46.0)
Shift work/irregular work (%)	101 (54.0)
Executive function (%)	95 (50.8)
Colleagues in similar professions > 3 (%)	97 (51.9)
Numbers of hours worked (SD)	34.0 (12.2)
Hours worked >36 (%)	81 (43.3)

[a] Monthly net individual income: low = < 1000, intermediate = >1000, <2000, high = > 2000
[b] Educational level: low = primary and lower-level vocational school, intermediate = lower-level secondary and intermediate-level vocational school, high = higher-level secondary and higher-level vocational school and university.

Table 8.2 Compensating mechanisms

	Compensated in normal hours [a] *(%)*	*Compensated in extra hours* [b] *(%)*	*Not compensated (%)*	*Compensation unknown (%)*
Absence n = 187	119 (64)	42 (23)	17 (9)	9 (5)
Period of absence n=184				
One day n=44	27 (61)	7 (16)	6 (14)	4 (9)
2 - 7 days n=71	48 (68)	16 (23)	5 (7)	2 (3)
7-13 days n=25	19 (76)	3 (12)	3 (12)	0
> 2 weeks n = 44	23 (52)	15 (34)	3 (7)	3 (7)
Chi-Square test	P = 0.214			
Income level n=186				
Low n=47	26 (55)	13 (28)	5 (11)	3 (6)
Intermediate n=120	83 (69)	23 (19)	9 (8)	5 (4)
High n=10	4 (40)	5 (50)	1 (10)	0
Missing n=9	5 (56)	1 (11)	2 (22)	1 (11)
Chi-Square test	P = 0.290			
Educational level				
Low n=38	23 (61)	9 (24)	4 (11)	2 (5)

Table 8.2 continued

	Compensated in normal hours [a] (%)	*Compensated in extra hours [b] (%)*	*Not compensated (%)*	*Compensation unknown (%)*
Intermediate n=96	64 (67)	21 (22)	4 (4)	7 (7)
High n=53	32 (60)	12 (23)	9 (17)	0
Chi-Square test	P = 0.109			
Occupational characteristics				
Physically demanding n=64	40 (63)	14 (22)	5 (8)	5 (8)
Mentally demanding n =73	56 (77)	11 (15)	5 (7)	1 (1)
Mixed n=50	23 (46)	17 (34)	7 (14)	3 (6)
Chi-Square test	P = 0.028			
Autonomy to plan tasks n=91	57 (63)	20 (22)	12 (13)	2 (2)
No autonomy to plan tasks n 96	62 (65)	22 (23)	5 (5)	7 (7)
Chi-Square test	0.120			
Ability to postpone n=9	7 (78)	1 (11)	1 (11)	0
No ability to postpone n =178	112 (63)	41 (23)	16 (9)	9 (5)
Chi-Square test	0.720			
Special skills needed n=165	119 (63)	40 (24)	16 (10)	5 (3)
No special skills needed n =22	15 (68)	2 (9)	1 (5)	4 (18)
Chi-Square test	P = 0.008			
Complex work n=86	51 (59)	25 (29)	9 (11)	1 (1)
No complex work n=101	68 (67)	17 (17)	8 (8)	8 (8)
Chi-Square test	P = 0.040			
Irregular/shift work n=101	56 (55)	26 (26)	13 (13)	6 (6)
No irregular/shift work n=83	61 (74)	5 (18)	4 (5)	3 (4)
Chi-Square test	P = 0.065			
Executive function n=95	62 (65)	15 (16)	11 (12)	7 (7)
No executive function n=90	56 (62)	27 (30)	6 (7)	1 (1)
Chi-Square test	P = 0.023			
<3 Colleagues n=90	61 (68)	19 (21)	7 (8)	3 (3)
>2 Colleagues n=97	58 (60)	23 (24)	10 (10)	6 (6)
Chi-Square test	P = 0.631			
Full-time employment n=81	52 (64)	18 (22)	6 (7)	5 (6)
< Full-time employment n=106	67 (63)	24 (23)	11 (10)	4 (4)
Chi-Square test	P = 0.800			
Sector of occupation				
Industry n=42	27 (64)	6 (14)	5 (12)	4 (10)
Health/wellbeing n=42	19 (45)	18 (43)	4 (10)	1 (2)
Public n=5	5	0	0	0
Education n=14	7 (50)	4 (29)	2 (14)	1 (7)
Commercial n=45	34 (18)	8 (76)	2 (4)	1 (2)
Others n=37	25 (67)	6 (16)	4 (11)	2 (5)
Chi-Square test	*P = 0.107*			

[a] Includes colleagues taking over in work in normal hours or the employee itself compensates missed work after returning to work, [b] Includes colleagues taking over work in extra hours or the employee compensates missed work in extra hours or extra workers are hired

Chapter 8

proportion of respondents regularly or often facing complex problems in their occupation reported that their missed work was compensated in extra hours compared with respondents seldom facing such problems (Chi-square, P = 0.040). Respondents in executive functions more often reported that missed work was not compensated or that they did not know how the work was compensated. Moreover, a lower proportion of respondents in executive functions reported that missed work was compensated in extra hours (Chi-square, P = 0.023).

Table 8.3 reports the outcomes of the binary logistic multivariable regression model investigating the association between compensation mechanisms, demographic characteristics and work related factors. It did not make a difference whether we only included variables in the model with a P-value of ≤ 0.05 (or categorical variables with at least one level with a P-value of ≤ 0.05) or variables with P-value ≤ 0.10. As can be seen, work involving complex problems was associated with an increased likelihood of absenteeism being compensated in extra hours (odds ratio 2.346). Similarly, diminished productivity was more often compensated in extra hours in jobs involving both physical and mental work compared with jobs with mainly mental tasks (odds ratio 4.073). The other way around, being in an executive function was associated with a lower likelihood of absenteeism being compensated in extra hours.

Table 8.3 Logistic regression of compensation of lost work time and work-related factors (n = 178).

	Odds ratio (95% CI)	P-value
Constant	1.269	0.786
Executive function	0.393 (0.176 - 0.878)	0.23
Complex work	2.346 (1.018 – 5.405	0.045
Physical/ mental occupation		
Mental	*Reference category*	
Physical	2.361 (0.830 – 6.718)	0.107
Mix	4.073 (1.429 – 11.604)	0.009
Income		
High	*Reference category*	
Low	0.343 (0.067 – 1.753)	0.199
Intermediate	0.137 (0.031 – 0.617)	0.010
Unknown	0.142 (0.11 – 1.799)	0.132
Absence		
> 7 days	*Reference category*	
One day	0.142 (0.011 – 1.799)	0.132
Two days	0.930 (0.337 – 2.564)	0.888
3-7 days	0.238 (0.074 – 0.765)	0.016
Hosmer & Lemeshow	P = 0.463	

Different from the results of the Chi-square tests, the multivariable analysis did not expose a significant association between compensation in normal or extra hours and 'needing special skills'. In the multivariable analysis, this relation may have been captured by the 'complex work' variable. Although, the correlation between 'complex work' and 'needing special skills' was low (Spearman's rho 0.246) the correlation was significant (P = 0.001). Another difference was that both income and period of absenteeism now did turn out to be significantly associated with compensation mechanisms. A higher income was associated with higher levels of compensation in extra hours. Similarly, a period of absenteeism of more than one week was more likely to lead to compensation in extra hours, compared with shorter absenteeism. The multivariable model did not fail the Hosmer&Lemeshow test (P = 0.463), suggesting a good model fit.

The potential impact of compensation mechanisms and multiplier effects on productivity cost estimates

We investigated how adjusting for compensation mechanisms and the effects of absenteeism on team productivity potentially affect productivity costs estimates (separately and simultaneously). Again, respondents indicating not to know how their absenteeism was compensated were not included in the analyses. As shown in table 8.4, adjusting regular productivity cost estimates for either compensation mechanisms or multiplier effects leads to a substantial alteration of the total amount of calculated productivity costs. If productivity costs were only calculated for absenteeism compensated during extra hours or not compensated absenteeism, only 43 percent of the initial calculated productivity costs remained. Productivity cost estimates per patient were 1,100 euros lower in the compensation mechanisms scenario than in the regular scenario. Productivity cost estimates including multiplier effects were 28 percent higher than the regular estimates. I.e., the decreasing effect of including compensation mechanisms was approximately twice as strong as the increasing effect of

Table 8.4 Six month Productivity costs based on 4 scenarios in 2009 euros (n=178)

Scenarios		Six months friction costs			
		Sum	*Mean*	*SD* [a]	*Difference (%)*
1	Regular approach	€ 349,748	€ 1,965	€ 3,040	
2	Compensation mechanisms	€ 152,116	€ 855	€ 2,507	-€ 1,110 (-57%)
3	Multiplier	€ 447,487	€ 2,514	€ 3,802	€ 549 (+28%)
4	Multiplier and compensation mechanisms	€ 249,855	€ 1,404	€ 3,166	-€ 561 (-29%)
5	Multiplier and partial compensation mechanisms	€ 370,491	€ 2,081	€ 3,540	€ 116 (+6%)

[a] SD = Standard deviation

including multiplier effects. Accounting for both compensation mechanisms and multiplier effects resulted in productivity cost estimates 29 percent lower than common estimates. Again accounting for compensation mechanisms and multiplier effects, but assuming that, per hour, compensation in normal working hours involves half of the costs of compensation in extra hours and removing the 0.8 elasticity, resulted in a productivity costs estimate of 2081 euros per patient. This approximates the baseline estimate of 1965 euros.

8.4 Discussion

The first aim of our study was to increase knowledge on how diminished productivity is compensated within firms. Over half of the respondents in the smoking cessation trial indicated that their absenteeism was compensated in normal working hours by either themselves or their colleagues. These findings are in line with the results from previous research by Severens et al. (17) and Jacob-Tacken et al. (18). The way lost work was compensated was associated with work type characteristics, such as being in an executive function, having a job that is both physically and mentally demanding, facing complex problems at work and needing special skills. The need of special skills and facing complex problems may increase the difficulty to postpone work and may decrease the possibility to work taken in over in normal hours by colleagues. A significant association between being in an executive function and compensation mechanisms was previously seen in the research by Jacob-Tacken et al. (18). In (some) executive functions it might be possible to postpone work and additionally, it may be difficult to find substitutes.

If a large amount of absenteeism is indeed, as it seems, compensated during normal working hours this raises the question of how such compensation would affect productivity cost estimates. Compensation during normal hours may be taken to suggest that normal productivity levels may be maintained or restored without resulting in productivity costs. If lost work time compensated in normal hours is kept out of productivity cost calculations it is indeed implicitly assumed that such compensation is, in fact, 'for free'. However, it seems unlikely that this assumption is correct. To clarify, if compensation in normal hours is possible due to the existence of an amount of internal labor slack, this implies that (some) firms may (deliberately) have more labor present to compensate reduced productivity in ill workers. This is obviously not costless. If internal labor slack is not present, compensating absenteeism in normal hours probably involves putting extra strain on employees to maintain normal productive output. Such extra strain may also not be costless in an economic sense, as it may cause stress, may involve fewer breaks and may induce (work-related) health problems potentially leading to new absenteeism or impaired functioning at work. In our fifth productivity cost scenario we arbitrarily assumed these costs to approximate about half of the costs of compensation in extra ours, but at present, it seems unclear what the exact costs are of compensation of lost work in normal hours. Hence, to use a zero cost estimate (i.e. keeping lost work time compensated in normal hours outside productivity cost calculations) seems a strong assumption. Clearly, the translation of compensation mechanisms into productivity cost estimates deserves more attention as it is not straightforward.

Next to increasing knowledge on compensation mechanisms, the second aim of our study was to provide some insight on the (combined) potential impact of compensation mechanisms and so-called 'multiplier effects' (related to team dependency and problems with finding good substitutes in cases of absenteeism) on productivity cost estimates. The results indicated that conventional friction cost estimates differ substantially from estimates adjusted for compensation mechanisms.

Compared with conventional estimates, 43% of productivity costs remained if assumed that compensation in normal working hours was costless. In the research of Jacob-Tacken et al. (18) this percentage ranged from 23 to 32%. These lower percentages are likely to be explained by differences in calculations. Our 43% included compensation in extra hours and uncompensated absenteeism, whereas the percentages of Jacob-Tacken et al. (18) only included compensation in extra hours. Conventional estimates, moreover, differ substantially from estimates accounting for multiplier effects, or compensation and multiplier effects simultaneously. The diminishing effect of adjusting for compensation mechanisms (naively assuming compensation in normal hours is for free) on regular estimates seems stronger than the amplifying effect of applying multipliers. However, it is important to note that, although we simultaneously adjusted productivity costs estimates for compensation in normal hours and the effects of absenteeism on team output, the applied adjustment factors were originally not simultaneously assessed. In reality, compensation mechanisms and team dependencies are likely to interact with each other. In some professions team work may increase the opportunity that work is compensated by colleagues in normal hours, however, at the same time such compensation may jeopardize the team's future health and productive output. To study potential interaction between compensation mechanisms and team dependencies it would be worthwhile to investigate these effects on productive output within one study sample, ideally in a context with objectively measurable costs and output of teams and individuals. This could demonstrate the influence of both team effects and compensation mechanisms, also in the context of different job types etc., more clearly. Potentially, (context specific) average multiplier effects and compensation effects could be estimated for use in future economic evaluations.

It is important to acknowledge the impact of some (implicit) assumptions on our results. For instance, within the friction cost approach it is common to apply an elasticity factor of 0.8 between working time and productivity, as was done here in the first four scenarios. The 0.8 elasticity was first applied by Koopmanschap et al. [56], as a general correction factor for the notion that a reduction of labor time causes a less than proportional decrease in productivity. However, it remains unclear what this 0.8 elasticity exactly represents, i.e. which mechanisms underlie it. To some extent this elasticity may for instance already embody internal labor slack and compensation of lost work during normal working hours. If this is the case, then correcting for compensation mechanisms on top of the general correction factor of 0.8 used in the friction cost approach may be considered a 'double correction' and, subsequently, will lead to too low productivity cost estimates. If considering adjusting productivity cost estimates for compensation mechanisms within the friction cost approach it may seem reasonable to not additionally apply the 0.8 elasticity factor. As shown in scenario 5, if the elasticity factor is not applied, and it is assumed that compensation in normal hours is not without costs, but amounts to half of the costs of compensation in extra hours, productivity cost estimates including multipliers approximate the baseline estimate of scenario 1. Although more research is needed to further investigate compensation mechanisms, multiplier effects and the costs of compensation, the standard productivity costs estimates (scenario 1, using a general 0.8 correction factor) may therefore be considered a reasonable base case, in which multiplier effects as well as compensation mechanisms are included.

While our study was mainly intended to increase the knowledge on compensation mechanisms and to demonstrate the potential impact of compensation mechanisms and multiplier effects on productivity cost estimates, it suffered from some important limitations. First, our sample size was relatively small. Although 1266 people were included in the smoking cessation trial, only 187 were in paid profession and had experienced absenteeism during the previous six months. Second, we used a convenient sample existing of smokers willing to quit instead of a general public sample. The results may, therefore, not be fully generalizable. Third, compensation mechanisms were based on self-reporting data among employees who had experienced the absenteeism themselves. In some cases, employers or colleagues may have better insight in how work was compensated during the employees' absenteeism. Fourth, the multiplier effects applied were extracted from previous research [(15)]. In half of the cases we needed to apply the median multiplier of 1.28 presented in the paper by Nicholson et al. [(15)] because the professions of our respondents did not match the professions presented in their empirical work. Moreover, the original multipliers were collected in the United States and therefore may not be directly transferable to the Dutch occupational situation (for an extensive discussion on transferability issues, see [(265)]). Since Dutch estimates of multipliers are lacking, we took the US figures as the best available estimates. Fifth, we only focused on absenteeism. Compensation mechanisms and multiplier effects are, however, potentially equally important in diminished functioning at work due to health problems (i.e. presenteeism).

Despite these limitations, our research indicates that both multiplier effects and compensation mechanisms are important to consider in the area of productivity cost research. This study was the first to simultaneously explore these effects. *How* to exactly consider compensation mechanisms and multiplier effects concurrently in future research remains uncertain. Nevertheless, adjusting productivity cost estimates for compensation mechanisms while ignoring the effects of ill health on team output, or the other way around, may potentially lead to incorrectly low or high productivity cost estimates and may therefore not be advisable. Although our research (and the previous research by Severens et al. [(17)] and Jacob-Tacken et al. [(18)]) implies that over half of the lost work is compensated in normal working hours, the actual effects on productive output and related productivity costs are unknown. Moreover, the presence of team dependencies and the subsequent tangible productivity costs are equally underexplored. Therefore, further research to increase insight in these processes is critical before considering adopting adjustments for compensation mechanisms and multiplier effects as common standard.

9.

Predicting productivity costs based on EQ-5D:

An explorative study

Based on: Krol M, Stolk E, Brouwer W.

Productivity costs predictions based on EQ-5D: an explorative study.

Submitted paper.

Summary

Background: Productivity costs are often ignored in economic evaluations. In order to facilitate productivity costs inclusion it has been suggested to estimate productivity costs indirectly using quality of life data. Objective: This study aimed to derive and validate an algorithm for predicting productivity on the basis of quality of life data using the EQ-5D-3L. Methods: A large representative sample of the Dutch general public (n = 1,100) was asked in a web-based questionnaire to state their expected level of productivity in terms of absenteeism and presenteeism for multiple EQ-5D health states. Based on these data, two generalized estimating equations (GEE) prediction models were constructed: i) a model predicting levels of absenteeism and ii) a model predicting presenteeism. The models were validated by comparing model predictions with conventionally measured productivity within a group of low back pain patients. Results: Predicted absenteeism levels based on EQ-5D health state closely resembled conventionally measured absenteeism levels. Productivity losses related to presenteeism seemed somewhat overestimated by our prediction model. Measured and predicted productivity were moderately but highly significantly correlated. Conclusions: Overall, it appears possible to make reasonable productivity predictions based on EQ-5D data. Further exploration and validation of prediction algorithms remains necessary, however, especially for presenteeism.

9.1 Introduction

Economic evaluations are increasingly used to assist national decision makers in selecting which new and established medical technologies should be included in national health care benefit packages. In an economic evaluation an intervention is compared to one or more alternatives in terms of costs and (health) effects. Which costs (and effects) are to be included in the evaluation depends on the perspective from which economic evaluations are performed.

Both in terms of theory and practice, there is dissensus regarding the appropriate perspective from which to perform economic evaluations. These opposing viewpoints appear to be related to different assumptions regarding the decision making context in which the results of economic evaluations need to be used (e.g. [(33)]). Broadly speaking, two dominant perspectives can be distinguished: (i) a health care perspective, which reflects the notion that health economic evaluations aim to inform health care decision makers (allocating some –fixed – health care budget), and (ii) a societal perspective, reflecting the notion that the underlying aim of the decision maker informed by economic evaluations is to improve social welfare. Here, we will not revisit this discussion, but merely take the viewpoint that from a conventional welfare-economic view the societal perspective can be considered and commonly is considered (e.g. [(2)]) an appropriate viewpoint since only by including all relevant societal costs and effects optimal (that is welfare improving) societal decision-making is possible [(34)]. If one adopts this viewpoint, this implies that all relevant effects and costs, also those falling outside the health care sector (and thus may be deemed irrelevant when taking a health care perspective) should be included in the evaluation, in principle. This obviously expands the scope of economic evaluations. In practice, it regularly proves difficult to adopt such a societal perspective and adequately include all relevant costs in an evaluation.

One cost category that is sometimes difficult to include is that of productivity costs, which can be defined as *"costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid."* [(8)] Despite the societal importance of unpaid production loss, in this chapter, we focus on productivity costs related to paid profession. Productivity costs are frequently neglected in economic evaluations, also by those claiming to take a societal perspective [(10,44)]. If included, however, productivity costs often reflect a large part of total costs in the evaluation and can have a strong impact on cost-effectiveness outcomes [(9,42,44)]. This makes the neglect of productivity costs in a majority of economic evaluations particularly worrisome.

The neglect of productivity costs in economic evaluations may partly be related to concerns regarding the equity implications of including productivity costs. Indeed, inclusion of such costs may have distributional consequences. For instance, it can lead to more favorable cost-effectiveness results in persons with a paid job than in those without paid work. It is questionable whether distributing health care accordingly will be in line with societal preferences. [13] However, while such concerns must be addressed, it must be noted that inclusion of productivity costs can also result in more favorable results in patients without paid work (e.g. if the treatment itself causes absenteeism) [44] and that fully ignoring these real societal costs as a response to potential conflicts between efficiency and equity seems hard to defend as well.

Another reason why productivity costs may be neglected in economic evaluations claiming to adopt a societal perspective may be the time and effort needed to collect the data required to calculate productivity costs. Collecting the appropriate data involves asking patients a broad range of questions regarding their employment status. Additionally it is necessary to repeatedly ask patients about their attendance at work and their level of functioning while being at work. Besides, it may not always be possible to collect necessary data regarding patients' productivity. For instance, when working with retrospective data, or when data collection is bounded to information available in medical files productivity data is commonly unavailable. Moreover, at this moment, there is no scientific consensus on how to identify, measure and value productivity losses and gains, which may make researchers reluctant to collect data [11,51,233].

In order to facilitate the inclusion of productivity costs and thus reducing the neglect of these important costs in economic evaluations that wish to adopt a societal perspective, also in cases where primary data collection is difficult, it has been suggested to create prediction models capable of predicting productivity costs based on patient characteristics such as age, sex and quality of life data [20]. Previous research indeed has shown that productivity is significantly correlated with levels of quality of life [19], but prediction models have not yet been created, to our knowledge. These models would facilitate inclusion of productivity costs on the basis of typically easily available information. Although using such models may be perceived to be a second best option compared to the collection of actual data on productivity, including a reasonable estimate of productivity costs seems preferable to neglecting these costs and, therefore, implicitly estimating them to be zero, as has also been argued in the context of other societal costs [266]. This point seems especially valid for productivity costs, since these can have such a profound effect on the results of economic evaluations (and therefore decision making). Obviously, the desirability of using predicted productivity costs rather than excluding these costs also depends on the accuracy of such predictions.

Creating a prediction model requires establishing a link between productivity (involving absenteeism and reduced productivity at work, commonly labeled 'presenteeism') and quality of life. Establishing such a link is not straightforward. It has been previously attempted by combining existing patient data on quality of life and productivity (19). However, it turned out to be difficult to create a prediction model with the available data due to limited sample size; difficulty in linking observed productivity to disease stage and a lack of variety in health states in terms of quality of life. A possible way of establishing the link between productivity and quality of life more directly is to ask individuals to state to what extent they expect to be productive in different health states, described with some generic quality of life measure. This has as advantage that it allows including a wider range of health states in terms of quality of life and therefore a better analysis of the statistical link between health state and productivity than commonly would be the case with data including information on actual absenteeism, presenteeism and quality of life. Notably, such an approach does require a more elaborate validation of the prediction model, since there may be discrepancies between expected and actual absenteeism and presenteeism.

This chapter presents the results of a study exploring the possibility to predict patients' productivity based on the three level version of the EuroQol's EQ-5D (24). The EQ-5D is a descriptive system with five dimensions and three levels per dimension, summing up to a total of 243 unique health states. Health states are expressed in five-digit codes. For example, 22311 describes the following state: Some problems with walking about, some problems with washing and dressing, unable to perform usual activities, no pain or discomfort and no anxiety or depression. In order to facilitate creating a productivity prediction model, a large sample of the general public was asked to state their expected level of productivity (in terms of both absenteeism and diminished functioning while attending work) for multiple EQ-5D health states. Based on these data, two models were constructed: i) a model predicting levels of absenteeism and ii) a model predicting presenteeism. Subsequently, these two models were validated by comparing model predictions with conventionally measured productivity within a group of low back pain patients, demonstrating the predictions to be reasonably close to actual losses.

9.2 Methods

Data collection

Respondents

To establish a link between productivity and EQ-5D health states we collected data on expected productivity through an online survey among a large sample of the Dutch general public. Invitations were sent out to potential respondents to obtain a representative sample of 1,000 members of the Dutch general public in paid profession aged between 18 and 65.

Questionnaire

The questionnaire was designed for self-completion. Respondents were first asked some general back ground characteristics such as age, sex, education and their self-assessed health today by means of the EQ-5D (3 level version). Subsequently, respondents were presented with 16 distinct EQ-5D health states. For each of these states respondents were asked to indicate whether they expected to be present at work (absenteeism), and if so, how they would expect to function while attending work in terms of productivity (presenteeism). Productivity loss at work was assessed using the quantity scale of the QQ-method. The QQ-method is an instrument developed to quantify productivity loss related to presenteeism (48). Respondents in our study were asked to state the quantity of production on a ten-point scale, where '0' represented not able to do any work and '10' represented being able to do the same amount of work as usual. Respondents not showing any variation in the responses in the absenteeism and presenteeism questions were excluded from further analyses.

In total, 96 EQ-5D health states out of the 243 possible states were included in our study. Health states were selected based on orthogonal design. An orthogonal array of strength 3 provided 54 different combinations of domains and levels (267). We used two generators to create two additional orthogonal sets of 54 health states. Subsequently, we removed implausible states (26), and we removed doubles. The 96 remaining EQ-5D states were divided into 6 subsets (See appendix 9A) of 16 EQ-5D states by means of block design in order to have an optimal variety of health states within these subsets. Respondents to the questionnaire were randomly assigned one of these six subsets. Health states within these subsets were presented to the respondent in a random order. With a sample of 1000 respondents, this would result in approximately 160 observations per each of the 96 health states.

Data analysis

Model construction

Based on the respondents' expected productivity in the EQ-5D states, two prediction models were constructed (one for absenteeism and one for presenteeism). For this purpose generalized equation estimations (GEE) were used. GEE are appropriate for analyzing correlated response data (which in this case refers to the 16 productivity estimates made by each individual respondent). Respondents were numbered to be included as subject variable to account for within-subject correlation. Since the 16 EQ-5D states were presented to the respondents randomly, an exchangeable working correlation matrix structure was specified. This structure assumes homogenous within-subject correlation between the 16 productivity estimates.

GEE models

Given the binomial outcome of the absenteeism question, a GEE binary logistic regression was applied to create the absenteeism model. The dependent variable in the absenteeism model was the dichotomous variable 'expected to present work in presented EQ-5D state'. Independent variables included were respondents' age and dummy variables for sex (males labeled '1' and females '0') and the domain levels of the EQ-5D states. Other (work-related) covariates were not included, since this study aimed to explore the possibilities to estimate productivity effects in absence on data on absenteeism and presenteeism. We assumed that in cases where data on absenteeism and presenteeism is lacking, other work related information will also be unavailable.

The absenteeism model is expressed in equation 1 where y is the logit of the probability of attending work. Mob_2 represents a level 2 score on the EQ-5D mobility domain and mob_3 a level 3 score on mobility. Likewise, sc_2 and sc_3 represent self-care domains, levels 2 and 3, respectively and da_2 and da_3 daily activities level 2 and 3 and so on. By means of equation 2 the probability P of attending work in a specific EQ-5D health state is calculated. Xβ denotes the linear combination of regression coefficients and covariates for each EQ-5D health state.

Equation 9.1 Absenteeism model

$$Y = \beta_0 + \beta_1 * mob_2 + \beta_2 * mob_3 + \beta_3 * sc_2 + \beta_4 * sc_3 + \beta_5 * da_2 + \beta_6 * da_3 + \beta_7 * pain_2 + \beta_8 * pain_3 + \beta_9 * mood_2 + \beta_{10} * mood_3 + \beta_{11} * age + \beta_{12} * male$$

Equation 9.2 Probability of attending work

$$P = \frac{e^{x\beta}}{e^{x\beta} + 1}$$

Equation 3 represents the presenteeism prediction model. A negative binomial distribution with a log link demonstrated to be the best fit model, based on the Quasi Likelihood under Independence Model Criterion and the Corrected Quasi Likelihood under Independence Model Criterion. The dependent variable in the presenteeism model was the expected level of functioning E_{lf} (ranging from zero and ten) in the given health states, when at work. Similar to the absenteeism model the following independent variables were included: respondents' age, sex and dummy variables for the domain levels of the EQ-5D states.

Equation 9.3 Presenteeism model

$$E_{\mathrm{lf}} = ln^{-1}[\beta_0 + \beta_1 * mob_2 + \beta_2 * mob_3 + \beta_3 * sc_2 + \beta_4 * sc_3 + \beta_5 * da_2 + \beta_6 * da_3 + \beta_7 * pain_2 + \beta_8 * pain_3 + \beta_9 * mood_2 + \beta_{10} * mood_3 + \beta_{11} * age + \beta_{12} * male]$$

Equation 4 combines the absenteeism and presenteeism prediction model into one predicted productivity measure (expected productivity in a EQ-5D state: E_{pr}) by multiplying the probability of attending work (*P)* by the expected level of functioning while attending work E_{lf}. Outcomes are divided by 10 to arrive at a 0 to 1 scale, where '0' illustrates zero productivity and '1' normal productivity (i.e. no absenteeism and no presenteeism).

Equation 9.4 Expected productivity level

$$E_{\mathrm{pr}} = \frac{P \cdot E_{if}}{10}$$

Data analysis was conducted in SPSS 18.0, Chicago, Illinois.

Validation

Internal validity

Internal validity of the absenteeism and presenteeism models was tested in several ways. First, it was examined whether the EQ-5D model parameters pointed in the right direction; i.e. a level two or three on the EQ-5D domains was expected to have a diminishing effect on the probability to attend work and on the level of functioning if attending work. Moreover, a level three score was expected to have a stronger negative effect than a level two score.

Subsequently, productivity levels were estimated for an average aged worker for all possible EQ-5D states based on the models' parameters. For illustrative purposes we chose this to be a 40 year old male, since in the Netherlands, males fulfill a larger fraction of paid productivity than females. The absenteeism model outcomes were presumed to cover the full range of probabilities, e.g. health state 11111 was hypothesized to have a probability of (almost) one of attending work and for health state 33333 this probability was hypothesized to be (almost) zero. The presenteeism model outcomes range was expected to start at 1 (normal functioning in 11111). However, it was not expected to range to zero, since in case of almost zero productivity, it is unlikely that people are still at work (and as long as they can get to work, their productivity is unlikely to be zero).

To test how well the constructed models predict within sample, we used the models to predict the sample populations estimated absenteeism, presenteeism and productivity levels in the given EQ-5D states. We subsequently calculated the Mean Absolute Error.

External validity

External validity was assessed on the predictive validity of the two models. It was evaluated through how well the models were able to predict productivity levels measured with a commonly used instrument in economic evaluations. In order to do so, conventionally measured absenteeism by means of the in the Netherlands quite often applied Health and Labor Questionnaire (HLQ) [84] was compared with predicted absenteeism based on EQ-5D health state. Similarly, measured presenteeism by means of the quantity scale of the QQ method, where '0' represented not able to do any work and '10' represented being able to do the same amount of work as usual [48] was compared with predicted presenteeism Moreover, the association between conventionally measured and predicted productivity was assessed by means of correlation coefficients. Given the skewed nature of productivity outcome measures (e.g. many individuals experience little to no absenteeism and few have very high levels of absenteeism), correlation was examined non-parametrically by determining Spearman's rank correlation coefficients and their 95% Bootstrap confidence intervals.

The external validity data set consisted of data of 500 patients' with low back pain included in a clustered randomized controlled trial to test the effects of an active implementation strategy for the Dutch physiotherapy guideline for low back pain. Details of this study and the piggy-back economic evaluation can be found elsewhere [268,269]. Patients were included in the study before receiving physiotherapy for their low back pain. Patients' were requested to complete a questionnaire at baseline, and at 6, 12, 26 and 52 weeks follow-up. Among other things, the questionnaire included questions on the patients' quality of life (measured through the EQ-5D) and on productivity. Absenteeism was measured using the Health and Labor Questionnaire [84] and included questions on attendance at work during the last 14 days. Presenteeism was measured by means of the Quality and Quantity method [48]. Patients were asked to state how much work they were able to perform on their last day of work, as compared with their normal functioning, and how they rated the quality of their work on that day. For our study we only included data of the quantity question.

A disadvantage of the validation dataset was that patients' stated EQ-5D reflected their quality of life on the day they completed the questionnaire, while the productivity questions related to their productivity during the last two weeks. Although the patients' EQ-5D on the day of the questionnaire does not necessarily reflect their quality of life during the last two weeks, we had to assume it was the same during the previous two weeks in order to test our prediction models. For this reason, baseline data seemed most appropriate; patients had not received treatment yet and the vast majority (91%) stated to suffer from low back pain for longer than one week. In the baseline data patients' EQ-5D would have been more likely to be constant during the preceding period than in the data collected during and after treatment.

Low back pain patients in the external validation data set were included if they were in paid profession and data was available on their age, sex, EQ-5D at baseline and their productivity during the past two weeks. Patients' productivity levels were constructed from the data in the following manner. Data was available on the number of days a week normally worked and the days absent during the last two weeks. Based on these numbers the percentage work attendance during the past two weeks compared to normal was calculated. This calculated percentage served as a proxy for the probability of attending work in the patients' EQ-5D states. Presenteeism was based on patients' stated quantitative level of functioning (compared with patients' normal functioning) during the last working day.

Table 9.1 characteristics of respondents to the online questionnaire

Respondents' characteristics	n =1013
Age (SD)	41.6 (12.40)
Females (%)	502 (49.6)
EQ-5D score (SD)	0.92 (0.14)
Education	
Lower (%)	117 (11.5)
Medium (%)	575 (56.8)
Higher (%)	321 (31.7)
Marital status	
Single (%)	158 (15.6)
Married/living together (%)	728 (71.8)
Divorced (%)	89 (8.8)
Widow/widower (%)	11 (1.1)
Other (%)	27 (2.7)
Work	
In paid employment (%)	931 (91.9)
Self-employed (%)	82 (8.1)

9.3 Results

Respondents

In total, 1,100 respondents completed the online questionnaire. 87 of these respondents were excluded from further analysis since they seemed insensitive to variation in the presented health states; 33 respondents stated to attend work in any given health state and 54 respondents stated always to call in sick. Excluded respondents did not differ from the not excluded respondents in terms of age, sex, education and income.

As seen in Table 9.1, the respondents were in good health on average. Most of them were married or living together, had a medium level of education and were in paid employment. A small minority was self-employed.

Prediction models

The absenteeism analysis was based on 16,208 estimations of respondents on whether they would expect to be able to attend work in the presented EQ-5D health states; i.e. this resulted in over 160 estimations for each of the 96 included EQ-5D health states. In 5,803 of these cases respondents expected to attend work and were subsequently asked how they expected to function while attending work as compared to their normal functioning. Table 9.2 shows the output of both GEE models expressed in equation 1 and the presenteeism prediction model expressed in equation 3. Sex proved insignificant in the absenteeism analysis and was therefore not included in this model. Likewise, age

Table 9.2 GEE Models

	Absenteeism n = 16,208					Presenteeism n = 5,803				
		95% confidence interval					95% confidence interval			
Variable	Coefficient	Lower limit	Upper Limit	SE	P-value	Coefficient	Lower limit	Upper Limit	SE	P-value
Intercept	2.414	2.087	2,741	.167	<.001	2.029	2.005	2.053	.012	<.001
Mobility 2	-.264	-.341	-,188	.039	<.001	-.045	-.057	-.033	.006	<.001
Mobility 3	-2.054	-2.217	-1,891	.083	<.001	-.179	-.215	-.144	.018	<.001
Self-care 2	-.316	-.405	-,227	.045	<.001	-.034	-.050	-.019	.008	<.001
Self-care 3	-1.193	-1.314	-1,072	.062	<.001	-.096	-.118	-.074	.011	<.001
Daily activities 2	-.395	-.495	-,296	.051	<.001	-.063	-.078	-.048	.008	<.001
Daily activities 3	-1.761	-1.889	-1,633	.066	<.001	-.173	-.194	-.151	.011	<.001
Pain 2	-.360	-.452	-,267	.047	<.001	-.065	-.081	-.050	.008	<.001
Pain 3	-1.740	-1.875	-1,604	.069	<.001	-.178	-.202	-.154	.012	<.001
Mood 2	-.229	-.322	-,136	.048	<.001	-.068	-.083	-.053	.008	<.001
Mood 3	-.942	-1.047	-,836	.054	<.001	-.135	-.153	-.117	.009	<.001
Age	-.009	-.016	-,003	.003	.006					
Male						.037	.010	.064	.0137	.006

was not significantly associated with productivity in the presenteeism analysis and excluded accordingly.

Validity

Internal validity

Internal validity of both the absenteeism and the presenteeism model appeared satisfactory. All EQ-5D parameters were highly significant in the GEE models and all pointed in the expected (negative) direction. Moreover, all level 3 scores had a stronger negative effect on outcomes than level 2 scores. A level 3 score in the domain mobility was the most influential parameter in both models. The probability of attending work in a given health state was negatively affected by age. This indicates, all other things equal, older persons would be more likely to call in sick in a specific health state. However, when at work, age did not seem to affect (expected) functioning. Females expected to be less productive than males did when being at work while ill.

To further explore internal validity we compared the models' estimated productivity levels for all EQ-5D states for a 40 year old male with our prior expectations (the full list of estimates can be found in appendix 9B). Different from our expectations, the productivity in health state 11111 (full health) did not equal 1. The probability to attend work in 'full health' was 0.89 and level of functioning while attending work was estimated to be 7.9. The 'level of functioning scale' ranged from 7.9 to 3.7. As hypothesized, the scale did not end at '0', since it would be unlikely that fully unproductive persons would be present at work. The highest point of the scale did not equal '10'. The probability of attending work in state 33333 was, as one would expect, 0.00.

We found no straightforward relationship between *utility score* and productivity. To illustrate, state 31311 has a Dutch national tariff value of 0.477 and an estimated productivity level of 0.08, whereas state 22123 has a tariff value of 0.166 yet an estimated productivity level of 0.32. Different levels of the EQ-5D domains impact quality of life differently than they impact productivity. Impaired mobility and not being able to perform daily activities have a relatively strong impact on expected productivity, whereas the impact of anxiety or depression on expected productivity is much smaller.

The mean absolute error comparing the respondents expected absenteeism, presenteeism and subsequent total productivity with the models predictions was low, (respectively 0.05, 0.32 and 0.03) indicating a good within sample prediction.

External validity

In order to examine external validity we investigated how well our prediction models predicted productivity levels in an external data set concerning a group of patients suffering from low back pain [269]. Data on productivity was available of 341 patients and included information on absenteeism, productivity on the last day at work and quality of life as measured using the EQ-5D. Characteristics of these patients are presented in table 9.3. Average age was approximately 40 years, their average Dutch national tariff EQ-5D score was 0.71 and most respondents suffered from low back pain for more than 1 month.

Based on the patients reported EQ-5D scores for each dimensions and their age and sex, we predicted the population's probability of attending work and their level of functioning if attending work with our two models. We tested whether our models' predictions were significantly associated with the conventionally measured productivity levels in the low back pain patients. As can be seen in table 9.4, the correlation coefficient between measured productivity by means of the HLQ and predicted productivity was significantly different from zero (all P-values < 0.001). Attending work and total productivity were moderately correlated (between 0.3 and 0.5) and the 'functioning when attending work' correlation was low (< 0.3).

We subsequently compared the model predictions with the patients reported absenteeism during the last two weeks and their reported level of functioning on their last working day. Additionally, we compared the conventionally measured productivity levels with the predicted levels (calculated by means

Table 9.3 Characteristics low back pain patients

Characteristic	n = 341
Age (SD)	39.9 (10.6)
Females (%)	154 (45)
Education	
Lower (%)	11 (3)
Medium (%)	201 (59)
Higher (%)	128 (38)
EQ-5D score (SD)	0.71 (0.2)
Duration of complaints	
< 7 days (%)	30 (9)
1 to 4 weeks (%)	87 (26)
1 to 6 months (%)	139 (41)
> 6 months (%)	81 (24)
Cannot recall (%)	4 (1)
Working hours a week (SD)	32.0 (11.0)
Absenteeism (%)	149 (44)
Level of functioning last working day (SD)	7.5 (2.7)

Table 9.4 Correlation between measured and predicted productivity

	R	Bootstrapped 95% confidence interval		P-value
		Lower limit	Upper limit	
Attending work	0.351	0.252	0.444	< 0.001
Functioning at work	0.257	0.154	0.363	< 0.001
Total productivity	0.375	0.276	0.469	< 0.001

of equation 4). In figures 9.1, 9.2 and 9.3 the measured and predicted levels of attending work in a particular EQ-5D state were plotted against each other. Within these figures distinction is made between EQ-5D states with 10 or more observations, EQ-5D states with 3 to 10 observations and EQ-5D states with only one or two observations. The diagonal illustrates the 'perfect fit line' between observed and predicted productivity. As figure 9.1 illustrates, when there are more observations per state, on average, predicted absenteeism seems reasonably comparable to measured absenteeism. However, when there are only a few observations, predicted and measured absenteeism differ more.

Notably, the measured percentage of attending work of the low back pain patients' stating to be in health state 11111 (i.e. full health) was 0.87 (similarly as the predicted percentage of attendance in full health) and not one. The average measured percentage of attending work of the low back pain patients was 0.74 (95% confidence interval 0.70-0.78). The mean predicted probability of attending work was 0.71 (95% confidence interval 0.69-0.73). Our model only slightly overestimated absenteeism levels for the low back pain patients.

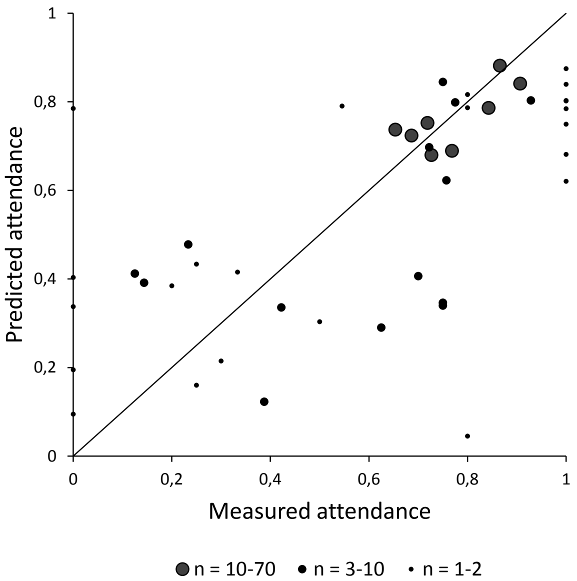

Figure 9.1 Measured versus predicted absenteeism in low back pain patients

The average measured level of functioning when attending work was 0.75 (95% confidence interval 0.72-0.78) and the average predicted level of functioning was 0.66 (95% confidence interval (0.66-0.67). As can also be seen in figure 9.2, our presenteeism prediction model overestimated production losses due to presenteeism.

The combined average measured productivity level was 0.59 (95% confidence interval 0.56-0.63) and the average predicted productivity level for the low back pain patients was 0.48 (95% confidence interval 0.46-0.49). I.e., total productivity estimates based on our absenteeism and presenteeism model seemed to overestimate productivity loss with about 11% for this particular patient group. As can be seen in the figures 9.1, 9.2 and 9.3, the prediction models are not able to adequately predict productivity on an individual level.

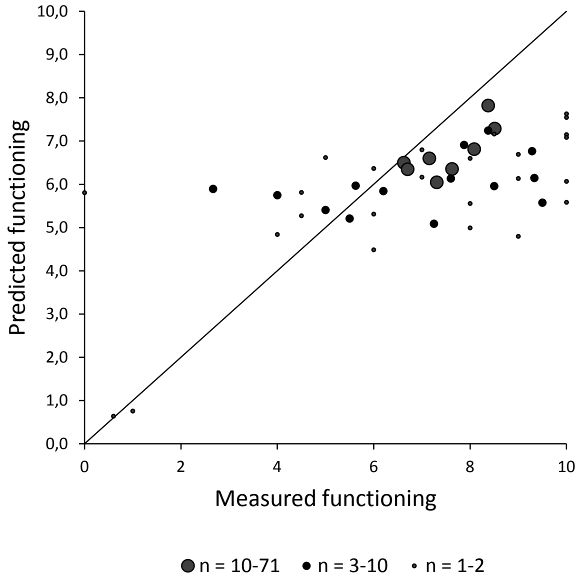

Figure 9.2 Measured versus predicted presenteeism in low back pain patients

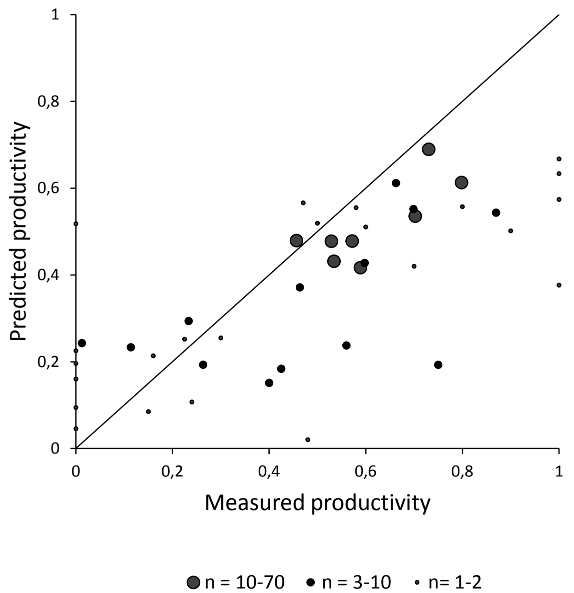

Figure 9.3 Measured versus predicted productivity in low back pain patients

9.4 Discussion

This study explored whether (impaired) productivity, and subsequently productivity costs, can be reasonably predicted based on patients' EQ-5D health states. The results of this large web-based study among the Dutch general public seem promising. Measured and predicted productivity were moderately (but highly significantly) correlated. Predicted absenteeism levels based on EQ-5D health state on average closely resembled conventionally measured absenteeism levels for a group of low back pain patients. Note that absenteeism normally induces the major part of total productivity costs and, in that sense, this finding is encouraging. Productivity losses related to presenteeism, and consequently total productivity losses, seemed to be somewhat overestimated by our prediction model.

A surprising finding was that respondents to the online survey as well the low back pain patients indicated that in full health (11111) they (would) function less than normal. This may be explained by several reasons: i) respondents are known to suffer from end-aversion bias [(109)], ii) respondents may have interpreted a '10' as optimal rather than as normal functioning, iii) the low back pain patients in health state 11111 may have had other (non-health) reasons for functioning less than normal and iv) state 11111 may leave room for mild problems not necessitating moving to level 2 on any domain, yet somewhat limiting functioning.

Limitations

Some limitations of our study need noting. A first obvious limitation of our study was that we based the prediction models on hypothetical estimates instead of actual absenteeism and presenteeism. Respondents to our survey were asked whether they expected to attend work in a given EQ-5D state and, if so, how they expected to function. Such a task might be experienced as quite difficult, especially given the rather abstract nature of EQ-5D state descriptions. Second, our prediction model for presenteeism seemed less accurate than the absenteeism model. Predicted productivity loss related to presenteeism was approximately 9% higher than measured productivity loss. Respondents of the general public may overestimate the effects of ill-health on productivity, perhaps due to underestimation of coping strategies and adaptation. In that sense, there is an analogy with the common observation that the general public values health state lower, on average, than patients do [(270)]. A related explanation might be the nature of the QQ-scale we used to measure presenteeism.

Respondents in our survey as well as in the study of low back pain were asked to compare their (expected) level of functioning with their normal functioning. If people with low back pain have a lower reference point for 'normal functioning', this may lead to relatively high levels of indicated productivity, whereas the scores would be relatively low for healthier individuals.

How the model's overestimation of productivity losses related to presenteeism would affect subsequent cost-effectiveness outcomes is unclear. In that way, two aspects are important to consider. First, the overestimation of the model was determined by comparing predicted outcomes with measured outcomes by means of self-assessment. Obviously, such measured outcomes will be surrounded with uncertainty.

Second, within an economic evaluation the overestimation of productivity losses using the prediction model would affect both the estimated productivity changes in the intervention and the control group. As a result, predicted incremental changes are a combination of the production losses in both arms with less predictable outcomes. This may mitigate the effect of the overestimation in certain cases, where the losses in both arms are overestimated.

An additional limitation of our study was that we only compared our models' productivity estimates with productivity measured with the HLQ [(84)] and the QQ-method [(48)]. Productivity cost calculations based on different measurement instruments sometimes differ substantial [(83)]. It is unclear how our model estimates relate to productivity measured with other instruments.

Implications

The findings of our study indicate that the absenteeism prediction model produces reasonable estimates based on comparing predicted absenteeism with measured absenteeism by means of the HLQ. If this finding proves generalizable, using these estimates may be preferred over the common alternative of using zero estimates when direct measurement of productivity is, for some reason, not possible. Since our presenteeism model seems to lead to an overestimation of productivity losses additional research is needed to investigate whether a different study design, preferably using actual data on presenteeism and EQ-5D health states, would lead to more accurate estimates.

It is important to note that our models are neither intended nor appropriate for individual productivity predictions. Productivity and quality of life are (evidently) not perfectly correlated. Productivity losses for instance also partly depend on job characteristics and personality. Our models are based on general public estimations with respondents in a large variety of professions and consequently, these models maybe less appropriate for productivity predictions for more homogeneous groups (e.g. groups in similar professions when evaluating some firm-based intervention).

Although the productivity prediction models we constructed could be used to provide rough estimates of absenteeism and presenteeism in cases where EQ-5D data is available, but direct productivity measurement is impossible, translating productivity estimates into productivity *cost* estimates may not be easy, also given the valuation approach applied. When applying the human capital approach [(88)] it is sufficient to have information on age, sex and total productivity levels during the time period of interest. Subsequently, productivity estimates can be translated into costs based on national statistics regarding average age and sex dependent employment rates and average income. However, calculating productivity costs when applying the friction cost approach would be more difficult, since this requires information on duration and frequency of individual absence periods [(9,90)]. In the friction cost method productivity costs are calculated for every period of absenteeism, up to a certain limit (i.e. the friction period). Accordingly, additional assumptions need to be made if productivity costs are based on predicted absenteeism, especially regarding length and frequency of absenteeism. Such assumptions could be based on disease characteristics (and potentially disease models).

It is important to investigate how productivity predictions can be best translated into productivity costs and subsequently to test the accuracy of such cost estimates. Moreover, given the promising results of this explorative study it would be worthwhile to conduct research aiming at constructing productivity prediction models based on actual rather than hypothetical productivity and to further test the validity of such models.

9.5 Conclusions

Based on our findings, productivity estimates based on quality of life data seems of additional value as a second best solution if the first best solution of direct measurement is impossible. Although the constructed productivity prediction models presented in this chapter are not adequate nor intended for predictions on an individual level, they seem to provide relatively reasonable estimates on group level. If nothing else, this makes future research in this area appear worthwhile. These findings moreover imply that productivity estimates based on EQ-5D data may be an improvement over ignoring productivity costs, implicitly assuming these costs are zero. Reasonable estimates seem better than unreasonable neglect.

Acknowledgements

We are grateful to Lucas Goossens and Mark Oppe for their useful comments during the process of the study on which this chapter is based.

Appendices

9 A EQ-5D states stratified by group

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
11133	11123	11113	11111	11221	12121
11233	11212	11231	11122	11311	12211
12123	12321	11312	11313	12312	12232
12323	12332	12112	11331	12333	12233
13232	13223	12313	11332	13213	13231
13311	21112	13312	13222	13322	21212
13313	21123	21332	13323	21121	21312
21233	21232	22111	21133	21211	22132
21311	21331	22122	22123	21323	22133
22223	23221	22232	22231	22222	22313
22311	23233	23213	23311	23231	22322
31121	23313	23321	31322	23332	22331
31231	31311	31232	32222	31223	23312
32231	32312	32321	33212	32121	31313
33211	33221	32331	33233	32213	32233
33312	33321	33333	33311	33232	32311

9 B Predicted productivity based on EQ-5D state for 40 year old males

EQ-5D state	Dutch national tariff [115]	Probability attending work	Level of functioning attending work	Total productivity
11111	1	0.89	7,9	0,7
11211	0.897	0.84	7,4	0,62
21111	0.893	0.86	7,6	0,65
21211	0.861	0.8	7,1	0,57
12111	0.847	0.85	7,6	0,65
11121	0.843	0.84	7,4	0,62
12211	0.815	0.79	7,2	0,57
22111	0.811	0.81	7,3	0,59
11221	0.811	0.79	6,9	0,55
21121	0.807	0.81	7,1	0,57
11112	0.805	0.86	7,4	0,64
22211	0.779	0.75	6,9	0,51
21221	0.775	0.74	6,6	0,49
11212	0.773	0.81	6,9	0,56
21112	0.769	0.83	7,1	0,58
12121	0.761	0.8	7,2	0,57
21212	0.737	0.76	6,6	0,5
12221	0.729	0.73	6,7	0,49
22121	0.725	0.75	6,8	0,51
12112	0.723	0.82	7,1	0,58
11122	0.719	0.81	6,9	0,56
22221	0.693	0.67	6,4	0,43
12212	0.691	0.75	6,7	0,5
22112	0.687	0.78	6,8	0,53

EQ-5D state	Dutch national tariff (115)	Probability attending work	Level of functioning attending work	Total productivity
11222	0.687	0.74	6,5	0,48
21122	0.683	0.77	6,6	0,51
22212	0.655	0.7	6,4	0,45
21222	0.651	0.69	6,2	0,43
11311	0.638	0.57	6,6	0,38
12122	0.637	0.76	6,7	0,51
12222	0.605	0.68	6,3	0,43
21311	0.602	0.51	6,4	0,32
22122	0.601	0.71	6,4	0,45
22222	0.569	0.62	6,0	0,37
12311	0.556	0.49	6,4	0,32
11321	0.552	0.48	6,2	0,3
13111	0.543	0.7	7,2	0,5
31111	0.534	0.5	6,6	0,33
22311	0.52	0.43	6,1	0,26
21321	0.516	0.42	6,0	0,25
11312	0.514	0.52	6,2	0,32
13211	0.511	0.61	6,7	0,41
23111	0.507	0.64	6,9	0,44
31211	0.502	0.4	6,2	0,25
13311	0.486	0.29	6,0	0,17
21312	0.478	0.45	5,9	0,27
31311	0.477	0.15	5,6	0,08
23211	0.475	0.55	6,4	0,35
12321	0.47	0.41	6,0	0,24
13121	0.457	0.62	6,7	0,42
32111	0.452	0.42	6,4	0,27
23311	0.45	0.24	5,8	0,14
31121	0.448	0.41	6,2	0,25
22321	0.434	0.34	5,8	0,2
12312	0.432	0.44	6,0	0,26
11322	0.428	0.43	5,8	0,25
13221	0.425	0.53	6,3	0,33
23121	0.421	0.56	6,4	0,36
32211	0.42	0.33	6,0	0,2
13112	0.419	0.65	6,7	0,44
31221	0.416	0.32	5,8	0,19
31112	0.41	0.44	6,2	0,27
13321	0.4	0.22	5,7	0,12
22312	0.396	0.37	5,7	0,21
32311	0.395	0.11	5,4	0,06
21322	0.392	0.36	5,6	0,2
31321	0.391	0.11	5,2	0,06
23221	0.389	0.46	6,0	0,28
13212	0.387	0.56	6,3	0,35
23112	0.383	0.59	6,4	0,38
33111	0.382	0.23	6,0	0,14

EQ-5D state	Dutch national tariff [115]	Probability attending work	Level of functioning attending work	Total productivity
31212	0.378	0.35	5,8	0,2
11113	0.37	0.75	6,9	0,52
11131	0.366	0.58	6,6	0,38
32121	0.366	0.34	6,0	0,2
23321	0.364	0.18	5,4	0,1
13312	0.362	0.24	5,6	0,14
31312	0.353	0.12	5,2	0,06
23212	0.351	0.49	6,0	0,3
33211	0.35	0.17	5,6	0,1
12322	0.346	0.35	5,6	0,2
11213	0.338	0.67	6,5	0,44
21113	0.334	0.7	6,6	0,46
11231	0.334	0.48	6,2	0,3
32221	0.334	0.26	5,6	0,14
13122	0.333	0.57	6,3	0,36
21131	0.33	0.51	6,3	0,32
32112	0.328	0.37	6,0	0,22
23312	0.326	0.2	5,4	0,11
33311	0.325	0.05	5,0	0,02
31122	0.324	0.36	5,8	0,21
11313	0.313	0.34	5,8	0,2
22322	0.31	0.29	5,4	0,16
32321	0.309	0.08	5,0	0,04
11331	0.309	0.19	5,6	0,11
21213	0.302	0.61	6,2	0,38
13222	0.301	0.47	5,9	0,28
21231	0.298	0.41	5,9	0,25
23122	0.297	0.5	6,0	0,3
33121	0.296	0.17	5,6	0,1
32212	0.296	0.28	5,6	0,16
31222	0.292	0.27	5,4	0,15
12113	0.288	0.69	6,7	0,46
11123	0.284	0.68	6,5	0,44
12131	0.284	0.5	6,4	0,32
21313	0.277	0.29	5,6	0,16
13322	0.276	0.18	5,3	0,1
21331	0.273	0.15	5,3	0,08
32312	0.271	0.09	5,0	0,05
31322	0.267	0.09	4,9	0,04
23222	0.265	0.4	5,6	0,23
33221	0.264	0.12	5,3	0,07
33112	0.258	0.19	5,6	0,11
12213	0.256	0.6	6,3	0,37
22113	0.252	0.63	6,4	0,4
11223	0.252	0.59	6,1	0,36
12231	0.252	0.4	6,0	0,24
21123	0.248	0.62	6,2	0,38

EQ-5D state	Dutch national tariff (115)	Probability attending work	Level of functioning attending work	Total productivity
22131	0.248	0.43	6,1	0,26
11132	0.242	0.52	6,2	0,32
32122	0.242	0.29	5,6	0,16
23322	0.24	0.15	5,1	0,07
33321	0.239	0.04	4,7	0,02
12313	0.231	0.28	5,6	0,15
11323	0.227	0.27	5,4	0,15
12331	0.227	0.15	5,4	0,08
33212	0.226	0.14	5,3	0,07
22213	0.22	0.53	6,0	0,32
13113	0.218	0.48	6,3	0,3
21223	0.216	0.52	5,8	0,3
22231	0.216	0.34	5,7	0,2
13131	0.214	0.29	6,0	0,18
11232	0.21	0.42	5,8	0,25
32222	0.21	0.21	5,2	0,11
31113	0.209	0.28	5,8	0,16
21132	0.206	0.46	5,9	0,27
31131	0.205	0.15	5,5	0,08
12123	0.202	0.61	6,3	0,38
33312	0.201	0.04	4,7	0,02
22313	0.195	0.23	5,4	0,12
21323	0.191	0.22	5,2	0,11
22331	0.191	0.12	5,1	0,06
13213	0.186	0.38	5,9	0,23
32322	0.185	0.07	4,7	0,03
11332	0.185	0.16	5,2	0,08
23113	0.182	0.41	6,0	0,25
13231	0.182	0.22	5,6	0,12
23131	0.178	0.24	5,7	0,14
31213	0.177	0.21	5,4	0,11
21232	0.174	0.36	5,5	0,2
31231	0.173	0.11	5,2	0,05
33122	0.172	0.14	5,3	0,08
12223	0.17	0.51	5,9	0,3
22123	0.166	0.54	6,0	0,32
13313	0.161	0.14	5,3	0,07
12132	0.16	0.44	6,0	0,26
13331	0.157	0.07	5,1	0,03
31313	0.152	0.06	4,9	0,03
23213	0.15	0.32	5,6	0,18
21332	0.149	0.13	5,0	0,06
31331	0.148	0.03	4,7	0,01
23231	0.146	0.18	5,4	0,1
12323	0.145	0.21	5,3	0,11
33222	0.14	0.1	4,9	0,05
22223	0.134	0.44	5,6	0,25

EQ-5D state	Dutch national tariff [115]	Probability attending work	Level of functioning attending work	Total productivity
13123	0.132	0.39	5,9	0,23
12232	0.128	0.35	5,6	0,2
32113	0.127	0.22	5,6	0,12
23313	0.125	0.11	5,0	0,05
22132	0.124	0.38	5,7	0,22
31123	0.123	0.21	5,4	0,12
32131	0.123	0.11	5,3	0,06
23331	0.121	0.05	4,8	0,03
33322	0.115	0.03	4,4	0,01
22323	0.109	0.17	5,0	0,09
12332	0.103	0.12	5,0	0,06
13223	0.1	0.3	5,5	0,17
23123	0.096	0.33	5,6	0,19
32213	0.095	0.16	5,2	0,08
22232	0.092	0.29	5,4	0,16
31223	0.091	0.15	5,1	0,08
32231	0.091	0.08	5,0	0,04
13132	0.09	0.25	5,6	0,14
31132	0.081	0.12	5,2	0,06
13323	0.075	0.1	4,9	0,05
32313	0.07	0.05	4,7	0,02
22332	0.067	0.09	4,8	0,05
31323	0.066	0.04	4,5	0,02
32331	0.066	0.02	4,5	0,01
23223	0.064	0.25	5,3	0,13
13232	0.058	0.18	5,3	0,1
33113	0.057	0.11	5,2	0,06
23132	0.054	0.2	5,4	0,11
33131	0.053	0.05	5,0	0,03
31232	0.049	0.09	4,9	0,04
11133	0.041	0.35	5,8	0,2
32123	0.041	0.17	5,2	0,09
23323	0.039	0.08	4,7	0,04
13332	0.033	0.05	4,7	0,03
33213	0.025	0.07	4,9	0,04
31332	0.024	0.02	4,3	0,01
23232	0.022	0.15	5,0	0,07
33231	0.021	0.03	4,7	0,02
11233	0.009	0.26	5,4	0,14
32223	0.009	0.12	4,9	0,06
21133	0.005	0.29	5,5	0,16
33313	-0.000	0.02	4,4	0,01
32132	-0.001	0.09	5,0	0,05
23332	-0.003	0.04	4,5	0,02
33331	-0.004	0.01	4,2	0
32323	-0.016	0.03	4,4	0,01
11333	-0.016	0.08	4,9	0,04

EQ-5D state	Dutch national tariff [115]	Probability attending work	Level of functioning attending work	Total productivity
21233	-0.027	0.22	5,2	0,11
33123	-0.029	0.08	4,9	0,04
32232	-0.033	0.06	4,7	0,03
12133	-0.041	0.28	5,6	0,16
21333	-0.052	0.07	4,6	0,03
32332	-0.058	0.02	4,2	0,01
33223	-0.061	0.05	4,6	0,02
33132	-0.071	0.04	4,7	0,02
12233	-0.073	0.21	5,2	0,11
22133	-0.077	0.23	5,3	0,12
33323	-0.086	0.01	4,1	0,01
12333	-0.098	0.06	4,7	0,03
33232	-0.103	0.03	4,4	0,01
22233	-0.109	0.17	5,0	0,08
13133	-0.111	0.14	5,2	0,07
31133	-0.12	0.06	4,8	0,03
33332	-0.128	0.01	3,9	0
22333	-0.134	0.05	4,5	0,02
13233	-0.143	0.1	4,9	0,05
23133	-0.147	0.11	5,0	0,06
31233	-0.152	0.04	4,5	0,02
13333	-0.168	0.03	4,4	0,01
31333	-0.177	0.01	4,1	0
23233	-0.179	0.08	4,7	0,04
32133	-0.202	0.05	4,7	0,02
23333	-0.204	0.02	4,2	0,01
32233	-0.234	0.03	4,4	0,01
32333	-0.259	0.01	3,9	0
33133	-0.272	0.02	4,4	0,01
33233	-0.304	0.01	4,1	0,01
33333	-0.329	0.00	3,7	0

10.

Discussion

10.1 Introduction

Economic evaluations are increasingly used to inform decision makers about the relative efficiency of new health interventions. Moreover, the role of economic evaluations in reimbursement decisions is increasingly formalized. An important cost category in economic evaluations is that of productivity costs. Productivity costs can strongly affect cost-effectiveness outcomes; however, the proper place of productivity costs in economic evaluations is still disputed. The inclusion of these costs is sometimes seen as unethical, since it may lead to favoring treatments targeted at illnesses affecting the working population over other treatments [(13)]. As described in chapter 2, the debate on in- or excluding productivity costs in an economic evaluation is closely related to the issue of the appropriate perspective to take in economic evaluations and (therefore) the relevant decision context and decision rule. Theoretically, there are compelling arguments to take a societal perspective in economic evaluations (and consequently to include productivity costs). Reimbursement decisions based on other perspectives may lead to inefficient use of resources and hence welfare decreasing decisions, since these decisions are based on incomplete information on societal costs and effects [(35)]. However, in practice health care decision-makers may favor economic evaluations performed from a health care perspective (in which case only costs falling on the health care budget are included) over evaluations performed from a societal perspective. Moreover, decisions based on economic evaluations adopting a societal perspective must be sensitive to the responsibility and budgetary constraints of health care decision makers. For instance, if the costs of some intervention fall inside the health care sector, but savings fall outside the health care sector, the intervention may be cost-effective from a societal perspective, but hard to implement for the health care decision maker. In theory, such savings outside the health care sector could then be (partially) transferred into the health care budget in order to prevent budgetary problems. Such transfers, however, may be difficult to realize in practice. [(33)] Moreover, the opportunity costs within the health care sector may be disproportionally high when high costs fall on the budget. The lack of consensus regarding the appropriate perspective and the inclusion of productivity costs in economic evaluations is reflected in variations in the recommendations in national health economic guidelines regarding these matters.

Next to debates on whether to include productivity costs, debates have focused on how productivity costs should actually be included [(2,15-18,21,39,47,48,56,90)].

As described in chapter 2, many issues on how to include productivity costs in economic evaluations have not yet been resolved. Moreover, new important questions have emerged from these earlier debates, of which five have been addressed in this thesis.

10.2 Questions addressed in this thesis

1. Do respondents to health state valuations include the effects of ill-health on income and if so, how does such inclusion affect valuations?

The recommendation of the United States Panel on Cost Effectiveness in Health and Medicine [(2)] to abandon valuing productivity costs in monetary terms was based on the hypothesis that these costs would already be captured in QALYs. Their recommendation received much criticism [(8,22,31)]. Indeed, it seems unlikely that respondents to health state valuations would be able to include all societal effects of productivity changes. If respondents would consider the effects of ill-health on income, they would probably focus on the effects on their net income. Moreover, the link between productivity and income on an individual level may be weak because of private and social security systems. [(22,31)] In spite of the theoretical criticisms of the US Panel approach, the suggestion of double counting resulting from some (partial) valuation of productivity costs in terms of QALYs when also applying the human capital or friction cost approach to value productivity losses could not be easily rejected, since it was unknown whether respondents in health state valuations include the effects of ill-health on income during the valuation process and, if so, how such considerations would affect health state valuations.

Chapters 3, 4 and 5 of this thesis showed that, in our studies, around half of the respondents to health state valuations spontaneously considered the potential effects of ill-health on income during health state valuation exercises. These outcomes indicate that if productivity costs are included, as commonly done, on the cost side of the cost-effectiveness ratio, this may indeed lead to double counting. However, it turned out that spontaneous inclusion or exclusion of income effects had only a limited effect on subsequent health state values. In other words, the QALY measure does not seem to be sensitive enough to capture the effects of ill-health and treatment on productivity and the 'pollution' created by double counting is small (if not non-existent). Therefore, it seems advisable to measure productivity costs in monetary terms on the cost side of the cost-effectiveness ratio where these costs do have a noticeable effect, rather than to try and capture (a part of) these costs on the effect side. Given the evidence regarding at best very small effects of taking the effects of ill-health on income into consideration, it does not seem to be necessary to explicitly instruct respondents to health state valuation to ignore potential income-effects and double counting does not appear to pose a threat in practice.

2. How common is productivity cost inclusion in economic evaluations?

Though the inclusion of productivity costs has been extensively debated in the literature over the last two decades, only little information was available on how often productivity costs are actually included in economic evaluations. Moreover, not much was known about the effects of inclusion or exclusion of productivity costs on cost-effectiveness outcomes. Such information is, however, critical in providing insight in the potential effects of the choice to include or exclude productivity costs. The research described in chapters 6 and 7 of this thesis indicated that in most published economic evaluations productivity costs are, in fact, ignored. Inclusion of productivity costs was more common in economic evaluations of therapies for depressive disorders targeted at patients in the working age groups, than in economic evaluations of treatments with expensive drugs administered in a hospital setting. In the former, just under one-third of economic evaluations included productivity costs and in the latter these cost were included in less than ten percent of the economic evaluations. The fact that inclusion of productivity costs was more common in the field of depressive disorders than in expensive drug treatments probably is partly explained by differences in the age and the severity of illness of the patient populations. Nevertheless, productivity costs in both fields were less frequently included than expected, indicating that productivity costs may regularly be 'unjustly' excluded, reasoning from the premise that decision makers should be informed about all relevant societal costs and benefits of an intervention.

It is also worth to note that sometimes productivity costs were included in economic evaluations when this was not expected, given the fact that the relevant national health economic guidelines prescribed adopting a health care perspective. Such unexpected inclusion in the studies reviewed in chapter 7 in all cases led to more favorable cost-effectiveness outcomes for the new drug. While this thesis cannot draw any firm conclusions regarding this issue, this finding may indicate the existence of some kind of 'perspective bias'; i.e. productivity costs may sometimes be included based on the (desired) expected effects on incremental cost-effectiveness outcomes. It is worth to investigate this in future studies.

3. How does productivity cost inclusion or exclusion affect cost-effectiveness outcomes?

As shown in chapters 6 and 7, the inclusion or exclusion of productivity costs can influence both incremental costs and incremental cost-effectiveness relatively strongly. Mostly, inclusion of productivity costs was to the advantage of the 'new' intervention rather than the alternative intervention in terms of incremental cost-effectiveness. Notably, the precise impact of productivity cost inclusion differed strongly between studies, both expressed as a fraction of total costs and as the effect on incremental cost (-effectiveness) outcomes. It was not possible to gain full insight in the nature of these differences, since in most studies the level of detail of underlying methods and figures was insufficient to allow further investigation. Besides, the limited information that *was* presented in the studies suggested that there was much variation in applied methods regarding the identification, measurement and valuation of productivity costs.

The strong potential impact of productivity costs on cost-effectiveness outcomes emphasizes that in countries where a health care perspective is prescribed (i.e. where productivity costs are structurally ignored) reimbursement decisions based on cost-effectiveness studies adhering to this perspective may lead to inefficient use of societal resources.

4. How do compensation mechanisms and multiplier effects potentially affect productivity cost estimates?

The infrequent inclusion of productivity costs in economic evaluations may be influenced by uncertainty regarding the accuracy of productivity cost estimates. Indeed, it has been questioned whether such estimates accurately reflect 'true' productivity costs. On the one hand, it has been suggested that common productivity costs estimates are, in fact, overestimations [17,18]. If absenteeism and presenteeism are compensated either by the ill employee herself or her colleagues during normal working hours or in extra working time, this may lower production losses. While these compensation mechanisms need not be costless, they may result in lower costs than when production would have been lost.

On the other hand, it has been advocated that common estimates of productivity costs are, in fact, underestimations [15,16]. In many professions team dependencies exist. If one of the team members falls ill, or is not able to function properly, the productive output of the whole team may be negatively affected. Thus production losses would be higher than expected by considering only the directly affected individual.

In chapter 8 of this thesis, results were presented of a randomized trial on smoking cessation treatments. Patients in this trial were questioned on how their work was compensated in cases of absenteeism due to ill-health. Based on their responses and previously published profession dependent effects of ill-health on coworkers (referred to as multipliers) productivity costs were calculated using multiple scenarios to explore the potential impact of compensation mechanisms and multiplier effects.

The results indicate that when simultaneously accounting for multiplier effects and compensation mechanisms, the amplifying effect on regular productivity cost estimates of applying multipliers is outweighed by the diminishing effect of adjusting for compensation mechanisms, assuming – naively – that compensation in normal hours is costless. However, in reality it is unlikely that compensation in normal hours is, in an economic sense, costless. Compensation in normal working hours implies that companies may have some type of internal labor reserve, and hence that they are not functioning at full capacity, which would result in economic costs. If such labor reserve is not present, compensation during normal working hours would result in putting extra non-costless strain on employees, potentially even leading to additional (work-related) ill-health.

5. Can productivity losses be predicted based on quality of life estimates?

Next to the controversies and debates regarding productivity costs inclusion and the lack of standardization, the neglect of productivity cost inclusion in economic evaluations may to some extent be explained by practical limitations. Researchers do not always have the opportunity to collect data on the effects of ill-health and treatment on patients' productivity. Sometimes only retrospective data is available and in other cases resources maybe insufficient to collect additional data. To date, if such problems regarding data collection emerge, productivity costs are likely to be simply ignored. However, previous research has indicated a relationship between quality of life and productivity [19], leading to the question whether it would be possible to estimate productivity changes based on quality of life data.

In Chapter 9 the results are presented of a study exploring the possibility to predict productivity levels in terms of absenteeism and diminished productivity while being at work based on quality of life information described with the EuroQol descriptive system (EQ-5D, 3-level version [24]). In order to do so, an online survey was conducted in which respondents were questioned on their expected productivity in numerous EQ-5D health states. Subsequently, based on the responses a prediction model was constructed for absenteeism and presenteeism. The validity of the models was tested among a group of low back pain patients. The results of the validation study were promising, although the model predicting presenteeism seemed to overestimate productivity losses somewhat. Nevertheless, applying prediction models to estimate productivity losses in cases where no actual data on productivity is available may be considered preferable over the alternative of ignoring the effect of ill-health on productivity altogether.

10.3 Limitations

Although this thesis has addressed some important issues regarding productivity cost inclusion in economic evaluations, other equally important issues have been neglected. Moreover, the theoretical and practical choices made during the process of constructing this thesis may be criticized.

For instance, all questionnaires used in the research on income considerations in health state valuation exercises presented in chapters 3, 4 and 5 were designed for self-completion and most questions were close-ended. Although this approach was beneficial in terms of sample size, it provided little information on the actual thought process of respondents during the health state valuation exercises. For this purpose, it would have been useful if the research had included a more qualitative part including semi-structured face-to-face interviews (e.g. such as done before with the Visual Analogue Scale [271]).

The research presented in chapter 6 and 7 focused on productivity cost inclusion in economic evaluations in specific disease areas (i.e., depressive disorders and disorders requiring expensive intramurally administrated drugs). The (lack of) productivity cost inclusion and the strong impact when included on cost-effectiveness outcomes may differ in other areas. In addition, the impact of productivity cost inclusion could only be determined based on outcomes of the few studies that had included productivity costs. Especially if in- and exclusion (to some extent) is the result of a perspective-selection-bias, the impact of productivity costs on cost-effectiveness described in this thesis may be a poor representation of its 'true' impact.

A limitation of the research presented in chapter 8 is that it has focused merely on compensation of diminished productivity in case of absenteeism. It would be worthwhile to further investigate how diminished productivity while being at work is compensated. Furthermore, the multipliers applied in this research were based on previous published estimates based on a study conducted in the United States. These multipliers may not adequately resemble potential effects of ill-health on coworkers in the Netherlands. Research in the line of Nicholson and Pauly and colleagues (15,16) on the effects of ill-health on coworkers' productivity among Dutch workers may provide more insight in this matter. Moreover, it is important to further explore how diminished productivity is compensated and especially the costs of such compensation mechanisms need further investigation.

In chapter 9 the constructed prediction models were based on respondents' hypothetical productivity in distinct EQ-5D health states, rather than actual productivity. To an extent this limitation may explain why the model predicting presenteeism seemed to overestimate productivity losses. Respondents who are inexperienced in the presented ill health states may have the tendency to overestimate the effects of ill-health on their productivity. Individuals actually experiencing the health problems may adjust to their limitations. Moreover, individuals in ill health may have to choose professions in which their health problems have a limited effect on their working abilities, particularly if these health problems are related to chronic conditions.

Next to the limitations of the research presented in this thesis, another limitation is that some important aspects regarding the inclusion of productivity costs in economic evaluations have not received the attention these aspects would deserve. For instance, most of the research presented in this thesis has limited its focus on productivity costs related to the effects of ill-health and treatment on paid work. In doing so, productivity costs related to unpaid professions remain as underexplored as before. Important questions to address in future research in the area of productivity costs related to unpaid work are: how can unpaid productivity best be identified and measured to be included in economic evaluations and which valuation approaches are most appropriate?

Furthermore, in this thesis relatively little attention has been paid to the question whether productivity costs should be included in economic evaluations. Although in this thesis it is advocated that the societal perspective is the most appropriate perspective to take in economic evaluations and this view is widespread (2,7,34,35,40,41), in practice many national health economic guidelines prescribe a health care perspective (at least for the reference case) (124,233,272).

In that context, an underlying underexplored issue is the question of how to deal with productivity costs in health care decision making. The role and place of productivity costs in actual health care decision making could be explored further. It could be investigated how, when and to what extent decision makers (wish to) include productivity costs in their decisions. It may well be that such costs receive different weight in health care decisions, than, for instance, health care costs. Exploring the practical possibilities for redirecting savings outside the health care system, occurring due to health care interventions financed from the health care budget, back into the health care budget is also interesting. Note that this does not imply that health care policy makers should only take these costs into account when this is possible. Such research may contribute to the awareness of the importance of productivity costs, stimulate their inclusion in economic evaluations and decision making, and, consequently, may help to improve overall welfare.

10.4 Implications for policy and future research

This thesis aims to contribute to the ongoing (predominantly methodological) debates in the area of productivity costs measurement and valuation. Moreover, the outcomes of the research in this thesis may have some implications for health policy and future research.

To start, the results of the empirical work presented in this thesis indicate that the QALY does not seem capable to fully capture the impact of ill-health and treatment on productivity and income. This finding invalidates the recommendation of the United States Panel to include productivity costs on the effect side of the cost-effectiveness ratio. Nevertheless, since approximately half of the respondents in health state valuation exercises do seem to consider the effects of ill-health on income, one may wonder which other non-health related effects respondents are including in their valuation process. It seems clear that respondents include more than anticipated, considering the QALY measure is normally expected to capture health related quality of life rather than broader notions of quality of life.

Another important finding is that the decision to include or exclude productivity costs in economic evaluations can have a very large effect on outcomes. Decision makers should be aware of the strong potential impact on incremental cost-effectiveness. If a societal perspective is prescribed, decision makers may have to put more emphasis on the actual and systematic inclusion of productivity costs. Although in particular circumstances it may prove to be impossible or irrelevant to include productivity measurement in economic evaluations, exclusion of productivity costs should always be explicitly justified. Otherwise, inclusion or exclusion because of strategic reasons rather than because of theoretical considerations may be the consequence. Moreover, in situations where inclusion is difficult for practical reasons, estimates based on productivity predictions may be more appropriate than simply ignoring the effects of ill-health and treatment on productivity. Currently, productivity costs seem to be often excluded without sufficient argumentation. For decision makers it is important to understand that reimbursement decisions based on economic evaluations ignoring actual costs (or savings) risk an inefficient use of societal resources.

Comparing results of economic evaluations including different aspects on the cost side of the cost-effectiveness ratio comes dangerously close to comparing apples to oranges. Where (for the sake of comparability) most national health economic guidelines are uniformly prescribing the QALY as preferred outcome measure[10], the same guidelines are less consistent and explicit in their guidance on cost-inclusion. This is probably strongly related to issues discussed in this thesis, such as the lack of consensus on appropriate perspective, on methods to measure and value productivity costs, the role and influence of compensation mechanisms and multiplier effects, and so on. More research and debate, working towards a further standardization of productivity costs identification, measurement and valuation, remains crucial therefore.

10.5 Concluding remarks

Despite the scientific research conducted in the area of productivity costs in economic evaluations over the years, consensus has not been reached on the issues of whether and how to include productivity costs in economic evaluations. This lack of consensus is reflected in the diversity in national health economic guidelines and has likely contributed to productivity costs being often ignored in economic evaluations. This thesis has provided more insight in the potential impact of productivity costs on cost-effectiveness outcomes. It seems clear that wrongful exclusion of productivity costs could lead to wel-

10 To note, the use of use the QALY measure in economic evaluations is not free of dispute. Several problems with the QALY have been discussed in the literature (see for instance Lipscomb et al. and Nord et al. [275,276]).

fare damaging decisions, since decisions regarding reimbursement of health interventions may then be based on inaccurate information on incremental cost-effectiveness ratios.

Considering the adoption of the societal perspective as most appropriate, it is pivotal to increase productivity cost inclusion in economic evaluations. In that context, it is important that decision makers are more aware of the strong potential impact of productivity costs on cost-effectiveness outcomes and that they put more emphasis on adhering to the prescribed perspective. Moreover, the lack of standardization of productivity cost methodology is a serious concern. In this thesis it is shown that in daily health economic practice a variety of methods is used to identify, measure and value productivity costs. Since the different methods can lead to a large difference in productivity cost estimates, the trustworthiness of outcomes inevitably may become a matter of debate. Moreover, the diversity in methodology seriously hampers the possibilities to compare outcomes between studies. This thesis has emphasized that, standardization is crucial for productivity cost inclusion to become standard practice instead of an exception.

This thesis has further increased to the theoretical and practical knowledge regarding productivity cost inclusion in economic evaluations. Although, given the findings of this thesis, achieving full consensus regarding the appropriateness of inclusion of productivity costs and best methods for identifying as well as measuring and valuing productivity costs may be challenging, it remains an ultimate goal. Difficulties in reaching consensus may partly be due to differences in basic assumptions regarding the context and goal of health care decision making. Especially such differences may prove to be difficult to bridge.

Moreover, in jurisdictions wishing to take a societal perspective and hence include productivity costs, consensus on how to include these costs may not be easily reached. Again there, normative judgments and schools of thought may lead to fundamental differences in opinion on how to measure and value productivity costs (and other costs and effects for that matter) [273]. The debates regarding the friction cost method versus the human capital method, and the underlying differences in schools of thought, may illustrate this. An intermediate solution in this matter would be calculating productivity costs with both approaches, which, as shown in chapters 6 and 7, is commonly done in studies applying the friction cost approach. Additionally, it is critical to standardize *how* both approaches are applied best, as variation within methods may be large as well.

Some form of consensus on main productivity cost issues (and *at least* on standards for transparency regarding applied methods and presentation of outcomes) could contribute to an increase in productivity costs inclusion in economic evaluations and to a higher level of quality and comparability of productivity cost estimates. Such a consensus, for instance laid down in the form of a consensus document, is already available for other (maybe less controversial) topics in the field of health technology assessment [(274)]. For consensus regarding productivity cost calculation to be successful in increasing comparability in productivity cost outcomes, recommendations need to be made on a relatively detailed level. Some of the issues that need to be standardized are: which productivity costs to include (e.g. paid and unpaid labor, absenteeism and presenteeism, effects on coworkers' productivity, costs and savings due to compensation mechanisms, etc.), which of the available instruments are appropriate to measure lost productivity, which values to attach to lost production units (e.g. a fraction of GDP, patients' income, workers' added value, or average incomes). It seems worthwhile attempting to take first steps on the road to standardization of productivity cost methodology.

The above remarks may illustrate that while productivity costs are obviously important, both societally and in the context of economic evaluation, their sound consideration in health care decision making remains an important objective. Hopefully this thesis ultimately contributes to that goal.

Summary

Economic evaluations are increasingly used to inform decision makers about the relative efficiency of new health interventions. While the influence of this information on subsequent decisions may still be limited [3-5], more and more countries appear to wish to include this information in the process of reimbursement decisions [6]. The increased use of economic evaluations in health care also emphasizes the need to ensure a sound methodology of these evaluations. Obviously, when the influence of outcomes of economic evaluations on subsequent decision making becomes stronger, the need for sound methodology becomes weightier as well. In that context it must be stressed that numerous methodological debates regarding how to best perform economic evaluations are ongoing. One of the areas of debate concerns (the inclusion of) productivity costs, which can be defined as *"Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid."* [8]. Productivity costs can strongly affect cost-effectiveness outcomes; however, the proper place of productivity costs in economic evaluations is still disputed.

This thesis aimed to answer five questions which emerged from previous research and debates concerning productivity costs. These five questions relate to the inclusion of productivity costs and the methodology to estimate these costs in economic evaluations of health care interventions:

1. Do respondents to health state valuations include the effects of ill-health on income and if so, how does such inclusion affect valuations?
2. How common is productivity cost inclusion in economic evaluations?
3. How does productivity cost inclusion or exclusion affect cost-effectiveness outcomes?
4. How is reduced productivity related to ill-health compensated and how do compensation mechanisms and multiplier effects potentially affect productivity cost estimates?
5. Can productivity losses be predicted based on quality of life estimates?

In chapter 2 of this thesis an overview was presented of the most important past and present debates and developments regarding productivity cost inclusion, identification, measurement and valuation. Reviewing the published literature, demonstrated that despite the progress and the substantial amount of scientific research conducted, consensus has not been reached on either the inclusion of productivity costs in economic evaluations or the methods used to produce productivity cost estimates. Such a lack of consensus has likely contributed to ignoring productivity costs in economic evaluations and is reflected in variations in national health economic guidelines. Given the current variety in applied productivity cost methodology, accurately reporting the methods used to estimate productivity costs in economic evaluations is critical to understand the nature of potential differences in outcomes. However, in published studies the level of detail provided regarding the exact methods used is often poor. Based on the studied literature several recommendations were made regarding future research topics to address, aiming to result in an increase in productivity cost inclusion in economic evaluations and in further development of productivity cost methodology.

The empirical work in chapters 3, 4 and 5 was aimed at answering research question 1. It was investigated whether respondents to health state valuations take the effects of ill-health on income into account when participating in health state valuation exercises. Additionally, it was investigated how spontaneous inclusion or exclusion of income-effects influences health state valuations. Moreover, the effects of explicitly instructing respondents to either include or exclude income-effects were explored. The research in chapters 3, 4 and 5 was initiated by the recommendation of the United States Panel on Cost Effectiveness in Health and Medicine in 1996 (2) to abandon valuing productivity costs, as it was commonly done, in monetary terms. This recommendation was based on the hypothesis that these costs would already be captured in quality adjusted life years (QALYs) and, therefore, measuring productivity costs additionally on the cost side of the cost-effectiveness ratio would lead to double counting. Although the panel's recommendation received much criticism (8,22,31) and it was argued that valuing productivity costs in QALYs would not lead to sound estimates, the suggestion of double counting could not be easily rejected, since it was unknown whether respondents in health state valuations include the effects of ill-health on income during the valuation process and, if so, how such considerations would affect these valuations.

Approximately half of the respondents in the empirical studies described in chapters 3, 4 and 5 spontaneously included the potential effects of ill-health on income during health state valuation exercises. When income-effects were included, it had little to no effect on health state valuations. Explicitly instructing respondents to exclude or include income-effects, likewise, had a limited effect on outcomes. Therefore, the QALY measure does not seem to be sensitive enough to capture the effects of ill-health and treatment on productivity and the 'pollution' created by double counting is small (if not non-existing). Therefore, it seems advisable to capture productivity costs in terms of monetary terms on the cost side.

In the studies described in chapters 6 and 7 of this thesis, aimed at answering research questions 2 and 3, it was investigated to what extent productivity costs are included in economic evaluations in order to increase insight in the current role of productivity costs in economic evaluations. Next, in evaluations where these costs were included, it was examined how these costs were identified, measured and valued and the impact of including productivity costs on cost-effectiveness outcomes was assessed. The findings illustrated that in most published economic evaluations productivity costs are not included. Inclusion of productivity costs was more common in economic evaluations of therapies for depressive disorders targeted at patients of working age, than in economic evaluations of treatments with expensive drugs administered in a hospital setting. Just under one-third of the former economic evaluations included productivity costs, while this percentage was less than ten of the latter economic evaluations. The fact that inclusion of productivity costs was more common in the field of depressive disorders than in expensive drug treatments is, to some extent, explained by differences in age and severity of illness of the patient populations. Nevertheless, productivity costs in both fields were less frequently included than expected, indicating that productivity costs may regularly be 'unjustly' excluded, reasoning from the premise that decision makers should be informed about all relevant societal costs and benefits of an intervention.

The results of the research described in chapters 6 and 7 also illustrated that the choice to include or exclude productivity costs can influence both incremental costs and incremental cost-effectiveness relatively strongly. Mostly, inclusion of productivity costs was to the advantage of the 'new' intervention rather than the comparator in terms of incremental cost-effectiveness. The precise impact of productivity cost inclusion differed strongly between studies, both expressed as a fraction of total costs and as the effect on incremental cost (-effectiveness) outcomes. It was not possible to gain full insight in the nature of these differences, since in most studies the level of detail in reporting underlying methods and figures was insufficient to allow further investigation. Besides, the limited information that *was* presented in the studies suggested that there was much variation in applied methods regarding the identification, measurement and valuation of productivity costs.

In chapter 8, related to research question 4, the influence of compensation mechanisms and multiplier effects on productivity costs estimates was investigated. In the literature, on the one hand, it has been suggested that common productivity cost estimates may overestimate actual productivity costs if absenteeism and presenteeism are (partly) compensated either by the ill employee herself or by her colleagues during normal working hours [17,18]. However, on the other hand, it has been advocated that common estimates of productivity costs are, in fact, underestimations of actual productivity costs [15,16], since in many professions team dependencies exist. Then, if one of the team members falls ill, the productive output of the whole team may be negatively affected. The empirical study described in chapter 8 investigated how ill workers' absenteeism is compensated. Additionally, it was explored how compensation mechanisms and effects of ill-health on team workers potentially affect common productivity cost estimates. Over half of the respondents indicated that their work was compensated during normal working hours by themselves or their colleagues. When simultaneously accounting for effects of ill-health on team workers' productivity and compensation mechanisms, the amplifying effect on regular estimates of accounting for team effects is outweighed by the diminishing effect of adjusting for compensation mechanisms. Note that this was done under the naive assumption that compensation in normal hours is costless. However, in reality it is unlikely that compensation in normal hours is, in an economic sense, costless.

As chapters 6 and 7 pointed out, productivity costs are often ignored in economic evaluations. This may have to do with practical difficulties in obtaining reliable estimates of productivity costs. In such cases, it has been suggested to estimate productivity costs indirectly using quality of life data. The study presented in chapter 9, answering research question 5, aimed to derive and validate an algorithm for predicting productivity on the basis of quality of life data. In order to do so, a large representative sample of the Dutch general public was asked in a web-based questionnaire to state their expected level of productivity in terms of absenteeism and presenteeism for multiple EQ-5D health states. Based on these data, two prediction models were constructed: i) a model predicting levels of absenteeism and ii) a model predicting presenteeism. The models were validated by comparing model predictions with conventionally measured productivity within a group of low back pain patients. Predicted absenteeism levels using EQ-5D health states, closely resembled conventionally measured absenteeism levels. Productivity losses related to presenteeism seemed somewhat overestimated by the prediction model. Measured and predicted productivity were moderately but highly significantly correlated. Overall, it appeared possible to make reasonable productivity predictions based on EQ-5D data. Further exploration and validation of prediction algorithms remains necessary, however, especially for presenteeism.

In chapter 10 the main findings of the research presented in this thesis were discussed. Moreover, the implications and the limitations of this thesis were addressed. Based on the findings (and referring back to the five main questions this thesis aimed to answer), some conclusions could be drawn. First, about half of the respondents to health state valuations seem to include income-effects. This seems to have little effect on subsequent valuations. Productivity costs therefore are preferably placed on the cost side of the cost-effectiveness ratio. Second, despite the scientific attention regarding productivity cost issues, productivity costs are ignored in the majority of economic evaluations, often without clear justification. Third, when productivity costs are included this can have a very strong impact on cost-effectiveness outcomes. The strong potential impact of productivity costs on cost-effectiveness outcomes emphasizes that when productivity costs are not included in economic evaluations, reimbursement decisions based on the results, may lead to inefficient use of societal resources. Fourth, lost production is commonly compensated during normal working hours. Productivity costs estimates are strongly influenced by adjustment for compensation mechanisms and multiplier effects. The validity of such adjustments needs further examination, however. Fifth, although the validity of predicting productivity losses based on quality of life data needs further investigation, applying prediction models in cases where no actual data on productivity is available seems preferable over ignoring the effect of ill-health on productivity altogether.

This thesis ends with some final remarks regarding the inclusion of productivity costs in economic evaluations and the methodology to estimate these costs. Consensus has not yet been reached on whether and how to include productivity costs in economic evaluations. Consequently, productivity costs are often ignored in economic evaluations. Considering the adoption of the societal perspective as most appropriate, it is pivotal to increase productivity cost inclusion in economic evaluations, especially, given the strong potential impact of productivity costs on cost-effectiveness outcomes. In that context, it is important that decision makers are more aware of this potential impact and that they put more emphasis on adhering to the prescribed perspective. Moreover, the lack of standardization of productivity cost methodology is a serious concern. In this thesis it is shown that in actual economic evaluations a variety of methods is used to identify, measure and value productivity costs. Since applying different methods can lead to large variations in productivity cost estimates, the trustworthiness of outcomes may become a matter of debate. Moreover, this diversity in methodology seriously hampers the possibilities to compare outcomes between studies. This thesis has emphasized that, standardization is crucial for productivity cost inclusion to become standard practice instead of an exception.

Concluding, this thesis has further increased the theoretical and practical knowledge regarding productivity cost inclusion in economic evaluations. Although, given the findings of this thesis, achieving full consensus regarding the appropriateness of inclusion of productivity costs and best methods for identifying as well as measuring and valuing productivity costs may be challenging, it remains an ultimate goal.

Samenvatting

Economische evaluaties worden steeds vaker gebruikt om besluitvormers te informeren over de relatieve efficiëntie van nieuwe gezondheidsinterventies. Hoewel de invloed van informatie over de kosteneffectiviteit op de uiteindelijke beslissingen nog beperkt lijkt te zijn [3-5], zijn er steeds meer landen die dit soort informatie meenemen in vergoedingsbeslissingen. [6] Het is daarom belangrijk dat de methoden en technieken die gebruikt worden in economische evaluaties betrouwbaar zijn. Het belang van goed onderbouwde methodologie gaat vanzelfsprekend zwaarder wegen naar mate de invloed van economische evaluaties op besluitvorming groter wordt.

In de wetenschappelijke literatuur zijn diverse debatten gaande over hoe economische evaluaties het beste kunnen worden uitgevoerd. Een van de onderwerpen van discussie is (de inclusie van) productiviteitskosten. Productiviteitskosten zijn de kosten die samenhangen met productiviteitsverliezen en het vervangen van medewerkers in geval van ziekte, arbeidsongeschiktheid en sterfte van betaald en onbetaald productieve personen. [8]

Productiviteitskosten kunnen de kosteneffectiviteit sterk beïnvloeden, maar de juiste plaats van deze kosten in economische evaluaties blijft een punt van aandacht en debat. Dit proefschrift heeft vijf onderzoeksvragen proberen te beantwoorden die zijn afgeleid van eerdere onderzoeken en debatten. Deze vijf vragen hangen samen met de inclusie van productiviteitskosten in economische evaluaties en de methodologie om deze kosten te schatten:

1. Denken respondenten tijdens het waarderen van gezondheidstoestanden aan de effecten van ziekte op hun inkomen; en zo ja, hoe beïnvloedt dit die waarderingen?
2. Hoe gebruikelijk is de inclusie van productiviteitskosten in economische evaluaties?
3. Hoe beïnvloedt het wel of niet meenemen van productiviteitskosten kosteneffectiviteitsuitkomsten?
4. Hoe wordt ‘verminderde productiviteit door ziekte’ gecompenseerd wat is de mogelijke impact van het rekening houden met compensatiemechanismen en effecten van ziekte op teamproductie op schattingen van productiviteitskosten?
5. Kunnen productiviteitsverliezen geschat worden op basis van gegevens over kwaliteit van leven?

In hoofdstuk 2 van dit proefschrift werd een overzicht gegeven van belangrijke debatten uit het verleden en het heden over productiviteitskosten. Ook werden de ontwikkelingen in termen van inclusie, identificatie, meting en waardering van productiviteitskosten samengevat. Het bestuderen van de gepubliceerde literatuur heeft duidelijk gemaakt dat er, ondanks de ontwikkelingen en de nodige aandacht voor productiviteitskosten, nog geen consensus bestaat over de inclusie van deze kosten noch de methoden om deze kosten te schatten. Dit gebrek aan consensus heeft er waarschijnlijk aan bijgedragen dat productiviteitskosten vaak genegeerd worden en hun inclusie soms ook in nationale gezondheidseconomische richtlijnen expliciet wordt ontmoedigd.

Gezien de grote variatie in de gebruikte productiviteitskostenmethodologie, is het belangrijk dat de gebruikte methoden nauwkeurig gerapporteerd worden. Alleen dan is het mogelijk om verschillen in uitkomsten te begrijpen. Het detailniveau in rapportage van gehanteerde methoden is in gepubliceerde studies vaak onvoldoende. Op basis van de bestudeerde literatuur is er in hoofdstuk 2 een aantal aanbevelingen gedaan met betrekking tot toekomstig onderzoek, met als doel het bevorderen van de inclusie van productiviteitskosten in economische evaluatie en het verder ontwikkelen van productiviteitskostenmethodologie.

Het empirische werk in hoofdstuk 3, 4 en 5 was gericht op het beantwoorden van onderzoeksvraag 1. Er werd onderzocht of respondenten van methoden om gezondheidstoestanden te waarderen de effecten van ziekte op hun inkomen meenemen wanneer ze deelnemen aan zulke waarderingsexercities. Er werd daarbij gekeken hoe het wel of niet denken aan inkomen de waarderingen beïnvloedt. Als laatste werd bestudeerd wat de effecten zijn van het expliciet instrueren van respondenten om inkomenseffecten wel of niet mee te nemen.

Het onderzoek in hoofdstuk 3, 4 en 5 komt voort uit de aanbeveling van 'the United States Panel on Cost Effectiveness in Health and Medicine' uit 1996 (2) om productiviteitskosten niet meer te waarderen in termen van geld, maar in termen van hun effect op de kwaliteit van leven. Deze aanbeveling was gebaseerd op de hypothese dat deze kosten al in de QALY[11] waarderingen zouden zitten. Het ook nog meenemen van productiviteitskosten aan de kostenkant van de kosteneffectiviteitsratio zou hierdoor tot dubbeltellingen leiden. Deze aanbeveling leidde tot veel kritiek (8,22,31). Er werd beargumenteerd dat het waarderen van productiviteitskosten in QALYs niet zou leiden tot betrouwbare resultaten. Desalniettemin kon de suggestie dat waarderingen in geld, naast het gebruik van QALYs, zouden leiden tot dubbeltellingen niet makkelijk weerlegd worden, omdat het niet bekend was of respondenten van gezondheidstoestandwaarderingen inkomenseffecten meenemen in het waarderingsproces en hoe dit deze waarderingen eventueel zou beïnvloeden.

11 QALY staat voor quality adjusted life year, ofwel een voor kwaliteit gecorrigeerd levensjaar.

In de empirische studies beschreven in hoofdstuk 3, 4 en 5 bleek ongeveer de helft van de respondenten spontaan aan inkomenseffecten te denken tijdens het waarderen van diverse gezondheidstoestanden. Deze inclusie van inkomenseffecten had echter weinig tot geen effect op de uiteindelijke waarderingen. Het expliciet instrueren van respondenten om deze effecten wel of niet mee te nemen had eveneens weinig effect op waarderingen. De QALY-maat lijkt dus weinig gevoelig voor het wel of niet meenemen van inkomenseffecten en van dubbeltellingen is dus (nagenoeg) geen sprake. Het lijkt daarom aan te raden om productiviteitskosten mee te nemen aan de kostenkant van de kosteneffectiviteitsratio.

In de studies die beschreven staan in hoofdstuk 6 en 7 (gericht op het beantwoorden van vraag 2 en 3) werd onderzocht in hoeverre productiviteitskosten momenteel worden meegenomen in economische evaluaties. Dit werd gedaan om meer zicht te krijgen op de huidige rol van productiviteitskosten in economische evaluaties. Vervolgens is, voor de evaluaties waarin productiviteitskosten werden meegenomen, onderzocht hoe deze kosten werden meegenomen, gemeten en gewaardeerd. Daarnaast is de impact van productiviteitskosten op de kosteneffectiviteitsuitkomsten berekend.

In het merendeel van de economische evaluaties bleken productiviteitskosten niet te zijn meegenomen. Inclusie van productiviteitskosten kwam vaker voor in economische evaluaties van interventies gericht op depressieve patiënten in de werkende leeftijd dan in economische evaluaties van dure geneesmiddelen toegepast in een ziekenhuissetting. In de eerste groep werden productiviteitskosten in minder dan een derde van de evaluaties meegenomen en in de laatste groep in minder dan tien procent. Dat productiviteitskosten vaker werden meegenomen in evaluaties van interventies gericht op depressie dan evaluaties van dure ziekenhuisgeneesmiddelen, werd gedeeltelijk verklaard door verschillen in leeftijd en ernst van ziekte, maar in beide groepen was de inclusie lager dan verwacht mocht worden. Dit impliceert dat productiviteitskosten regelmatig ‘onterecht’ genegeerd worden, beredeneerd vanuit de gedachte dat dat besluitvormers geïnformeerd dienen te worden over alle relevante maatschappelijk kosten en baten van een interventie.

De beschreven resultaten in hoofdstuk 6 en 7 laten bovendien zien dat de keuze om productiviteitskosten wel of niet mee te nemen een groot effect kan hebben op zowel de incrementele kosten als de kosteneffectiviteit. In termen van incrementele kosteneffectiviteit was de inclusie van productiviteitskosten, in de meeste gevallen, in het voordeel van de ‘nieuwe interventie’ in plaats van de vergelijkende behandeling. De precieze impact van productiviteitskosten verschilde sterk tussen studies, zowel uitgedrukt als een fractie van de totale kosten als het effect op de incrementele kosten(effectiviteit). Het was niet mogelijk om te bepalen hoe deze verschillen precies tot stand kwamen, omdat in de meeste studies te weinig details werden verschaft over de onderliggende methoden en cijfers om dit verder te kunnen onderzoeken. De informatie die wel werd verschaft liet bovendien zien dat er veel variatie bestaat in de toegepaste methoden op het vlak van identificatie, meten en waarderen van productiviteitskosten.

In hoofdstuk 8 (gerelateerd aan onderzoeksvraag 4) is de invloed van compensatiemechanismen en de effecten van ziekte op teamproductie op de productiviteitskosten onderzocht. In de literatuur is er gesuggereerd dat, indien afwezigheid en/of verminderde productiviteit op het werk (gedeeltelijk) gecompenseerd wordt in normale werkuren door de zieke werknemer of een collega, gebruikelijke productiviteitskostenschattingen een overschatting zijn van de werkelijke kosten [17,18]. Er is echter ook beargumenteerd dat huidige schattingen juist een onderschatting van werkelijke kosten zijn [15,16], omdat werknemers in veel beroepen in hun werk afhankelijk zijn van hun collega's. Als in dat geval één van de teamgenoten ziek wordt, kan dit een negatief effect hebben op de productiviteit van het gehele team. In de empirische studie beschreven in hoofdstuk 8 werd onderzocht hoe het verzuim van zieke medewerkers wordt gecompenseerd. Daarbij werd onderzocht wat de mogelijke invloed van compensatiemechanismen en effecten op teamproductie is op productiviteitskostenschattingen.

Meer dan de helft van de respondenten gaf aan dat hun gemiste werk door verzuim gecompenseerd werd in normale werkuren door henzelf of collega's. Wanneer er bij het berekenen van productiviteitskosten tegelijk rekening gehouden werd met het effect van verzuim op teamproductie en compensatiemechanismen, was het verlagende effect van compensatiemechanismen op de schattingen sterker dan het verhogende effect van de invloed op teamproductie. Nota bene, deze berekeningen waren gebaseerd op de naïeve assumptie dat compensatie in normale werkuren kosteloos is. Dit laatste is vanuit een economisch perspectief echter onwaarschijnlijk.

Zoals aangetoond in hoofdstuk 6 en 7 worden productiviteitskosten vaak niet meegenomen in economische evaluaties. Dit hangt mogelijk samen met praktische moeilijkheden met het verkrijgen van betrouwbare productiviteitskostenschattingen. Er is derhalve wel geopperd dat productiviteitskosten dan wellicht indirect geschat zouden kunnen worden op basis van kwaliteit van leven data. In hoofdstuk 9 (in antwoord op onderzoeksvraag 5) werd een studie gepresenteerd die als doel had een algoritme te ontwikkelen en te valideren om productiviteit te schatten op basis van kwaliteit van leven data. Om dit te doen werd een grote representatieve groep van het algemene Nederlandse publiek middels een internet vragenlijst gevraagd naar hun verwachte niveau van productiviteit (in termen van verzuim en verminderde productiviteit op het werk) in diverse EQ-5D gezondheidstoestanden.

Op basis van deze gegevens werden twee predictiemodellen geconstrueerd: i) een predictiemodel voor verzuim en ii) een predictiemodel voor verminderd functioneren op het werk. Beide modellen werden gevalideerd door de modelschattingen te vergelijken met conventioneel gemeten productiviteit in een groep patiënten met lage rugpijn. Het geschatte verzuim op basis van EQ-5D gezondheidstoestanden lag dicht bij het normaal gemeten verzuim van de lage rugpijn patiënten. Productiviteitsverliezen op het werk werden enigszins overschat door het predictiemodel. Gemeten en geschatte productiviteit was matig maar sterk significant gecorreleerd. Het bleek derhalve mogelijk om tot redelijke productiviteitsschattingen te komen op de basis van EQ-5D data. Verder onderzoek naar en verdere validatie van dergelijke predictiemodellen blijft echter noodzakelijk, zeker voor verminderde productiviteit op het werk.

In hoofdstuk 10 werden de belangrijkste onderzoeksbevindingen bediscussieerd en werden de implicaties en beperkingen van dit proefschrift behandeld. Gebaseerd op de resultaten (en terugkomend op de vijf belangrijkste vragen die dit proefschrift probeerde te beantwoorden) konden enkele conclusies worden getrokken. Ten eerste, ongeveer de helft van de respondenten van gezondheidstoestandwaarderingen lijkt inkomenseffecten mee te nemen. Dit blijkt weinig effect te hebben op deze waarderingen. Productiviteitskosten zouden daarom bij voorkeur meegenomen moeten worden aan de kostenkant van de kosteneffectiviteitsratio. Ten tweede, ondanks de wetenschappelijke aandacht voor productiviteitskosten worden deze kosten in de meeste economische evaluaties niet meegenomen, vaak zonder duidelijke argumentatie. Ten derde, wanneer productiviteitskosten worden meegenomen kan dit de kosteneffectiviteitsuitkomsten zeer sterk beïnvloeden. Deze sterke potentiele impact op kosteneffectiviteitsuitkomsten maakt duidelijk dat wanneer productiviteitskosten niet worden meegenomen in economische evaluaties, vergoedingsbeslissingen gebaseerd op de resultaten van deze evaluaties kunnen leiden tot inefficiënt gebruik van maatschappelijke middelen. Ten vierde, verloren productiviteit wordt vaak gecompenseerd in normale werkuren. Productiviteitskostenschattingen worden sterk beïnvloed door eventuele correcties voor compensatiemechanismen en effecten op teamproductie. De validiteit van zulke correcties dient echter nader onderzocht te worden. Ten vijfde, hoewel de validiteit van het schatten van productiviteitsverliezen gebaseerd op kwaliteit van leven data nog verder onderzocht moet worden, lijkt het toepassen van predictiemodellen, wanneer data over werkelijke productiviteit niet beschikbaar is, te verkiezen boven het geheel negeren van het effect van gezondheid op productiviteit.

Dit proefschrift eindigde met enkele opmerkingen over de inclusie van productiviteitskosten in economische evaluaties en de methodologie om deze kosten te schatten. Er bestaat nog altijd geen consensus over of productiviteitskosten thuishoren in economische evaluaties en hoe deze kosten geschat zouden moeten worden. Productiviteitskosten worden dan ook vaak genegeerd in economische evaluaties. Vanuit een maatschappelijk perspectief is het echter essentieel dat de inclusie van productiviteitskosten toeneemt, vooral gezien de sterke potentiële impact van productiviteitskosten op kosteneffectiviteitsuitkomsten. Het is dan ook belangrijk dat besluitvormers zich meer bewust zijn van deze potentiële impact en dat men meer benadrukt dat men zich dient te houden aan het voorgeschreven perspectief. Daarnaast is het gebrek aan standaardisatie van productiviteitskostenmethodologie een belangrijk probleem. In dit proefschrift is aangetoond dat in economische evaluaties een diversiteit aan methoden wordt gebruikt om productiviteitskosten te identificeren, meten en waarderen. Het toepassen van verschillende methoden kan leiden tot een grote variatie in productiviteitskostenschattingen en komt de betrouwbaarheid niet ten goede. Daarbij bemoeilijkt de diversiteit in methoden de mogelijkheid om uitkomsten van studies met elkaar te vergelijken. Dit proefschrift heeft benadrukt dat standaardisatie essentieel is om te bewerkstelligen dat het includeren van productiviteitskosten in economische evaluaties de standaard wordt in plaats van een uitzondering.

Tot slot, dit proefschrift heeft bijgedragen aan de theoretische en praktische kennis over de inclusie van productiviteitskosten in economische evaluaties. Hoewel het (gegeven de bevindingen van dit proefschrift) een uitdaging zal zijn om volledige consensus te bereiken wat betreft de inclusie van productiviteitskosten in economische evaluaties en de methodologie om deze kosten te identificeren, te meten en te waarderen, blijft dit een belangrijk en nastrevenswaardig doel.

Curriculum vitae
PhD portfolio
Dankwoord

Curriculum vitae

Marieke Krol was born in Hoogezand-Sappemeer in 1978. In 1996 she graduated from her Atheneum education at the Openbare Scholengemeenschap Nieuwediep in Den Helder. In 2001 she obtained a Bachelor's degree in nursing from The Hague University of Professional Education. For five years she worked as a nurse at the psychiatric department of GGZ Delfland in the Reinier de Graaf Gasthuis in Delft. In 2004 she received her Master's degree in Health Economics Policy and Law from Erasmus University Rotterdam, which she later complemented with a Health Economics specialization.

In 2006, she started working as a researcher at the institute for Medical Technology Assessment and at the institute of Health Policy & Management. Her research mainly focuses on methodological issues of economic evaluations (involving issues on costs as well as effects of health care interventions). She is furthermore active in teaching in the Health Economics, Policy and Law Master. She is coordinator of the Health Technology Assessment course as well as the coordinator of the HTA specialization within this program.

She has published papers in peer-reviewed journals, such as Pharmacoeconomics, Value in Health, Social Science and Medicine and Medical Decision Making.

PhD portfolio

Training

- Modeling health care costs and counts, International Health Economics Association 8th World Congress, Toronto, 2011
- Academic writing, Erasmus University Rotterdam, 2010
- Q-methodology, Erasmus University Rotterdam, 2007
- Advanced Modelling Methods for Economic Evaluation, University of Glasgow, 2007

Teaching

- Introduction in HTA/economic evaluation, lecture, course Health Technology Assessment (Master program Health Economics and the Master program Health Economics, Policy and Law, Erasmus University Rotterdam), 2012
- Contingent valuation of benefits of health care, practicum, course Health Technology Assessment 2010 - present
- Quality assessment of economic evaluations, practicum, course Health Technology Assessment 2010 - present
- Essay guidance and evaluation, course Health Technology Assessment 2008 - present
- Utility measurement exercises, practicum, course Health Technology Assessment 2008 - present
- Cost-effectiveness calculations, practicum, course Health Technology Assessment 2007 - present
- Utility measurement, lecture, course Health Technology Assessment 2010, 2011
- Outcomes beyond the QALY, lecture, course Health Technology Assessment 2010
- Quality of life research, lecture, course Kwaliteit en Doelmatigheid (Bachelor Health Sciences, Policy and Management, Erasmus University Rotterdam), 2007, 2008, 2010
- Working group supervisor, course Kwaliteit en Doelmatigheid 2007, 2008

Coordination

- Coordinator of the HTA specialization of the Master program ‘Health Economics, Policy and Law’, Erasmus University Rotterdam, 2012
- Coordinator of the HTA course of the Master program ‘Health Economics’ and ‘Health Economics, Policy and Law’ Erasmus University Rotterdam, 2009 - present

Thesis supervision

- Supervision and co-supervision of several theses of the Bachelor ‘Health Science, Policy and Management’, the Master ‘Health Economics’ and the master ‘Health Economics, Policy and Law’, Erasmus University Rotterdam, 2007 - present

Teaching at other institutes

- Quality of life, lecture, course 'Economic Evaluation' (Research Master Lifestyle and Chronic Disorders), Vrije Universiteit, Amsterdam, 2010
- Two-day HTA course for Romanian professionals in health care (e.g. members of the ministry of health and employees of pharmaceutical companies), Bucharest, 2012

Conference organization

- Organizer of the joined NVTAG/ZonMw/iBMG symposium on 'Outcomes Beyond the QALY', Rotterdam, 2012
- Member of the organizing committee of the Low Lands Health Economists' Study Group (annual Dutch/Flemish health economic conference) Egmond aan Zee, 2010

Presentations

- Moderator of the workshop 'The value of work in Health Technology Assessment' at the Fit For Work Annual Summit, Brussels, 2012
- Productivity predictions based on EQ-5D, 12th European Conference on Health Economics, 2012
- Significant others in health state valuations, International Health Economics Association 8th World Congress, 2011
- Significant others in health state valuations, Low Lands Health Economists' Study Group, 2011
- The place of productivity costs in economic evaluations of expensive drugs, 8th European Conference on Health Economics, 2010
- Do productivity costs matter? A study on the impact of productivity costs in cost-effectiveness outcomes, 8th European Conference on Health Economics, *poster*, 2010
- Is there more to life than money? A new approach to estimating the monetary value of a QALY, Low Lands Health Economists' Study Group, 2009
- Indirect costs of beta thalassemia major: results from the ITHACA study, International Health Economics Association 6th World Congress, *poster*, 2007
- A systematic review of economic analyses of pharmaceutical therapies for advanced colorectal cancer, International Health Economics Association 6th World Congress, *poster*, 2007
- Presentation on 'Outcomes research' at the outcomes research course organized by Stichting Post-Universitair Onderwijs Ziekenhuisfarmacie, Utrecht, 2007
- Breaking the silence: the effects of incorporating income in TTO Exercises, International Society for Pharmacoeconomics and Outcomes Research 9th Annual European Congress, 2006

publications

- Krol M, Brouwer WBF, Rutten FFH. Productivity costs in economic evaluations: Past, present, future. *Accepted for publication in Pharmacoeconomics*.
- Heintz E, Krol M, Levin LA. The impact of patients' subjective life expectancy on time trade-off (TTO) valuations. Med.Decis.Making. *In press*.
- Krol M, Brouwer WB, Severens JL, Kaper J, Evers SM. Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects. Soc.Sci.Med. 2012 Dec;75(11):1981-1988.
- Tilling C, Krol M, Tsuchiya A, Brazier J, Exel J, Brouwer W. Does the EQ-5D reflect lost earnings? Pharmacoeconomics 2012 Jan;30(1):47-61.
- Krol M, Papenburg J, Koopmanschap M, Brouwer W. Do productivity costs matter? The impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders. Pharmacoeconomics 2011 Jul 1;29(7):601-619.
- Tilling C, Krol M, Tsuchiya A, Brazier J, Brouwer W. In or out? Income losses in health state valuations: a review. Value Health. 2010 Mar-Apr;13(2):298-305.
- Krol M, Sendi P, Brouwer W. Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. Value Health. 2009 Jan;12(1):172-180.
- Tilling C, Krol M, Tsuchiya A, Brazier J, van Exel NJ, Brouwer WBF. Measuring the value of life: Exploring a new method for deriving the monetary value of a QALY. HEDS discussion paper 2009;09-14.
- Scalone L, Mantovani LG, Krol M, Rofail D, Ravera S, Bisconte MG, et al. Costs, quality of life, treatment satisfaction and compliance in patients with beta-thalassemia major undergoing iron chelation therapy: the ITHACA study. Curr.Med.Res.Opin. 2008 Jul;24(7):1905-1917.
- Krol M, Koopman M, Uyl-de Groot C, Punt CJ. A systematic review of economic analyses of pharmaceutical therapies for advanced colorectal cancer. Expert Opin. Pharmacother. 2007 Jun;8(9):1313-1328.
- Krol M, Brouwer W, Sendi P. Productivity costs in health-state valuations : does explicit instruction matter? Pharmacoeconomics 2006;24(4):401-414.

Dankwoord

'Nait soez'n, moar doun'. Een vleugje Groningse mentaliteit kwam tijdens het schrijven van dit proefschrift zeer van pas. Dat dit echter niet voldoende is om een proefschrifttraject succesvol af te sluiten, weet iedereen die hier ooit aan is begonnen. Aan de totstandkoming van dit proefschrift hebben veel mensen een bijdrage geleverd, sommigen aan de inhoud van dit boekje, anderen aan het proces, een aantal aan beide.

Werner, jij valt duidelijk in deze laatste categorie. Zonder jou was dit proefschrift er niet geweest en mijn dank aan jou is groot. Na het schrijven van twee masterscripties onder jouw hoede was dit proefschrift een prachtig vervolg. Het was een plezier en een eer om hier met jou aan te werken.

De promotiecommissie dank ik van harte voor het lezen en beoordelen van dit proefschrift en voor het opponeren bij de verdediging.

Ook wil ik de sponsors van het Broader Social Benefits project (AstraZeneca, GlaxoSmithKline, Janssen, Merck en Pfizer) bedanken voor de financiële steun aan het onderzoek gepresenteerd in hoofdstuk 2 van dit proefschrift en ZonMw voor de financiële steun aan het onderzoek waarop hoofdstuk 7 is gebaseerd.

De meeste hoofdstukken van dit proefschrift zijn ontstaan uit de samenwerking met gedreven collega-wetenschappers. Ik ben de coauteurs van de diverse hoofdstukken erg dankbaar. Hun enthousiasme, kennis en inzicht waren een grote verrijking.

Ik ben veel dank verschuldigd aan de zeer vele (anonieme) respondenten die deelnamen aan de empirische studies. Hun belangeloze inspanningen waren essentieel voor het tot stand komen van dit proefschrift.

Mijn ouders, broer, zus en alle belangrijke mensen om hen heen zijn mij immer tot steun (in veel meer dingen dan enkel het voltooien van dit proefschrift). Jullie zijn onmisbaar. Pap, jouw aandeel in de dataverzameling in hoofdstuk 3 en 4 was fantastisch. Ook mijn schoonouders ben ik zeer dankbaar. Jan, bedankt voor alle hulp bij het verbouwen van 't Drijfhuis terwijl ik rustig aan dit proefschrift werkte.

Werken bij GE-iMTA/iBMG is na zes jaar nog altijd leuk, leerzaam en uitdagend. Dit is te danken aan de vele inspirerende en gezellige collega's.

Een aantal van hen wil ik in het bijzonder bedanken. Chantal, zonder jou waren de afgelopen jaren bij iMTA (maar ook daarbuiten...) niet hetzelfde geweest. Samen een kamer delen was niet altijd even efficiënt, maar elke dag leuk. Marten, ontzettend bedankt voor het feit dat je er altijd bent. Je bent een grote steun geweest op vele vlakken. Marc, samen met jou lesgeven en het vak Health Technology Assessment coördineren is een groot plezier. Jouw bevlogenheid wat betreft onderzoek en onderwijs is een voorbeeld. Lucas, het is super om met jou te discussiëren over de vorm en inhoud van ons beider werk en over allerlei andere dingen. Ik had me geen fijnere 'nieuwe' kamergenoot kunnen wensen. Siok Swan, jij was er vanaf het eerste uur. Aio's bestonden er bij iMTA nog niet in 2006 en er was letterlijk slechts een handjevol junioren. Bedankt voor alle leuke momenten door de jaren heen. Job, bedankt voor je voortdurende bereidheid jouw ideeën te delen als ik behoefte heb aan een frisse, scherpe blik op van alles en nog wat. Elly, dat geldt ook voor jou. Leona, bedankt voor alle gezelligheid op donderdag-krokettendag (en vele andere dagen). Sander, zonder jou had dit proefschrift er niet hetzelfde uitgezien. Bedankt voor al je input rondom de vormgeving en natuurlijk voor alle leuke jaren in de activiteitencommissie. Karin en Maddy, bedankt voor de assistentie bij het inplannen van afspraken in de zeer volle agenda van Werner.

Naast familie en collega's wil ik een aantal vriendinnen bedanken. Tjitske, ik ben ontzettend blij dat je er altijd voor me bent (weliswaar op enige afstand) en dat je mijn paranimf bent. Jocé, wat was het leuk om met jou samen te werken aan hoofdstuk 6 en 7 van dit proefschrift. Ik mis je ontzettend aan de andere kant van de oceaan. Caro, Gerda, Ingrid, Marin, Sandra en Suus, bedankt voor de gezellige ochtenden bij Knus, etentjes en borrelavonden.

Maarten, Myrthe-Minou en Sylvijn, jullie geven elke dag kleur! Mijn liefde voor jullie is niet in woorden te bevatten. Mijn leven met jullie mogen delen is een feest. Ik ben dankbaar voor elke dag samen.

References

References

(1) Maynard A. Rationing health care: an exploration. Health Policy 1999 Oct;49(1-2):5-11.

(2) Gold M, Siegel J, Russell L, Weinstein M. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.

(3) Eddama O, Coast J. A systematic review of the use of economic evaluation in local decision-making. Health Policy 2008 May;86(2-3):129-141.

(4) Williams I, Bryan S. Understanding the limited impact of economic evaluation in health care resource allocation: a conceptual framework. Health Policy 2007 Jan;80(1):135-143.

(5) Brousselle A, Lessard C. Economic evaluation to inform health care decision-making: promise, pitfalls and a proposal for an alternative path. Soc.Sci.Med. 2011 Mar;72(6):832-839.

(6) Banta D. The development of health technology assessment. Health Policy 2003 Feb;63(2):121-132.

(7) Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford ; New York: Oxford University Press; 2005.

(8) Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion. Health Econ. 1997 Sep-Oct;6(5):511-514.

(9) Koopmanschap MA, Rutten FF. The impact of indirect costs on outcomes of health care programs. Health Econ. 1994 Nov-Dec;3(6):385-393.

(10) Stone PW, Chapman RH, Sandberg EA, Liljas B, Neumann PJ. Measuring costs in cost-utility analyses. Variations in the literature. Int.J.Technol.Assess.Health Care 2000 Winter;16(1):111-124.

(11) Tranmer JE, Guerriere DN, Ungar WJ, Coyte PC. Valuing patient and caregiver time: a review of the literature. Pharmacoeconomics 2005;23(5):449-459.

(12) Olsen J, Richardson J. Production gains from health care: what should be included in cost-effectiveness analysis. Soc.Sci.Med. 1999;49:17-26.

(13) Williams A. Cost-effectiveness analysis: is it ethical? J.Med.Ethics 1992 Mar;18(1):7-11.

(14) Pauly MV, Nicholson S, Xu J, Polsky D, Danzon PM, Murray JF, et al. A general model of the impact of absenteeism on employers and employees. Health Econ. 2002 Apr;11(3):221-231.

(15) Nicholson S, Pauly MV, Polsky D, Sharda C, Szrek H, Berger ML. Measuring the effects of work loss on productivity with team production. Health Econ. 2006 Feb;15(2):111-123.

(16) Pauly MV, Nicholson S, Polsky D, Berger ML, Sharda C. Valuing reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. Health Econ. 2008 Apr;17(4):469-485.

(17) Severens JL, Laheij RJ, Jansen JB, Van der Lisdonk EH, Verbeek AL. Estimating the cost of lost productivity in dyspepsia. Aliment.Pharmacol.Ther. 1998 Sep;12(9):919-923.

(18) Jacob-Tacken KH, Koopmanschap MA, Meerding WJ, Severens JL. Correcting for compensating mechanisms related to productivity costs in economic evaluations of health care programmes. Health Econ. 2005 May;14(5):435-443.

(19) Lamers LM, Meerding WJ, Severens JL, Brouwer WB. The relationship between productivity and health-related quality of life: an empirical exploration in persons with low back pain. Qual.Life Res. 2005 Apr;14(3):805-813.

(20) Brouwer WB, Meerding WJ, Lamers LM, Severens JL. The relationship between productivity and health-related QOL: an exploration. Pharmacoeconomics 2005;23(3):209-218.

(21) Weinstein MC, Siegel JE, Garber AM, Lipscomb J, Luce BR, G. MW,Jr, et al. Productivity costs, time costs and health-related quality of life: a response to the Erasmus Group. Health Econ. 1997 Sep-Oct;6(5):505-510.

(22) Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. Health Econ. 1997 May-Jun;6(3):253-259.

(23) Nyman J. Productivity costs revisited: Toward a new US policy. Health Econ. 2011 (Epub ahead of print).

(24) Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann.Med. 2001 Jul;33(5):337-343.

(25) Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics 1995 Jun;7(6):490-502.

(26) Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ. 1996 Mar-Apr;5(2):141-154.

(27) Torrance GW. Preferences for health states: a review of measurement methods. Mead Johnson Symp.Perinat.Dev.Med. 1982;(20)(20):37-45.

(28) Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv.Res. 1972 Summer;7(2):118-133.

(29) Gafni A. The standard gamble method: what is being measured and how it is interpreted. Health Serv.Res. 1994 Jun;29(2):207-224.

(30) Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. Qual.Life Res. 1996 Dec;5(6):521-531.

(31) Meltzer D, Johannesson M. Inconsistencies in the “societal perspective” on costs of the Panel on Cost-Effectiveness in Health and Medicine. Med.Decis.Making 1999 Oct-Dec;19(4):371-377.

(32) Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. Health Econ. 2011 Jan;20(1):2-15.

(33) Claxton K, Walker S, Palmer S, Sculpher M. Appropriate Perspectives for Health Care Decisions. CHE 2010;CHE Reasearch paper 54.

(34) Jonsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. Eur.J.Health Econ. 2009 Oct;10(4):357-359.

(35) Johannesson M, Jönsson B, Jönsson L, Kobelt G, Zethraeus N. Why Should Economic Evaluations of Medical Innovations Have a Societal Perspective? OHE briefing 2009(51).

(36) Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. Med.Decis.Making 2000 Jul-Sep;20(3):332-342.

(37) Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ. 2010 Apr;19(4):422-437.

(38) Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. J.Health Econ. 2012 Jan;31(1):158-168.

(39) Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: A critical review. Soc.Sci.Med. 2011 Jan;72(2):185-192.

(40) Johannesson M. A note on the depreciation of the societal perspective in economic evaluation of health care. Health Policy 1995 Jul;33(1):59-66.

(41) Brouwer WBF, Exel JA, Baltussen RMPM, Rutten FFH. A Dollar Is a Dollar-or Is It? Value Health. 2006 Sep-Oct;9(5):341-347.

(42) Lindholm L, Lofroth E, Rosen M. Does productivity influence priority setting? A case study from the field of CVD prevention. Cost.Eff.Resour.Alloc 2008 Mar 17;6:6.

(43) Pritchard C, Sculpher M. Productivity costs: principles and practice in economic evaluation. 2000.

(44) Krol M, Papenburg J, Koopmanschap M, Brouwer W. Do productivity costs matter?: the impact of including productivity costs on the incremental costs of interventions targeted at depressive disorders. Pharmacoeconomics 2011 Jul 1;29(7):601-619.

(45) Krol M, Papenburg J, Tan S, Brouwer W, Hakkaart L. A Noticeable Difference? Productivity costs related to paid and unpaid work in economic evaluations on expensive drugs. Submitted paper.

(46) Brooks A, Hagen SE, Sathyanarayanan S, Schultz AB, Edington DW. Presenteeism: critical issues. J.Occup.Environ.Med. 2010 Nov;52(11):1055-1067.

(47) Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J.Health Econ. 1995 Jun;14(2):171-189.

(48) Brouwer WB, Koopmanschap MA, Rutten FF. Productivity losses without absence: measurement validation and empirical evidence. Health Policy 1999 Jul;48(1):13-27.

(49) Posnett J, Jan S. Indirect cost in economic evaluation: the opportunity cost of unpaid inputs. Health Econ. 1996 Jan-Feb;5(1):13-23.

(50) Brouwer WB, van Exel NJ, Koopmanschap MA, Rutten FF. Productivity costs before and after absence from work: as important as common? Health Policy 2002 Aug;61(2):173-187.

(51) Koopmanschap M, Burdorf A, Jacob K, Meerding WJ, Brouwer W, Severens H. Measuring productivity changes in economic evaluation: setting the research agenda. Pharmacoeconomics 2005;23(1):47-54.

(52) Prasad M, Wahlqvist P, Shikiar R, Shih YC. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. Pharmacoeconomics 2004;22(4):225-244.

(53) Mattke S, Balakrishnan A, Bergamo G, Newberry SJ. A review of methods to measure health-related productivity loss. Am.J.Manag.Care 2007 Apr;13(4):211-217.

(54) Lofland JH, Pizzi L, Frick KD. A review of health-related workplace productivity loss instruments. Pharmacoeconomics 2004;22(3):165-184.

(55) Krol M, Stolk E, Brouwer W. Productivity costs predictions based on EQ-5D: an explorative study. Submitted paper.

(56) Koopmanschap MA, Rutten FF. Indirect costs in economic studies: confronting the confusion. Pharmacoeconomics 1993 Dec;4(6):446-454.

(57) Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. Reply to Johanneson's and Karlsson's comment. J.Health Econ. 1997;16:257-259.

(58) Johannesson M, Karlsson G. The friction cost method: a comment. J.Health Econ. 1997 Apr;16(2):249-55; discussion 257-259.

(59) Brouwer WB, Koopmanschap MA. The friction-cost method: replacement for nothing and leisure for free? Pharmacoeconomics 2005;23(2):105-111.

(60) Meltzer D, Weckerle C, Chang L. Do people consider financial effects in answering quality of life questions? Med.Decis.Making 1999;19:517.

(61) Myers J, McCabe S, Gohmann S. Quality-of-life assessment when there is a loss of income. Med. Decis.Making 2007 Jan-Feb;27(1):27-33.

(62) Richardson J, Peacock SJ, Iezzi A. Do quality-adjusted life years take account of lost income? Evidence from an Australian survey. Eur.J.Health.Econ. 2009 Feb;10(1):103-109.

(63) Davidson T, Levin LA. Do individuals consider expected income when valuing health states? Int.J.Technol.Assess.Health Care 2008 Fall;24(4):488-494.

(64) Sendi P, Brouwer WB. Is silence golden? A test of the incorporation of the effects of ill-health on income and leisure in health state valuations. Health Econ. 2005 Jun;14(6):643-647.

(65) Krol M, Brouwer W, Sendi P. Productivity costs in health-state valuations : does explicit instruction matter? Pharmacoeconomics 2006;24(4):401-414.

(66) Krol M, Sendi P, Brouwer W. Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. Value Health. 2009 Jan;12(1):172-180.

(67) Brouwer WB, Grootenboer S, Sendi P. The incorporation of income and leisure in health state valuations when the measure is silent: an empirical inquiry into the sound of silence. Med. Decis.Making 2009 Jul-Aug;29(4):503-512.

(68) Brouwer WB, Rutten FF. The missing link: on the line between C and E. Health Econ. 2003 Aug;12(8):629-636.

(69) Culyer AJ, Bombard Y. An Equity Framework for Health Technology Assessments. Med.Decis. Making 2012 May-Jun;32(3):428-441.

(70) Brouwer WB, Rutten FF. Health economics. A bridge over troubled water. Eur.J.Public Health 2001 Jun;11(2):234-236.

(71) Belgian Health Care Knowledge Centre. Guidelines for pharmacoeconomic evaluations in Belgium KCE reports 78C. 2008.

(72) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008.

(73) Edling A, Stenberg A. General guidelines for economic evaluations from the Pharmaceutical Benefits Board. The Pharmaceutical Benefits Board 2003.

(74) College voor zorgverzekeringen. Guidelines for pharmacoeconomic research, updated version. 2006.

(75) Olsen JA. Production gains: should they count in health care evaluations? Scot.J.Polit.Econ. 1994;41(1):69-84.

(76) Koopmanschap MA, Rutten FF. The consequence of production loss or increased costs of production. Med.Care 1996 Dec;34(12 Suppl):DS59-68.

(77) Cyert R, March J. A behavioral theory of the firm. Englewood Cliffs: Prentice-Hall; 1963.

(78) Capri S, Ceci A, Terranova L, Merlo F, Mantovani L. Guidelines for economic evaluations in Italy: Recommendations from the Italian group of pharmacoeconomic studies. Drug Inf.J. 2001;35:189-201.

(79) Norwegian Medicines Agency. Norwegian guidelines for pharmacoeconomic analysis in connection with applications for reimbursement. 2002.

(80) Krol M, Brouwer WB, Severens JL, Kaper J, Evers SM. Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects. Soc. Sci.Med. 2012 Dec;75(11):1981-1988.

(81) Swiebel J. Unpaid Work and Policy-Making: Towards a Broader Perspective of Work and Employment. Discussion Paper of the United Nations Department of Economic and Social Affairs 1999;4.

(82) Meerding WJ, IJzelenberg W, Koopmanschap MA, Severens JL, Burdorf A. Health problems lead to considerable productivity loss at work among workers with high physical load jobs. J.Clin. Epidemiol. 2005 May;58(5):517-523.

(83) Zhang W, Gignac MA, Beaton D, Tang K, Anis AH, Canadian Arthritis Network Work Productivity Group. Productivity loss due to presenteeism among patients with arthritis: estimates from 4 instruments. J.Rheumatol. 2010 Sep;37(9):1805-1814.

(84) van Roijen L, Essink-Bot ML, Koopmanschap MA, Bonsel G, Rutten FF. Labor and health status in economic evaluation of health care. The Health and Labor Questionnaire. Int.J.Technol. Asses.Health Care 1996 Summer;12(3):405-415.

(85) Zhang W, Bansback N, Boonen A, Severens JL, Anis AH. Development of a composite questionnaire, the valuation of lost productivity, to value productivity losses: application in rheumatoid arthritis. Value Health. 2012 Jan;15(1):46-54.

(86) Reid M. Economics of Household Production. New York: John Wiley; 1934.

(87) Gerard K. Cost-utility in practice: a policy maker's guide to the state of the art. Health Policy 1992 Jul;21(3):249-279.

(88) Weisbrod BA. The Valuation of Human Capital. J.Polit.Econ. 1961 Oct.;69(5):425-436.

(89) Gerard K, Donaldson C, Maynard AK. The cost of diabetes. Diabet.Med. 1989 Mar;6(2):164-170.

(90) Koopmanschap MA, van Ineveld BM. Towards a new approach for estimating indirect costs of disease. Soc.Sci.Med. 1992 May;34(9):1005-1010.

(91) Johannesson M. Avoiding double-counting in pharmacoeconomic studies. Pharmacoeconomics 1997 May;11(5):385-388.

(92) Tilling C, Krol M, Tsuchiya A, Brazier J, Brouwer W. In or out? Income losses in health state valuations: a review. Value Health. 2010 Mar-Apr;13(2):298-305.

(93) Waldman JD, Kelly F, Arora S, Smith HL. The shocking cost of turnover in health care. Health Care Manage.Rev. 2004 Jan-Mar;29(1):2-7.

(94) Koopmanschap MA, van Exel JN, van den Berg B, Brouwer WB. An overview of methods and applications to value informal care in economic evaluations of healthcare. Pharmacoeconomics 2008;26(4):269-280.

(95) van Exel J, Bobinac A, Koopmanschap M, Brouwer W. The invisible hands made visible: recognizing the value of informal care in healthcare decision-making. Expert Rev.Pharmacoecon. Outcomes Res. 2008 Dec;8(6):557-561.

(96) van den Berg B, Brouwer W, van Exel J, Koopmanschap M, van den Bos GA, Rutten F. Economic valuation of informal care: lessons from the application of the opportunity costs and proxy good methods. Soc.Sci.Med. 2006 Feb;62(4):835-845.

(97) van den Berg B, Al M, Brouwer W, van Exel J, Koopmanschap M. Economic valuation of informal care: the conjoint measurement method applied to informal caregiving. Soc.Sci.Med. 2005 Sep;61(6):1342-1355.

(98) van den Berg B, Brouwer WB, Koopmanschap MA. Economic valuation of informal care. An overview of methods and applications. Eur.J.Health Econ. 2004 Feb;5(1):36-45.

(99) Brouwer WB, van Exel NJ, Koopmanschap MA, Rutten FF. The valuation of informal care in economic appraisal. A consideration of individual choice and societal costs of time. Int.J.Technol. Assess.Health Care 1999 Winter;15(1):147-160.

(100) Gerard K, Mooney G. QALY league tables: handle with care. Health Econ. 1993 Apr;2(1):59-64.

(101) Russell L. Is prevention better than cure? Washington, DC: The Brookings Institute; 1986.

(102) Sculpher M. 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 pp 94-112; 2001.

(103) Brouwer WBF, Rutten FFH, Koopmanschap MA. Costing in economic evaluation. In: Drummond MF, McGuire A, editors. Economic evaluation in health care: merging theory with practice New York: Oxford University Press; 2001. p. 68-93.

(104) Brouwer WB, Koopmanschap MA. On the economic foundations of CEA. Ladies and gentlemen, take your positions! J.Health Econ. 2000 Jul;19(4):439-459.

(105) Rothermich EA, Pathak DS. Productivity-cost controversies in cost-effectiveness analysis: review and research agenda. Clin.Ther. 1999 Jan;21(1):255-267.

(106) Sendi P, Brouwer W. Leisure time in economic evaluation: theoretical and practical considerations. Expert Rev.Pharmacoecon.Outcomes Res. 2004 Feb;4(1):1-3.

(107) EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990 Dec;16(3):199-208.

(108) Dolan P. Modeling valuations for EuroQol health states. Med.Care 1997 Nov;35(11):1095-1108.

(109) Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? Med.Decis.Making 2001 Jul-Aug;21(4):329-334.

(110) Statistics Netherlands. Available at: http://www.cbs.nl. Accessed December/12, 2005.

(111) Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ. 2002 Jul;11(5):447-456.

(112) Sculpher MJ, O'Brien BJ. Income effects of reduced health and health effects of reduced income: implications for health-state valuation. Med.Decis.Making 2000 Apr-Jun;20(2):207-215.

(113) Torrance GW. Measurement of health state utilities for economic appraisal. J.Health Econ. 1986 Mar;5(1):1-30.

(114) Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ. 2006 Jul;15(7):653-664.

(115) Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ. 2006 Oct;15(10):1121-1132.

(116) Arnesen T, Trommald M. Roughly right or precisely wrong? Systematic review of quality-of-life weights elicited with the time trade-off method. J.Health Serv.Res.Policy 2004 Jan;9(1):43-50.

(117) Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med.Care 2005 Mar;43(3):203-220.

(118) Greiner W, Claes C, Busschbach JJ, von der Schulenburg JM. Validating the EQ-5D with time trade off for the German population. Eur.J.Health.Econ. 2005 Jun;6(2):124-130.

(119) Rice DP, Cooper BS. The economic value of human life. Am.J.Public Health Nations Health 1967 Nov;57(11):1954-1966.

(120) Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ. 2002 Jun;11(4):341-353.

(121) Attema AE, Brouwer WB. The correction of TTO-scores for utility curvature using a risk-free utility elicitation method. J.Health Econ. 2009 Jan;28(1):234-243.

(122) van Nooten FE, Koolman X, Brouwer WB. The influence of subjective life expectancy on health state valuations using a 10 year TTO. Health Econ. 2009 May;18(5):549-558.

(123) Hjelmgren J, Berggren F, Andersson F. Health economic guidelines--similarities, differences and some implications. Value Health. 2001 May-Jun;4(3):225-250.

(124) ISPOR. Pharmacoeconomic Guidelines Around The World. Available at: http://www.ispor.org/PEguidelines/index.asp. Accessed 08/18, 2011.

(125) Hylan TR, Buesching DP, Tollefson GD. Health economic evaluations of antidepressants: a review. Depress.Anxiety 1998;7(2):53-64.

(126) Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch.Gen.Psychiatry 1994 Jan;51(1):8-19.

(127) Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. Am.J.Psychiatry 2000 Dec;157(12):1925-1932.

(128) World Health Organization. Investing in mental health. 2003.

(129) Kessler RC, Barber C, Birnbaum HG, Frank RG, Greenberg PE, Rose RM, et al. Depression in the workplace: effects on short-term disability. Health Aff.(Millwood) 1999 Sep-Oct;18(5):163-171.

(130) Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? J.Clin. Psychiatry 2003 Dec;64(12):1465-1475.

(131) Sobocki P, Ekman M, Agren H, Jönsson B, Rehnberg C. Model to assess the cost-effectiveness of new treatments for depression. Int.J.Technol.Assess.Health Care 2006 Fall;22(4):469-477.

(132) Valenstein M, Vijan S, Zeber J, Boehm K, Buttar A. The cost-utility of screening for depression in primary care. Ann.Intern.Med. 2001 Mar 6;134(5):345-360.

(133) Zhang M, Rost K, Fortney J. Earnings changes for depressed individuals treated by mental health specialists. Am.J.Psych. 1999;156(1):108-114.

(134) Ramsey JB. Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. Journal of the Royal Statistical Society 1969;31(2):350-371.

(135) Andrews G, Sanderson K, Corry J, Lapsley H. Using epidemiological data to model efficiency in reducing the burden of depression*. J.Ment.Health Policy Econ. 2000 Dec 1;3(4):175-186.

(136) Vos T, Corry J, Haby M, Carter R, Andrews G. Cost-effectiveness of cognitive-behavioural therapy and drug interventions for major depression. Aust.N.Z.J.Psychiatry 2005 Aug;39(8):683-692.

(137) Brown M, Nimmerrichter A, Guest J. Cost-effectiveness of mirtazapine compared to amitriptyline and fluoxetine in the treatment of moderate and severe depression in Austria. Eur. Psychiatry 1999 Jul;14(4):230-244.

(138) Hemels M, Kasper S, Walter E, Einarson T. Cost-effectiveness analysis of escitalopram: a new SSRI in the first-line treatment of major depressive disorder in Austria. Curr.Med.Res.Opin. 2004 Jun;20(6):869-878.

(139) Hemels M, Kasper S, Walter E, Einarson T. Cost-effectiveness of escitalopram versus citalopram in the treatment of severe depression. Ann.Pharmacother. 2004 Sep;38(9):1545.

(140) Demyttenaere K, Hemels M, Hudry J, Annemans L. A cost-effectiveness model of escitalopram, citalopram,and venlafaxine as first-line treatment for major depressive disorder in Belgium. Clin.Ther. 2005 Jan;27(1):111-124.

(141) Machado M, Iskedjian M, Ruiz I, Einarson T. The economic impact of introducing serotonin-noradrenaline reuptake inhibitors into the Brazilian national drug formulary: cost-effectiveness and budget-impact analyses. Pharmacoeconomics 2007;25(11):979-990.

(142) Browne G, Steiner M, Roberts J, Gafni A, Byrne C, Dunn E, et al. Sertraline and/or interpersonal psychotherapy for patients with dysthymic disorder in primary care: 6-month comparison with longitudinal 2-year follow-up of effectiveness and costs. J.Affect.Disord. 2002 Apr;68(2-3):317-330.

(143) Einarson T, Addis A, Iskedjian M. Pharmacoeconomic analysis of venlafaxine in the treatment of major depressive disorder. Pharmacoeconomics 1997 23-10-1997;12(2 Pt 2):286-296.

(144) Araya R, Flynn T, Rojas G, Fritsch R, Simon G. Cost-effectiveness of a primary care treatment program for depression in low-income women in Santiago, Chile. Am.J.Psychiatry 1997 Aug;12(2 Pt 2):286-296.

(145) Hosák L, Tuma I, Hanus H, Straka L. Costs and outcomes of use of amitriptyline, citalopram and fluoxetine in major depression: exploratory study. Acta medica (Hradec Králové) 2000;43(4):133-137.

(146) Sørensen J, Stage K, Damsbo N, Le Lay A, Hemels M. A Danish cost-effectiveness model of escitalopram in comparison with citalopram and venlafaxine as first-line treatments for major depressive disorder in primary care. Nord.J.Psychiatry 2007;61(2):100-108.

(147) Brown M, van Loon J, Guest J. Cost-effectiveness of mirtazapine relative to fluoxetine in the treatment of moderate and severe depression in France. Eur.J.Psychiat. 2000;14(1):15-25.

(148) Brown M, van Loon J, Guest J. Cost-effectiveness of mirtazapine relative to amitriptyline in the treatment of moderate and severe depression in France. Eur.J.Psychiat. 1999;13(4):197-208.

(149) Nuijten M, Hadjadjeba L, Evans C, van den Berg J. Cost effectiveness of fluvoxamine in the treatment of recurrent depression in France. Pharmacoeconomics 1998 Oct;14(4):433-445.

(150) Fantino B, Moore N, Verdoux H, Auray J. Cost-effectiveness of escitalopram vs. citalopram in major depressive disorder. Int.Clin.Psychopharmacol. 2007 Mar;22(2):107-115.

(151) Dardennes R, Berdeaux G, Lafuma A, Fagnani F. Comparison of the cost-effectiveness of milnacipran (a SNRI) with TCAs and SSRIs: a modeling approach. Eur.Psychiatry 1999 Jun;14(3):152-162.

(152) Dardennes R, Lafuma A, Fagnani F, Pribil C, Bisserbe J, Berdeaux G. Economic assessment of a maintenance treatment strategy in prevention of recurrent depressive disorder. Value Health. 2000 Jan-Feb;3(1):40-47.

(153) Kulp W, von der Schulenburg J, Greiner W. Cost-effectiveness of outpatient treatment in depressive patients with escitalopram in Germany. Eur.J.Health Econ. 2005 Dec;6(4):317-321.

(154) Casciano J, Arikian S, Tarride J, Doyle J, Casciano R. A pharmacoeconomic evaluation of major depressive disorder (Italy). Epidemiol.Psichiatr.Soc. 1999 Jul-Sep;8(3):220-231.

(155) François C, Toumi M, Aakhus A, Hansen K. A pharmacoeconomic evaluation of escitalopram, a new selective serotonin reuptake inhibitor. Comparison of cost-effectiveness between escitalopram, citalopram, fluoxetine, and venlafaxine for the treatment of depression in Norway. Eur.J.Health Econ. 2003;4(1):12-19.

(156) Serrano-Blanco A, Gabarron E, Garcia-Bayo I, Soler-Vila M, Caramés E, Peñarrubia-Maria M, et al. Effectiveness and cost-effectiveness of antidepressant treatment in primary health care: a six-month randomised study comparing fluoxetine to imipramine. J.Affect.Disord. 2006 Apr;91(2-3):153-163.

(157) Sacristán J, Gilaberte I, Boto B, Buesching D, Obenchain R, Demitrack M, et al. Cost-effectiveness of fluoxetine plus pindolol in patients with major depressive disorder: results from a randomized, double-blind clinical trial. Int.Clin.Psychopharmacol. 2000 Mar;15(2):107-113.

(158) Löthgren M, Hemels M, Francois C, Jonsson B. A cost-effectiveness analysis of escitalopram as first line treatment of depression in Sweden. Primary Care Psychiatry 2004;9(4):153-161.

(159) Nuijten M. Assessment of clinical guidelines for continuation treatment in major depression. Value Health. 2001 Jul-Aug;4(4):281-294.

(160) Bosmans J, Brook O, van Hout H, de Bruijne M, Nieuwenhuyse H, Bouter L, et al. Cost effectiveness of a pharmacy-based coaching programme to improve adherence to antidepressants. Pharmacoeconomics 2007;25(1):25-37.

(161) Bosmans J, Hermens M, de Bruijne M, van Hout H, Terluin B, Bouter L, et al. Cost-effectiveness of usual general practitioner care with or without antidepressant medication for patients with minor or mild-major depression. 2008 Nov;111(1):106-112.

(162) Smit F, Willemse G, Koopmanschap M, Onrust S, Cuijpers P, Beekman A. Cost-effectiveness of preventing depression in primary care patients: randomised trial. Br.J.Psychiatry 2006 Apr;188:330-336.

(163) Stant A, Ten Vergert E, den Boer P, Wiersma D. Cost-effectiveness of cognitive self-therapy in patients with depression and anxiety disorders. Acta Psychiatr.Scand. 2008 Jan;117(1):57-66.

(164) van Baardewijk M, Vis P, Einarson T. Cost effectiveness of duloxetine compared with venlafaxine-XR in the treatment of major depressive disorder. Curr.Med.Res.Opin. 2005 Aug;21(8):1271-1279.

(165) Schene A, Koeter M, Kikkert M, Swinkels J, McCrone P. Adjuvant occupational therapy for work-related major depression works: randomized trial including economic evaluation. Psychol. Med. 2007 Mar;37(3):351-362.

(166) Borghi J, Guest J. Economic impact of using mirtazapine compared to amitriptyline and fluoxetine in the treatment of moderate and severe depression in the UK. Eur.Psychiatry 2000;15:378-387.

(167) Fernandez J, Montgomery S, Francois C. Evaluation of the cost effectiveness of escitalopram versus venlafaxine XR in major depressive disorder. Pharmacoeconomics 2005;23(2):155-167.

(168) Kendrick T, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, et al. Cost-effectiveness of referral for generic care or problem-solving treatment from community mental health nurses, compared with usual general practitioner care for common mental disorders: Randomised controlled trial. Br.J.Psychiatry 2006 Jul;189:50-59.

(169) Romeo R, Patel A, Knapp M, Thomas C. The cost-effectiveness of mirtazapine versus paroxetine in treating people with depression in primary care. Int.Clin.Psychopharmacol. 2004 May;19(3):125-134.

(170) Wade A, Toumi I, Hemels M. A pharmacoeconomic evaluation of escitalopram versus citalopram in the treatment of severe depression in the United Kingdom. Clin.Ther. 2005 Apr;27(4):486-496.

(171) Wade A, Toumi I, Hemels M. A probabilistic cost-effectiveness analysis of escitalopram, generic citalopram and venlafaxine as a first-line treatment of major depressive disorder in the UK. Curr.Med.Res.Opin. 2005 Apr;21(4):631-642.

(172) King M, Sibbald B, Ward E, Bower P, Lloyd M, Gabbay M, et al. Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care. Health Technol.Assess. 2000;4(19):1-83.

(173) McCrone , Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro D, et al. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. Br.J.Psychiatry 2004 Jul;185:55-62.

(174) Boath E, Major K, Cox J. When the cradle falls II: The cost-effectiveness of treating postnatal depression in a psychiatric day hospital compared with routine primary care. J.Affect.Disord. 2003 Apr;74(2):159-166.

(175) Freeman H, Arikian S, Lenox-Smith A. Pharmacoeconomic analysis of antidepressants for major depressive disorder in the United Kingdom. Pharmacoeconomics 2000 Aug;18(2):143-148.

(176) Freemantle N, Mason J, Watt I. Evidence into practice. Prescribing selective serotonin reuptake inhibitors. Int.J.Technol.Assess.Health Care 1998 Spring;14(2):387-391.

(177) Kendrick T, Peveler R, Longworth L, Baldwin D, Moore M, Chatwin J, et al. Cost-effectiveness and cost-utility of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine: randomised controlled trial. Br.J.Psychiatry 2006 Apr;188:337-345.

(178) Knapp M, Romeo R, Mogg A, Eranti S, Pluck G, Purvis R, et al. Cost-effectiveness of transcranial magnetic stimulation vs. electroconvulsive therapy for severe depression: A multi-centre randomised controlled trial. J.Affect.Disord. 2008 Aug;109(3):273-285.

(179) Lenox-Smith A, Conway P, Knight C. Cost effectiveness of representatives of three classes of antidepressants used in major depression in the UK. Pharmacoeconomics 2004;22(5):311-319.

(180) McLoughlin D, Mogg A, Eranti S, Pluck G, Purvis R, Edwards D, et al. The clinical effectiveness and cost of repetitive transcranial magnetic stimulation versus electroconvulsive therapy in severe depression: a multicentre pragmatic randomised controlled trial and economic analysis. Health Technol.Assess. 2007 Jul;11(24):1-54.

(181) Miller P, Chilvers C, Dewey M, Fielding K, Gretton V, Palmer B, et al. Counseling versus antidepressant therapy for the treatment of mild to moderate depression in primary care: economic analysis. Int.J.Technol.Assess.Health Care 2003 Winter;19(1):80-90.

(182) Morrell C, Spiby H, Stewart P, Walters S, Morgan A. Costs and effectiveness of community postnatal support workers: randomised controlled trial. BMJ 2000 Sep 9;321(7261):593-598.

(183) Petrou S, Cooper P, Murray L, Davidson L. Cost-effectiveness of a preventive counseling and support package for postnatal depression. Int.J.Technol.Assess.Health Care 2006 Fall;22(4):443-453.

(184) Peveler R, Kendrick T, Buxton M, Longworth L, Baldwin D, Moore M, et al. A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. Health Technol.Assess. 2005 May;9(16):1-134

(185) Scott J, Palmer S, Paykel E, Teasdale J, Hayhurst H. Use of cognitive therapy for relapse prevention in chronic depression. Cost-effectiveness study. Br.J.Psychiatry 3 Mar;182:221-227.

(186) Simpson S, Corney R, Fitzgerald P, Beecham J. A randomized controlled trial to evaluate the effectiveness and cost-effectiveness of psychodynamic counselling for general practice patients with chronic depression. Psychol.Med. 2003 Feb;33(2):229-239.

(187) Simpson S, Corney R, Fitzgerald P, Beecham J. A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression. Health Technol.Assess. 2000;4(36):1-83.

(188) Tome M, Isaac M. Cost effectiveness study of a year follow-up of selective serotonin reuptake inhibitor (SSRI) and augmentor combination compared with SSRI and placebo. Int.Clin.Psychopharmacol. 1998 Jul;13(4):175-182.

(189) Tome M, Isaac M. Cost-benefit & cost-effectiveness analysis of the rapid onset of selective serotonin reuptake inhibitors by augmentation. Int.J.Psychiatry Med. 1997;27(4):377-390.

(190) Antonuccio D, Thomas M, Danton W. A cost-effectiveness analysis of cognitive behavior therapy and fluoxetine (prozac) in the treatment of depression. Behavior Therapy 1997;28:187-210.

(191) Wang P, Patrick A, Avorn J, Azocar F, Ludman E, McCulloch J, et al. The costs and benefits of enhanced depression care to employers. Arch.Gen.Psychiatry 2006 Dec;63(12):1345-1353.

(192) Doyle J, Casciano J, Arikian S, Tarride J, Gonzalez M, Casciano R. A multinational pharmacoeconomic evaluation of acute major depressive disorder (MDD): a comparison of cost-effectiveness between venlafaxine, SSRIs and TCAs. Value Health. 2001 Jan-Feb;4(1):16-31.

(193) Alegría M, Frank R, McGuire T. Managed care and systems cost-effectiveness: treatment for depression. Med.Care 2005 Dec;43(12):1225-1233.

(194) Croghan T, Melfi C, Crown W, Chawla A. Cost-effectiveness of antidepressant medications. J.Ment.Health Policy Econ. 1998 Oct 1;1(3):109-117.

(195) Dobrez D, Melfi C, Croghan T, Kniesner T, Obenchain R. Antidepressant treatment for depression: total charges and therapy duration*. J.Ment. Health Policy Econ. 2000 Dec 1;3(4):187-197.

(196) Frank R, McGuire T, Normand S, Goldman H. The value of mental health care at the system level: the case of treating depression. Health Aff.(Millwood.) 1999 Sep-Oct;18(5):71-88.

(197) Johnson R, McFarland B, Nichols G. Changing patterns of antidepressant use and costs in a health maintenance organisation. Pharmacoeconomics 1997 Mar;11(3):274-286.

(198) Kozel F, George M, Simpson K. Decision analysis of the cost-effectiveness of repetitive transcranial magnetic stimulation versus electroconvulsive therapy for treatment of nonpsychotic severe depression. CNS spectrums 2004;9(6):476-482.

(199) Lave J, Frank R, Schulberg H, Kamlet M. Cost-effectiveness of treatments for major depression in primary care practice. Arch.Gen.Psychiatry 8 Jul;55(7):645-651.

(200) Liu C, Hedrick S, Chaney E, Heagerty P, Felker B, Hasenberg N, et al. Cost-effectiveness of collaborative care for depression in a primary care veteran population. Psychiatr.Serv. 2003 May;54(5):698-704.

(201) Malone D. A budget-impact and cost-effectiveness model for second-line treatment of major depression. J.Manag.Care Pharm. 2007 Jul;13(6 Suppl A):S8-18.

(202) Miller N, Markowitz J, Kocsis J, Leon A, Brisco S, Garno J. Cost effectiveness of screening for clinical trials by research assistants versus senior investigators. J.Psychiatr.Res. 1999 Mar-Apr;33(2):81-85.

(203) Obenchain R, Melfi C, Croghan T, Buesching D. Bootstrap analyses of cost effectiveness in antidepressant pharmacotherapy. Pharmacoeconomics 1997 May;11(5):464-472.

(204) Pyne J, Rost K, Farahati F, Tripathi S, Smith J, Williams D, et al. One size fits some: the impact of patient treatment attitudes on the cost-effectiveness of a depression primary-care intervention. Psychol.Med. 2005 Jun;35(6):839-854.

(205) Pyne J, Rost K, Zhang M, Williams D, Smith J, Fortney J. Cost-effectiveness of a primary care depression intervention. J.Gen.Intern.Med. 2003 Jun;18(6):432-441.

(206) Revicki D, Brown R, Keller M, Gonzales J, Culpepper L, Hales R. Cost-effectiveness of newer antidepressants compared with tricyclic antidepressants in managed care settings. J.Clin. Psychiatry 1997 Feb;58(2):47-58.

(207) Revicki D, Siddique J, Frank L, Chung J, Green B, Krupnick J, et al. Cost-effectiveness of evidence-based pharmacotherapy or cognitive behavior therapy compared with community referral for major depression in predominantly low-income minority women. Arch.Gen.Psychiatry 2005 Aug;62(8):868-875.

(208) Rost K, Pyne J, Dickinson L, LoSasso A. Cost-effectiveness of enhancing primary care depression management on an ongoing basis. Ann.Fam.Med. 2005 Jan-Feb;3(1):7-14.

(209) Schoenbaum M, Miranda J, Sherbourne C, Duan N, Wells K. Cost-effectiveness of interventions for depressed Latinos. J.Ment.Health Policy Econ. 2004 Jun;7(2):69-76.

(210) Schoenbaum M, Unützer J, Sherbourne C, Duan N, Rubenstein L, Miranda J, et al. Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. JAMA 2001 Sep 19;286(11):1325-1330.

(211) Simon G, Katon W, VonKorff M, Unützer J, Lin E, Walker E, et al. Cost-effectiveness of a collaborative care program for primary care patients with persistent depression. Am.J.Psychiatry 2001 Oct;158(10):1638-1644.

(212) Simon G, Manning W, Katzelnick D, Pearson S, Henk H, Helstad C. Cost-effectiveness of systematic depression treatment for high utilizers of general medical care. Arch.Gen.Psychiatry 2001 Feb;58(2):181-187.

(213) Simon G, Von Korff M, Ludman E, Katon W, Rutter C, Unützer J, et al. Cost-effectiveness of a program to prevent depression relapse in primary care. Med.Care 2002 Oct;40(10):941-950.

(214) Simon G, VonKorff M, Rutter C, Wagner E. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. BMJ 2000 Feb 26;320(7234):550-554.

(215) Schoenbaum M, Sherbourne C, Wells K. Gender patterns in cost effectiveness of quality improvement for depression: results of a randomized, controlled trial. J.Affect.Disord. 2005 Aug;87(2-3):319-325.

(216) Trivedi M, Wan G, Mallick R, Chen J, Casciano R, Geissler E, et al. Cost and effectiveness of venlafaxine extended-release and selective serotonin reuptake inhibitors in the acute phase of outpatient treatment for major depressive disorder. J.Clin.Psychopharmacol. 2004 Oct;24(5):497-506.

(217) Wells K, Schoenbaum M, Duan N, Miranda J, Tang L, Sherbourne C. Cost-effectiveness of quality improvement programs for patients with subthreshold depression or depressive disorder. Psychiatr.Serv. 2007 Oct;58(10):1269-1278.

(218) Sullivan P, Valuck R, Saseen J, MacFall H. A comparison of the direct costs and cost effectiveness of serotonin reuptake inhibitors and associated adverse drug reactions. CNS drugs 2004;18(13):911-932.

(219) Hunt SM, McKenna SP. The QLDS: a scale for the measurement of quality of life in depression. Health Policy 1992 Oct;22(3):307-319.

(220) Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch.Gen.Psychiatry 1961 Jun;4:561-571.

(221) Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br.J.Psychiatry 1979 Apr;134:382-389.

(222) Hamilton M. A rating scale for depression. J.Neurol.Neurosurg.Psychiatry. 1960 Feb;23:56-62.

(223) Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav.Sci. 1974 Jan;19(1):1-15.

(224) Rascati KL, Smith MJ, Neilands T. Dealing with skewed data: an example using asthma-related costs of medicaid clients. Clin.Ther. 2001 Mar;23(3):481-498.

(225) BorstkankerVereniging Nederland. Rapport onderbehandeling van borstkankerpatiënten met uitgezaaide HER2-positieve tumoren. Mei 2005.

(226) Nederlandse zorgautoriteit. http://www.nza.nl/137706/145406/CI-1135.pdf.

(227) Delwel G. Leidraad voor Uitkomstenonderzoek 'ten behoeve van de beoordeling doelmatigheid intramurale geneesmiddelen'. Diemen: College voor Zorgverzekeringen; 2008.

(228) Australian Government Department of Health and Ageing: Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee. December 2008.

(229) Grocott R, Metcalfe S, Schoeler R, Priest V, Hall C, Brougham M, et al. Prescription for Pharmacoeconomic Analysis: Methods for cost-utility analysis. PHARMAC Pharmaceutical Management Agency 2007.

(230) Hoffmann C, Graf von der Schulenburg J. The influence of economic evaluation studies on decision making. A European survey. The EUROMET group. 2000 Jul;52(3):179-192.

(231) Birnbaum H. Friction-cost method as an alternative to the human-capital approach in calculating indirect costs. Pharmacoeconomics 2005;23(2):103-104.

(232) Tilling C, Krol M, Tsuchiya A, Brazier J, van Exel J, Brouwer W. Does the EQ-5D Reflect Lost Earnings? Pharmacoeconomics 2012 Jan;30(1):47-61.

(233) Knies S, Severens JL, Ament AJ, Evers SM. The transferability of valuing lost productivity across jurisdictions. differences between national pharmacoeconomic guidelines. Value Health. 2010 Aug;13(5):519-527.

(234) Eurostat: Annual average rate of change in Harmonized Indices of Consumer Prices (HICPs). Available at: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tsieb060. Accessed 03/07, 2011.

(235) van den Hout WB, Goekoop-Ruiterman YP, Allaart CF, de Vries-Bouwstra JK, Hazes JM, Kerstens PJ, et al. Cost-utility analysis of treatment strategies in patients with recent-onset rheumatoid arthritis. Arthritis Rheum. 2009 Mar 15;61(3):291-299.

(236) Norum J, Holtmon M. Adjuvant fluorouracil, epirubicin and cyclophosphamide in early breast cancer: is it cost-effective? Acta Oncol. 2005;44(7):735-741.

(237) Walsh CA, Minnock P, Slattery C, Kennedy N, Pang F, Veale DJ, et al. Quality of life and economic impact of switching from established infliximab therapy to adalimumab in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007 Jul;46(7):1148-1152.

(238) Maniadakis N, Fragoulakis V, Pectasides D, Fountzilas G. XELOX versus FOLFOX6 as an adjuvant treatment in colorectal cancer: an economic analysis. Curr.Med.Res.Opin. 2009 Mar;25(3):797-805.

(239) Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden. Int.J.Technol.Assess.Health Care 2009 Apr;25(2):181-189.

(240) Dewilde S, Turk F, Tambour M, Sandström T. The economic value of anti-IgE in severe persistent, IgE-mediated (allergic) asthma patients: adaptation of INNOVATE to Sweden. Curr.Med.Res.Opin. 2006 Sep;22(9):1765-1776.

(241) Gani R, Giovannoni G, Bates D, Kemball B, Hughes S, Kerrigan J. Cost-effectiveness analyses of natalizumab (Tysabri) compared with other disease-modifying therapies for people with highly active relapsing-remitting multiple sclerosis in the UK. Pharmacoeconomics 2008;26(7):617-627.

(242) Bristow RE, Santillan A, Salani R, Diaz-Montes TP, Giuntoli RL, Meisner BC, et al. Intraperitoneal cisplatin and paclitaxel versus intravenous carboplatin and paclitaxel chemotherapy for Stage III ovarian cancer: a cost-effectiveness analysis. Gynecol.Oncol. 7 Sep;106(3):476-481.

(243) Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. J.Rheumatol. 2009 Jan;36(1):16-26.

(244) Spalding JR, Hay J. Cost effectiveness of tumour necrosis factor-alpha inhibitors as first-line agents in rheumatoid arthritis. Pharmacoeconomics 2006;24(12):1221-1232.

(245) Kobelt G, Andlin-Sobocki P, Brophy S, Jönsson L, Calin A, Braun J. The burden of ankylosing spondylitis and the cost-effectiveness of treatment with infliximab (Remicade). Rheumatology (Oxford) 2004 Sep;43(9):1158-1166.

(246) Kobelt G, Eberhardt K, Geborek P. TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: costs and outcomes in a follow up study of patients with RA treated with etanercept or infliximab in southern Sweden. Ann.Rheum.Dis. 2004 Jan;63(1):4-10.

(247) Manns BJ, Lee H, Doig CJ, Johnson D, Donaldson C. An economic evaluation of activated protein C treatment for severe sepsis. N.Engl.J.Med. 2002 Sep 26;347(13):993-1000.

(248) Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. Am.J.Med. 2002 Oct 1;113(5):400-408.

(249) Norum J, Olsen JA, Wist EA, Lønning PE. Trastuzumab in adjuvant breast cancer therapy. A model based cost-effectiveness analysis. Acta Oncol. 2007;46(2):153-164.

(250) Kobelt G, Berg J, Lindgren P, Jonsson B, Stawiarz L, Hillert J. Modeling the cost-effectiveness of a new treatment for MS (natalizumab) compared with current standard practice in Sweden. Mult. Scler. 2008 Jun;14(5):679-690.

(251) Kobelt G, Sobocki P, Mulero J, Gratacos J, Collantes-Estevez E, Braun J. The cost-effectiveness of infliximab in the treatment of ankylosing spondylitis in Spain. Comparison of clinical trial and clinical practice data. Scand.J.Rheumatol. 2008;37(1):62-71.

(252) Kobelt G, Sobocki P, Sieper J, Braun J. Comparison of the cost-effectiveness of infliximab in the treatment of ankylosing spondylitis in the United Kingdom based on two different clinical trials. Int.J.Technol.Assess.Health Care 2007;23(3):368-375.

(253) Kobelt G, Andlin-Sobocki P, Maksymowych WP. The cost-effectiveness of infliximab (Remicade) in the treatment of ankylosing spondylitis in Canada. J.Rheumatol. 2006 Apr;33(4):732-740.

(254) Kobelt G, Jönsson L, Young A, Eberhardt K. The cost-effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. Rheumatology (Oxford) 2003 Feb;42(2):326-335.

(255) Boonen A, van der Heijde D, Severens JL, Boendermaker A, Landewé R, Braun J, et al. Markov model into the cost-utility over five years of etanercept and infliximab compared with usual care in patients with active ankylosing spondylitis. Ann.Rheum.Dis. 2006 Feb;65(2):201-208.

(256) Lidgren M, Jönsson B, Rehnberg C, Willking N, Bergh J. Cost-effectiveness of HER2 testing and 1-year adjuvant trastuzumab therapy for early breast cancer. Ann.Oncol. 2008 Mar;19(3):487-495.

(257) McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics 2008;26(9):733-744.

(258) Raad voor de Volksgezondheid en Zorg. Zinnige en duurzame zorg. 2006, http://www.rvz.net/data/download/advies_Zinnige_zorg.pdf http://www.rvz.net/data/download/advies_Zinnige_zorg.pdf.

(259) Norum J, Risberg T, Olsen JA. A monoclonal antibody against HER-2 (trastuzumab) for metastatic breast cancer: a model-based cost-effectiveness analysis. Ann.Oncol. 2005 Jun;16(6):909-914.

(260) Kaper J, Wagena EJ, Willemsen MC, van Schayck CP. Reimbursement for smoking cessation treatment may double the abstinence rate: results of a randomized trial. Addiction 2005 Jul;100(7):1012-1020.

(261) Karasek RA. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. Administrative Science Quarterly 1979 June;24:285-308.

(262) Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. Expert Rev.Pharmacoecon.Outcomes Res. 2005 Feb;5(1):23-28.

(263) Rodenburg-van Dieten HEM. Richtlijnen voor farmaco-economisch onderzoek; evaluatie en actualisatie. Diemen: College voor zorgverzekeringen; 2005.

(264) Hakkaart-van Rooijen L, Tan SS, Bouwmans CAM. Dutch costing manual. Rotterdam: institute for Medical Technology Assessment (by order of the Board of Health Insurances); 2010.

(265) Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. Value Health. 2009 Jun;12(4):409-418.

(266) Johannesson M, Meltzer D, O'Conor RM. Incorporating future costs in medical cost-effectiveness analysis: implications for the cost-effectiveness of the treatment of hypertension. Med. Decis.Making 1997 Oct-Dec;17(4):382-389.

(267) Sloane NJA. A library of orthogonal arrays. Available at: http://www2.research.att.com/~njas/oadir/. Accessed 30/6, 2011.

(268) Bekkering GE, van Tulder MW, Hendriks EJ, Koopmanschap MA, Knol DL, Bouter LM, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys.Ther. 2005 Jun;85(6):544-555.

(269) Hoeijenbos M, Bekkering T, Lamers L, Hendriks E, van Tulder M, Koopmanschap M. Cost-effectiveness of an active implementation strategy for the Dutch physiotherapy guideline for low back pain. Health Policy 2005 Dec;75(1):85-98.

(270) Krabbe PF, Tromp N, Ruers TJ, van Riel PL. Are patients' judgments of health status really different from the general population? Health.Qual.Life.Outcomes 2011 May;11:9:31.

(271) van Osch SM, Stiggelbout AM. Understanding VAS valuations: qualitative data on the cognitive process. Qual.Life Res. 2005 Dec;14(10):2171-2175.

(272) Drummond M, Rutten FF. New guidelines for economic evaluation in Germany and the United Kingdom: Are we any closer to developing international standards? OHE briefing 2008(46).

(273) Sculpher MJ, Drummond MF. Analysis sans frontieres: can we ever make economic evaluations generalisable across jurisdictions? Pharmacoeconomics 2006;24(11):1087-1099.

(274) ISPOR Good Outcomes Research Practices Index. 2011; Available at: http://www.ispor.org/workpaper/practices_index.asp. Accessed 11/04, 2011.

(275) Lipscomb J, Drummond M, Fryback D, Gold M, Revicki D. Retaining, and enhancing, the QALY. Value Health. 2009 Mar;12 Suppl 1:S18-26.

(276) Nord E, Daniels N, Kamlet M. QALYs: some challenges. Value Health. 2009 Mar;12 Suppl 1:S10-5.