Economic Evaluations in Healthcare: Methodological Considerations and Applications on Cost-Effectiveness of ACE Inhibitor Treatment for Renal Disease Prevention

Charles Christian Adarkwah-Yiadom
2015
The research for this thesis was performed at the Department of Health Services Research, within the School for Public Health and Primary Care (CAPHRI), Faculty of Health, Medicine and Life Sciences, Maastricht University.

Printing: Gildeprint
Layout: Tiny Wouters
Cover: Florian Afflerbach

ISBN: 9789462330474
Copyright © Charles Christian Adarkwah-Yiadom, Maastricht, 2015

No parts of this thesis may be reproduced, stored in a retrieved system, or transmitted, in any forms or by any means, electronically, mechanically, by photocopying, recording or otherwise, without the prior written permission of the author. For articles published, the copyright has been transferred to the respective publisher.
Economic Evaluations in Healthcare: Methodological Considerations and Applications on Cost-Effectiveness of ACE Inhibitor Treatment for Renal Disease Prevention

DISSEPTION

to obtain the degree of Doctor at Maastricht University, on the authority of the Rector Magnificus Prof. dr. L.L.G. Soete, in accordance with the decision of the Board of Deans, to be defended in public on Thursday 10 September 2015, at 10:00 hours

by

Charles Christian Adarkwah-Yiadom
SUPERVISOR

Prof. dr. S.M.A.A. Evers

CO-SUPERVISOR

Dr. M. Hiligsmann

ASSESSMENT COMMITTEE

Prof. dr. C.D. Dirksen (chair)
Dr. H.-P. Dauben (DIMDI, Cologne, Germany)
Prof. dr. W. Karges (University of Aachen, Germany)
Dr. B.L.T. Ramaekers
Prof. dr. N.C. Schaper
DEDICATED TO MY
BELOVED PARENTS
Content

Chapter 1 Introduction and aim of the thesis 9

PART I Methodological aspects 23

Chapter 2 Should cost-effectiveness analysis include the cost of consumption activities? An empirical investigation. Health Econ. 2015 Feb 15. doi: 10.1002/hec.3162. 25


Chapter 4 Risk of bias in model-based economic evaluations: the ECOBIAS checklist. submitted to Expert Rev Pharmacoecon Outcomes Res 67

PART II Economic evaluation studies 91


Chapter 8 General discussion and implications 161

Valorisation 177
Summary 185
Acknowledgement 193
Curriculum Vitae 197
Chapter 1

General introduction and outline
Introduction

Burden of renal disease

There has been a notable increase in the incidence and prevalence of patients undergoing renal replacement therapy (RRT) in Western countries in recent years.\(^1\)\(^-\)\(^5\) RRT encompasses dialysis and renal transplantation. In Germany, more than 80,000 people currently receive RRT,\(^6\) while more than 15,000 patients require RRT in the Netherlands, where the proportion of transplant patients is continuously increasing and represents almost 60% of all patients in this population.\(^7\) End stage renal disease (ESRD) can be defined by either the need for renal transplantation or long-term dialysis.\(^8\) It is associated with declined quality of life, as patients have to come in to see their doctor on a regular basis, up to three times a week for dialysis or for additional blood tests associated with either dialysis or renal transplantation. ESRD is a final outcome of a gradual decline of renal function due to various reasons, such as diabetes or hypertension. ESRD is not only a huge burden from the patients’ perspective, but also for healthcare systems, with a proportion of the national expenditures devoted to ESRD in European countries ranging from 0.7% in the UK to 1.8% in Belgium.\(^9\)\(^,\)\(^10\) From a health care perspective, the costs of ESRD treatment amount to €42,000 per patient per year in the Netherlands,\(^9\)\(^,\)\(^11\)\(^,\)\(^12\) and up to €45,000 in Germany,\(^13\) respectively. The prevention of ESRD is therefore not only important from patient and medical perspective, but also from an economic viewpoint. There are various reasons for renal insufficiency, among those diabetes, vascular nephropathy and different forms of inflammatory kidney diseases (nephritides).\(^14\)\(^,\)\(^15\) Diabetes is considered the most common cause for ESRD in most Western countries, accounting for about 40% of all ESRD cases.\(^6\)\(^,\)\(^16\)\(^,\)\(^17\) In addition, the prevalence of especially type 2 diabetes and its secondary complications is expected to rise due to an aging population and a growing obesity rate.\(^18\)\(^-\)\(^21\) As a logical consequence, higher costs for diabetes treatment itself and especially for treatment of secondary complications (such as nephropathy, retinopathy or neuropathy) will challenge healthcare systems in the future.

Looking at this patient population in terms of quality of life, a cross-sectional study showed that diabetes is associated with a decrease in quality of life. Here, adult diabetic patients (n=292) with a disease duration of at least one year and a mean age of 62 years (range 21 – 85) were interviewed using the time tradeoff (TTO) method.\(^22\) Arnesen et al.\(^23\) investigated quality of life in
ESRD patients, while Hoerger et al. did a survey on patients with advanced renal insufficiency using the time tradeoff (TTO) method in 65 patients, and they both demonstrated a significant loss of utility in these states of disease. Renal disease, i.e. nephropathy, leads to a gradual decline of the renal function and is initially characterized by the presence of protein in the urine, which is called micro- or macroalbuminuria, depending on the total amount. Creatinine is a lab marker representative for the renal function. A rise in the creatinine level indicates a decline of the renal function and can therefore provide evidence of advanced renal insufficiency. Angiotensin-converting enzyme (ACE) inhibitors, which are drugs initially approved for hypertension or heart failure treatment, slow down the progression of diabetic nephropathy independent of elevated blood pressure. Angiotensin II receptor blockers (ARBs) have similar effects on renal outcomes, but are more expensive. Both drug classes were also proven effective in delaying the onset of ESRD in non-diabetic nephropathy patients.

Health technology assessment (HTA) and economic evaluation

Health technology assessment (HTA) provides a broad perspective on the evaluation of new or existing health technologies that in principle can cover every activity in healthcare that is patient orientated. HTA can be seen as a multi-disciplinary field of policy analysis that studies the medical, social, ethical, as well as the economic implications of the development, diffusion and use of health technology. From the assessment of individual patient interventions (such as a specific drug), to disease management and population based care programmes, the purpose of HTA is to provide decision makers with the information they need.

With respect to HTA, technology assessment is not limited to new drugs, screening or diagnostic activities in healthcare, but it also includes evaluation of the organization of care and its infrastructure. For instance, HTA can be seen as a bridge between scientific evidence and policy decision-making. HTA and cost-effectiveness analyses are becoming more and more important as healthcare costs are increasing in Western countries and resources are scarce. This means that healthcare must compete financially with other
sectors, such as education, housing and national defense. In addition, there are new and more expensive interventions, creating a tension between clinical treatment possibilities and scarce financial resources. As a result, the demand for evidence based on economic evaluation is increasing, e.g. for legislation regarding the reimbursement of pharmaceuticals and other interventions (e.g. NICE), and cost information is being demanded by research funding organizations.

It would be impossible for a society to funnel the majority of its resources towards healthcare and to finance all treatment options that provide some benefit without experiencing significant and unacceptable sacrifices in other important social sectors. Societies have to set priorities when deciding what gets subsidized.

One focus of HTA is on performing economic evaluation; it is only one part, but is a significant part, of HTA. Economic evaluation can be seen as the “comparative analysis of alternative courses of action in terms of both their costs and consequences”. The basic goals of an economic evaluation are to compare the costs with the consequences of the alternatives that are considered, in order to help decision makers to efficiently allocate scarce resources. According to Drummond et al., there are basically four types of full economic evaluations: cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis and cost-benefit analysis. A cost-minimization analysis (CMA) compares alternative strategies where all relevant outcomes are equal (i.e. equal patient quality of life). A cost-effectiveness analysis (CEA) compares alternatives and measures (in natural units) the primary objectives of an intervention (e.g. life-years saved). A cost-utility analysis (CUA) compares alternatives similar to a CEA, but uses a more generic outcome. Hence, a cost-benefit analysis (CBA) compares alternatives by using monetary outcomes (e.g. euros).

In a CEA, for example, the incremental effectiveness of an intervention is quantified and compared with its incremental costs (Figure 1.1).
The results of an economic evaluation are often presented in an aggregated way, using an incremental cost-effectiveness ratio (ICER). This ICER represents how much it will cost to improve one point on an outcome scale. The ICER is calculated as follows:

\[
\text{ICER} = \frac{C_A - C_B}{E_A - E_B},
\]

where \(C_A\) is the total cost of group A, \(C_B\) is the total cost of the comparator group, \(E_A\) is the effect at follow-up for group A and \(E_B\) is the effect at follow-up for the comparator group.

Be aware that the numerator, the \(C\), includes the resources saved by the intervention or the comparator. So \(C\) stands for costs minus savings in both groups. In a CEA, the main endpoint used in the denominator of the ICER can be any ordinal measure, i.e. any morbidity and mortality outcome. In a CUA, utilities or QALYs are used in the denominator.\(^3\)\(^4\)

Basically, there are two approaches to performing an economic evaluation study: a trial-based economic evaluation study, and a model-based economic evaluation study. A trial-based economic evaluation study is mostly piggy-backed onto an ongoing trial, in which input (costs) and output (effects) are measured and valued. In a model-based economic evaluation, data from a wide range of sources (RCTs, observational studies, trial-based economic evaluations, etc.) are synthesized using an economic model. These two
methods of economic analysis are seen as complementary, and each has its own advantages and disadvantages. 

A model-based approach can be interesting when accounting for scenarios where the natural history of a disease lasts for years, meaning it covers a period of time which is much longer than a clinical trial could cover in a real life setting. The progression of renal disease is a good example of a medical condition which might develop over decades.

Hence, looking at cost-effectiveness of ACE inhibitors in renal disease, we opted for a model-based approach to account for the long timeframe and disease duration. Modelling studies often use Markov models. A Markov model is an iterative process where patients are assumed to stay in one cycle (i.e. a defined health state) for a certain amount of time and then make a transition to another cycle. Those models are useful when a decision problem involves risk that is continuous over time, when the timing of events is important and when important events may happen more than once. Different clinical strategies and time-points for starting ACE inhibitor therapy were compared, and the models were constructed with the best available evidence regarding input parameters such as preference weights, transition probabilities, costs, etc. The outcomes were measured by means of costs per quality adjusted life year (QALY).

To calculate QALYs, preference weights, among other data, are needed. In health economics, the assessment of quality of life plays a major role, as outcomes are not only assessed in terms of enhanced survival (i.e. additional years to life), but also enhanced quality of life (QoL). In detail, health-related QoL analyses measure the impact of treatments on the physical, social and emotional aspects of life from the patients’ perspective. There are many techniques available to assess QoL, e.g. by using questionnaires such as the EQ-5D or the SF-36.

Besides the questionnaires mentioned above, there are other methods to assess the QoL, among those the time trade-off (TTO) method. The TTO is the most commonly used method to elicit quality-of-life weights for QALYs. The TTO technique determines the proportion of remaining life years in poor health that patients are willing to give up or trade in exchange for perfect health. Based on patient responses, utility scores are calculated. The TTO method is essentially relevant to our analyses and is further investigated in this thesis, as there is still a debate if costs for consumption and leisure
activities should be included in the ICER. Those utility measures in economic evaluations are becoming more and more important given the fact that decision makers are asked to optimize the allocation of scarce healthcare resources across disease areas and patient groups.\textsuperscript{45}

Conducting an economic evaluation, whether it is a trial-based or a model-based one, is not without precariousness. Hence, a large number of guidelines exist that set parameters regarding how to perform a good and valid economic evaluation.\textsuperscript{36,44-49} Nevertheless, even following a guideline does not prevent various types of bias, as guidelines show what to do but not exactly how to do it and in which way. Investigating how to minimize bias in trial- and model-based economic evaluation would be very worthwhile for both researchers and policymakers.

Objectives of the thesis

This thesis has two different aims: first to assess ACE inhibitors and ARBs in renal disease, i.e. nephropathy, in various patient populations and second to contribute to the clarification of methodological aspects in terms of bias and quality of life as measured using the TTO method.

These aims can be further divided into three objectives:

1. To assess whether costs for consumption and leisure activities need to be considered in the incremental cost-effectiveness ratio (ICER) in cost-effectiveness analyses. (Part Ia – methodological aspects)

2. To review evidence on bias types in trial-based as well as model-based economic evaluations (Part Ib – methodological aspects)

3. To assess the cost-effectiveness of ACE inhibitors and angiotensin II receptor blockers in diabetic as well as non-diabetic renal disease in different countries (Part II – economic evaluation studies)

Outline of the thesis

Figure 1.2 gives an overview of the outline. The first part, which is covered in Chapters 2 to 4, focuses on methodological aspects. Chapter 2 deals with the consideration of the costs of consumption and leisure time activities when
using the quality-adjusted life year as a measure of health outcome from a societal perspective in cost-effectiveness analysis. The purpose of this study was to investigate whether the effects of ill health on consumptive activities are spontaneously considered in a health state valuation exercise and how much this matters. Our survey included patients with inflammatory bowel disease in Germany (n=104). Patients were randomly provided or not provided with explicit instructions regarding the consideration of consumption and leisure effects in a time trade-off (TTO) exercise. The decision against performing a study on diabetic patients, or even patients with renal insufficiency, was made because a population of young patients with relatively high disease burden was needed to answer our research question. This neither accounts for the diabetic nor for the renal disease population covered in the economic evaluation studies (Chapters 5 to 7). The main focus of Chapter 2 is not on the disease itself, but on how patients consider loss of consumption and leisure activities due to a chronic disease when being confronted with a TTO exercise, which is the method of assessing quality of life applied in the economic evaluation studies.

Chapter 3 focuses on bias forms in trial-based economic evaluations. This chapter gives an overview of the risks of bias in trial-based economic evaluations. Key sources for bias are identified and strategies are shown for how key sources for bias can be revealed and overcome (i.e. what bias-reducing strategies might be employed) in future trial-based economic evaluations. The biases identified in Chapter 3 do also account for model-based studies. Hence, Chapter 4 focuses on specific bias forms in model-based economic evaluation. It is obvious that in model-based economic evaluation some other types of bias may occur apart from the bias types presented and discussed in Chapter 3. We aim to report on specific bias forms related to model-based economic evaluation, to illustrate potential problems and to show how those biases might influence the results of economic evaluations. These model-specific biases are related to structural assumptions, the model type, time horizon or data selection, etc. Finally, a checklist on bias in economic evaluation (ECOBIAS checklist) is presented, merging results from Chapters 3 and 4.

The next three chapters, Chapters 5 to 7, cover modelling studies on the cost-effectiveness of ACE inhibitors in renal disease. Chapter 5 assesses the cost-effectiveness of ACE inhibitors and angiotensin II receptor blockers in newly
diagnosed type 2 diabetes patients in Germany. A sample of newly diagnosed type 2 diabetic patients was chosen and a Markov model was constructed, simulating the progression through all stages of albuminuria, up to ESRD and finally death. Here, a calculation of ESRD costs for the German setting was also performed. Chapter 6 focuses on the Dutch setting and the prevention of diabetic nephropathy. In the Netherlands, not only costs but also quality of life parameters differ from the German setting. Furthermore, treatment regiments are quite different. For example, a higher percentage of patients with ESRD receive a renal transplantation, while in Germany a higher percentage receives dialysis. Chapter 7 deals with the cost-effectiveness of ACE inhibitors in non-diabetic advanced renal disease in the Netherlands. Although diabetes is the most common cause for ESRD in the Western world, there are many other possible causes as mentioned above. There is a debate on whether patients, especially those with advanced renal disease, should be treated with ACE inhibitors, as this is not without the risk of side effects. Hence, these patients are likely to be underexplored, at least in terms of cost-effectiveness studies. So, we built a lifetime-model, which was used by patients with non-diabetic advanced renal disease that followed to ESRD and finally death; one group was treated with ACE inhibitors and the other group was not.

All models are based on a Markov model, and various sensitivity analyses, univariate as well as multivariate, were performed. Finally, Chapter 8 discusses the research findings presented in the previous chapters with respect to their implications for practice, future research and policy-making.
Chapter 1: General introduction and outline of the thesis

Part I – Methodological aspects
- Chapter 2: Consideration of the consumption costs in cost-effectiveness analysis
- Chapter 3: Bias in trial-based economic evaluation
- Chapter 4: Bias in model-based economic evaluation

Part II – Economic evaluation studies
- Chapter 5: Cost-effectiveness of ACE inhibitors in type 2 diabetes in Germany
- Chapter 6: Cost-effectiveness of ACEI for diabetic nephropathy prevention in NL
- Chapter 7: Cost-effectiveness of ACEI in non-diabetic advanced renal disease

Chapter 8: General discussion

Figure 1.2 Outline of the thesis
References


Part I

Methodological aspects
Chapter 2

Should cost-effectiveness analysis include the cost of consumption activities? An empirical investigation

Adarkwah CC
Sadoghi A
Gandjour A

*Health Econ* 2015 Feb 15. doi: 10.1002/hec.3162
Abstract

Objective
There has been a debate on whether cost-effectiveness analysis should consider the cost of consumption and leisure time activities when using the quality-adjusted life year as a measure of health outcome under a societal perspective. The purpose of this study was to investigate whether the effects of ill health on consumptive activities are spontaneously considered in a health state valuation exercise and how much this matters.

Methods
The survey enrolled patients with inflammatory bowel disease in Germany (n=104). Patients were randomized to explicit and no explicit instruction for the consideration of consumption and leisure effects in a time trade-off (TTO) exercise.

Results
Explicit instruction to consider non-health-related utility in TTO exercises did not influence TTO scores. However, spontaneous consideration of non-health-related utility in patients without explicit instruction (60% of respondents) led to significantly lower TTO scores.

Conclusions
Results suggest an inclusion of consumption costs in the numerator of the cost-effectiveness ratio, at least for those respondents who spontaneously consider non-health-related utility from treatment. Results also suggest that exercises eliciting health valuations from the general public may include a description of the impact of disease on consumptive activities. Given the importance of this question for the conduct of cost-effectiveness analysis in health care, confirmation in additional studies that are conducted outside Germany and consider other health-state valuation techniques and diseases is recommended.
Introduction

There has been a debate on whether cost-effectiveness analysis should include health-related costs in added life years that are unrelated to the disease or intervention in question. Several authors have argued in favor of this inclusion.\textsuperscript{1-5} A perhaps even more controversial debate has been around the question whether cost-effectiveness analysis should consider the cost of consumption and leisure time activities when using the quality-adjusted life year (QALY) as a measure of health outcome under a societal perspective. For example, Meltzer\textsuperscript{1} argues that consumption costs and earnings during added years of life should be included in cost–utility analyses. Yet, Nyman\textsuperscript{2} suggests otherwise based on the argument that none of the existing direct or indirect utility measures explicitly recognizes variations in consumption or foregone leisure. The underlying principle is that costs should be excluded from cost–utility analyses if they represent resources that produce utility that is not being measured in the denominator of the cost–utility ratio.

For life-extending treatments, however, handling of the cost of consumption and leisure time activities seems less controversial. As satisfying primary needs such as food, shelter, and clothing contributes to the additional survival time and the associated utility, the costs associated with these resources should also be included in cost–utility analyses, according to Nyman’s principle.\textsuperscript{3,6} For a similar notion see Liu et al. (2012, p. 494).\textsuperscript{7}

For quality-of-life enhancing interventions some empirical studies have been conducted on how much the effects of ill health on leisure time matter in health state valuations. In the following we describe some of the most recent empirical evidence on leisure time effects. Krol et al. conducted a survey on 222 people of the general public to assess the incorporation of leisure in time trade-off (TTO) exercises.\textsuperscript{8} They found that 88% of the respondents spontaneously included the effects of ill health on leisure time. They also showed that respondents including leisure time gave significantly lower TTO values to one out of 3 health states. The authors conclude that it may be necessary to explicitly instruct respondents to include the effects of ill health on leisure time. Additional research on the effects of ill health on leisure time was conducted using a visual analogue scale (VAS) among non-patients.\textsuperscript{9-11}

To the best of our knowledge, for the effects of ill health on consumptive activities no such empirical evidence has been gathered so far. Therefore, the
The purpose of our study was to conduct an empirical survey on whether the effects of ill health on consumptive activities are spontaneously considered in a health state valuation exercise and how much this matters. To this end, participants were randomized to explicit and no explicit instruction for the consideration of consumption and leisure effects. The investigation thus aimed at providing insight into the question whether cost-effectiveness analysis should consider the cost of consumption and leisure time activities in the numerator of the cost-effectiveness ratio.

As a health-state valuation technique we used the TTO questionnaire. The TTO is the most commonly used method to elicit quality-of-life weights for QALYs. Also, it is the most commonly used technique in EQ-5D valuation studies.

In contrast to the above-mentioned studies, we recruited actual patients and not representatives of the general population because only patients have the necessary experience to judge how disease affects consumption and leisure time activities. We note that arguments have been brought forward to justify the valuation of health states by community members, for example, the popular argument that community members should have a say in resource allocation. The question of whose preference assessment should count when calculating QALYs for the purpose of economic evaluations is intensely debated in the health economics literature. Yet, regardless of one’s position in this debate, our study is also relevant for the elicitation of population preferences. That is, if patients considered non-health-related disutility from disease to be a significant factor in the valuation of health states, then exercises eliciting health valuations from the general public may also include a description of the impact of disease on consumptive activities.

We enrolled patients with inflammatory bowel disease (IBDs) because they are relatively young and therefore perhaps more aware of forgone consumption and leisure opportunities in life.

IBDs are chronic diseases of the gastrointestinal tract. IBD is a generic term for Crohn’s Disease (CD) and Ulcerative Colitis (UC). It is most prevalent in Europe. IBDs are characterized by episodes of remission and relapses, mostly associated with intestinal symptoms, which include bloody stools, abdominal pain and diarrhea, as well as extra-intestinal symptoms, ie, weight loss, fatigue, and joint-pain. These factors all lead to a decrease of patient’s
well-being and ability to work and also affect a patient’s leisure-time and consumption activities.

Methods

The survey was conducted between March 2012 and February 2013. Patients were recruited at two German university hospital outpatient clinics specialized in IBD treatment, where patients come in either on a regular basis for routine follow-up visits or show up in case or worsening of the disease, i.e., a flare or any sort of complication. Thus, patients could be identified in advance to assure that they do meet the inclusion criteria. No patient refused to participate in the study.

Patients were eligible if they were between 18 and 45 years of age and had a memorable disease flare within the last two years of at least moderate activity, measured either by the Crohn’s Disease Activity Index (CDAI) (>220) or the Mayo Score for UC patients (>6). After signing the informed consent patients were randomized into two groups (A and B). Randomization was computer-generated, with allocation concealment by opaque sequentially numbered sealed envelopes.

Patients in both groups were individually instructed to fill out a questionnaire while a physician or nurse was available for questions. In both questionnaires respondents were first asked some background questions about age, gender, profession, family status, disease duration, time from last flare, Beck Depression Inventory (BDI) score (an instrument for measuring the severity of depression), and disease activity indices.

Patients in group A were then confronted with the TTO question first. Respondents were asked how many years (out of 10) in the given health state with possible flares and complications they would be willing to give up in order to be in full health. Patients were informed to trade off months or weeks as well. Afterwards, patients were asked if consequences of complete healing on their life circumstances had been considered. If yes, patients were instructed to write down their thoughts.

Patients in group B were asked upfront what they would do differently regarding leisure-time and consumption activities as well as family and job in case they would become healthy. The TTO question as described above was
asked afterwards with the instruction to explicitly consider changes concerning leisure-time and consumption activities as well as family and job. In order to adjust TTO scores for imbalances in participant characteristics, we performed a multiple regression analysis using Stata 12 (StataCorp LP, College Station, TX, USA). Independent variables were the version of the questionnaire, age, gender, profession, family status, disease duration, time from last flare, BDI score, and disease activity indices. Gender, profession and family status were coded as categorical variables. To avoid overfitting we used a simple dichotomy for profession and family status. We considered $p<0.05$ to be statistically significant. Details of the statistical approach are shown in the appendix.

**Results**

A total of 104 patients participated in the study. See Table 2.1 for characteristics of study participants. Except for one data item concerning age and three data items concerning consequences of complete healing on life circumstances, questionnaires were filled out completely. Figure 2.1 estimates the probability density functions of TTO scores by the Kernel method. It shows that both groups have a similar distribution of TTO scores. The two-sample Kolmogorov-Smirnov test for equality of distribution functions showed no difference ($p=1.000$). The null hypothesis of meeting the Tobit model assumptions was not rejected. Table 2.2 and 2.3 show the results of the Tobit and the marginal Tobit model using TTO score as a dependent variable. Explicit instruction to consider consumption and leisure effects did not have a significant impact on the TTO score, as shown by the first rows of Tables 2.2 and 2.3. Table 2.4 shows corrected Tobit model coefficients using the formula suggested by Pullenayegum et al.24 However, among patients who were not explicitly instructed to consider consumption and leisure effects (group A) consideration of consumption and leisure effects (60% of respondents in group A) significantly reduced the TTO score in both the Tobit and the marginal Tobit model ($p=0.037$ and $p=0.033$, respectively). The corrected Tobit model coefficient for consideration of consumption and leisure effects using the formula suggested by Pullenayegum et al. was 0.26, meaning that consideration of consumption and leisure effects reduced the
TTO score by 0.26 (the unadjusted difference was 0.14). According to the qualitative statements by respondents, this disutility was not driven by changes in expected earnings. Furthermore, in all models BDI score and age had a significant impact on the TTO score. That is, a higher depression score and lower age were associated with a lower TTO score.

Table 2.1 Characteristics of study participants (n=104).

<table>
<thead>
<tr>
<th></th>
<th>Group A (no explicit instruction)</th>
<th>Group B (explicit instruction)</th>
<th>P value for difference of means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number</td>
<td>50</td>
<td>54</td>
<td>0.840</td>
</tr>
<tr>
<td>Male (%)</td>
<td>50 (male)</td>
<td>52 (male)</td>
<td></td>
</tr>
<tr>
<td>Mean age</td>
<td>27</td>
<td>29</td>
<td>0.191</td>
</tr>
<tr>
<td>Family status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>single</td>
<td>70</td>
<td>70</td>
<td>1.000</td>
</tr>
<tr>
<td>married</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>divorced</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>separated</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Profession (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employed</td>
<td>74</td>
<td>74</td>
<td>0.705</td>
</tr>
<tr>
<td>not job seeking</td>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>unemployed</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>student</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>retired</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BDI score</td>
<td>9.5</td>
<td>13</td>
<td><strong>0.029</strong></td>
</tr>
<tr>
<td>Mean duration of disease (month)</td>
<td>35</td>
<td>36</td>
<td>0.854</td>
</tr>
<tr>
<td>Mean time from last flare (weeks)</td>
<td>50</td>
<td>54</td>
<td>0.825</td>
</tr>
</tbody>
</table>

*p* values marked in bold are significant at a 0.05 level. BDI = Beck Depression Inventory
Figure 2.1 Probability density functions of time trade-off (TTO) scores estimated by the Kernel method. Comparison groups were explicit and no explicit instruction to consider changes in leisure-time and consumption activities.

Table 2.2 Results of the Tobit model analyzing the relationship between time trade-off score and individual variables.

|                       | Coefficient | Standard error | t    | p>|t | 95% CI (lower bound) | 95% CI (upper bound) |
|-----------------------|-------------|----------------|------|----|----------------------|----------------------|
| Version*              | 0.036       | 0.082          | 0.43 | 0.665 | -0.128               | 0.200                |
| CDAI score (current)  | -0.001      | 0.001          | -1.29 | 0.199 | -0.002               | 0.000                |
| CDAI score (last flare)| -0.001      | 0.001          | -0.63 | 0.532 | -0.002               | 0.001                |
| Mayo Score (current)  | -0.023      | 0.027          | -0.85 | 0.395 | -0.078               | 0.031                |
| Mayo Score (last flare)| -0.011      | 0.024          | -0.44 | 0.658 | -0.058               | 0.037                |
| BDI score             | -0.017      | 0.004          | -3.81 | 0.000 | -0.025               | -0.008               |
| Age                   | 0.016       | 0.007          | 2.17  | 0.033 | 0.001                | 0.031                |
| Gender                | 0.048       | 0.078          | 0.62  | 0.539 | -0.107               | 0.204                |
| Family status         | 0.080       | 0.097          | 0.83  | 0.409 | -0.112               | 0.272                |
| Profession            | -0.103      | 0.100          | -1.03 | 0.305 | -0.302               | 0.096                |
| Duration of disease   | -0.004      | 0.008          | -0.51 | 0.611 | -0.020               | 0.012                |
| Time from last flare  | 0.001       | 0.001          | 0.89  | 0.375 | -0.001               | 0.003                |
| Constant              | 0.604       | 0.328          | 1.84  | 0.068 | -0.046               | 1.255                |

*p values marked in bold are significant at a 0.05 level. *Dummy variable where 1 = no explicit instruction for the consideration of consumption and leisure effects; 0 if otherwise; CI = confidence interval; CDAI = Crohn’s Disease Activity Index; BDI = Beck Depression Inventory.
Table 2.3 Results of the marginal effects Tobit model analyzing the relationship between time trade-off score and individual variables

| Variable                        | Coefficient | Standard Error | z     | p>|z| | 95% CI (Lower Bound) | 95% CI (Upper Bound) |
|--------------------------------|-------------|----------------|-------|-----|----------------------|----------------------|
| Version                        | 0.032       | 0.074          | 0.430 | 0.666 | -0.112               | 0.176                |
| CDAl score (current)            | -0.001      | 0.001          | -1.290| 0.197| -0.002               | 0.000                |
| CDAl score (last flare)         | 0.000       | 0.001          | -0.630| 0.531| -0.002               | 0.001                |
| Mayo Score (current)            | -0.021      | 0.024          | -0.850| 0.394| -0.069               | 0.027                |
| Mayo Score (last flare)         | -0.009      | 0.021          | -0.440| 0.658| -0.051               | 0.032                |
| BDI score                       | -0.015      | 0.004          | -3.700| 0.000| -0.023               | -0.007               |
| Age                             | 0.014       | 0.007          | 2.210 | 0.027| 0.002                | 0.027                |
| Gender                          | 0.043       | 0.069          | 0.620 | 0.536| -0.093               | 0.178                |
| Family status                   | 0.070       | 0.083          | 0.850 | 0.398| -0.093               | 0.234                |
| Profession                      | -0.093      | 0.091          | -1.020| 0.307| -0.270               | 0.085                |
| Duration of disease             | -0.004      | 0.007          | -0.510| 0.607| -0.017               | 0.010                |
| Time from last flare            | 0.001       | 0.001          | 0.910 | 0.365| -0.001               | 0.002                |

* Dummy variable where 1 = no explicit instruction for the consideration of consumption and leisure effects; 0 if otherwise; CI = confidence interval; BDI = Beck Depression Inventory; CDAl = Crohn’s Disease Activity Index.

Table 2.4 Corrected Tobit model coefficients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>0.030</td>
</tr>
<tr>
<td>CDAl score (current)</td>
<td>-0.001</td>
</tr>
<tr>
<td>CDAl score (last flare)</td>
<td>0.000</td>
</tr>
<tr>
<td>Mayo Score (current)</td>
<td>-0.020</td>
</tr>
<tr>
<td>Mayo Score (last flare)</td>
<td>-0.009</td>
</tr>
<tr>
<td>BDI score</td>
<td>-0.014</td>
</tr>
<tr>
<td>Age</td>
<td>0.014</td>
</tr>
<tr>
<td>Gender</td>
<td>0.040</td>
</tr>
<tr>
<td>Family status</td>
<td>0.066</td>
</tr>
<tr>
<td>Profession</td>
<td>-0.090</td>
</tr>
<tr>
<td>Duration of disease</td>
<td>-0.003</td>
</tr>
<tr>
<td>Time from last flare</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* Dummy variable where 1 = no explicit instruction for the consideration of consumption and leisure effects; 0 if otherwise; BDI = Beck Depression Inventory; CDAl = Crohn’s Disease Activity Index.
Discussion

Explicit instruction to consider non-health-related utility in TTO exercises did not influence TTO scores. However, spontaneous consideration of non-health-related utility in the group without explicit instruction (60% of respondents of this group) led to significantly lower TTO scores. Hence, the first finding could be explained by spontaneous consideration of non-health-related utility even without explicit instruction (thus offsetting the effect of explicit instruction). Spontaneous consideration of non-health-related utility thus leads to a utility gain from treatment compared to no spontaneous consideration. This is consistent with the notion that health gains can increase the marginal utility of consumption. In turn, this would justify an inclusion of consumption costs in the numerator of the ICER not only for life extending but also for quality-of-life enhancing interventions, at least for those respondents who spontaneously consider non-health-related utility from treatment. Results also suggest that exercises eliciting health valuations from the general public for the purpose of conducting economic evaluations from a societal perspective may include a description of the impact of disease on consumptive and leisure activities.

An alternative approach not considered in this paper would be to instruct patients to exclude non-health-related utility. However, this approach was indirectly considered as it is equivalent to the situation of no explicit instruction and no spontaneous consideration of non-health-related utility.

One strength of our study is that we randomized patients to the two groups, thus controlling for omitted variable bias. Still, we would like to point out several limitations. First, results are based on a single chronic disease group. Therefore, it is not known whether results are transferable to other diseases. On the other hand, mixing patients with different diseases in one study would have introduced other problems, e.g., the question whether results from such a study would hold for every disease. Second, our sample purposely included relatively young patients and while we adjusted for age in our analysis, another sample with more elderly patients may provide a different result. Third, our sample size may have lacked the power to detect additional significant associations. Finally, results may not be transferable from country to country.
Given the importance of this question for the conduct of cost-effectiveness analysis in health care, confirmation in additional studies that are conducted outside Germany and consider other health-state valuation techniques and diseases is recommended. Future research may also include elderly patients.
Appendix – Statistical Approach

Many measures of preferences have the property that a proportion of the population achieves the upper bound of 1. Therefore, the utility distribution will be non-normal and linearity of utility–covariate associations will be questionable. The Tobit model provides a consistent and efficient alternative for estimating such data with a bounded nature. Yet, the Tobit model is sensitive to the assumptions of normality and homoscedasticity of residuals. Both heteroscedasticity and non-normality therefore cause inconsistency in the Tobit estimator.

The Tobit model assumes that there is a latent (i.e., unobservable) variable that can extend below zero and above 1 (i.e., there is no bounding at zero and 1) but has been observed subject to censoring at zero and 1. Because of the latent variable extending below zero, the simple Tobit model does not allow interpreting the effect of independent variables on the dependent variable. The marginal effects Tobit model is required to estimate the coefficients and is conditional upon the latent variable being at or above zero. To this end, the constant is dropped.

When preference weights are assumed to be bounded at 1, meaning that preference weights cannot exceed 1, a formula to calculate corrected coefficients for the Tobit model has been recently published. We note that there is a debate about whether preference weights are, in fact, bounded at 1. It has been argued that the definition of full health (i.e., a preference weight of 1) depends on the type of preference-based questionnaire used and therefore leaves the possibility of scores above 1 open.

In order to check for bias in the Tobit estimator we tested the hypothesis of homoscedasticity and normality of residuals. To this end, we performed the conditional moment (CM) test.
Cost-effectiveness of angiotensin-converting enzyme inhibitors

References

Chapter 2


Chapter 3

Risk of bias in trial-based economic evaluations: Identification of sources and bias-reducing strategies

Evers SM
Hiligsmann M
Adarkwah CC

Psychol Health 2015;30:52-71
Chapter 3

Abstract

Objective
The objectives of this article are first to give an overview of the risks of bias in trial-based economic evaluations and, second, to identify how key sources for bias can be revealed and overcome (i.e. what bias-reducing strategies might be employed) in future trial-based economic evaluations in the field of health psychology.

Design
Narrative review discussing sources of bias in trial-based economic evaluations and bias-reducing strategies.

Results
We identified 11 biases and assigned them to a particular trial phase. A distinction is made between pre-trial biases, biases during the trial and biases that are relevant after the actual trial. All potential forms of bias are discussed in detail and strategies are shown to detect and overcome these biases.

Conclusion
In order to avoid bias in trial-based economic evaluations, one has to be aware of all the possible forms of bias. All stakeholders have to examine trial-based economic evaluations in a rigorous and stringent manner. This article can be helpful in this examination as it gives an overview of the possible biases which researchers should take into account.
Introduction

In the current economic climate, it is of crucial importance to prove that innovative health behaviour change interventions are not only effective, but also cost-effective. Economic evaluation studies compare the costs and outcomes of an innovative health behaviour change intervention with a control intervention, mostly care as usual.

Basically, there are two approaches to performing an economic evaluation study: a trial-based economic evaluation study, and a model-based economic evaluation study. A trial-based economic evaluation study is mostly piggy-backed onto a behavioural/educational trial. In general this means that a resource-use measure (most often a cost questionnaire; for a repository see www.DIRUM.org) and a quality of life instrument (most often the EuroQol questionnaire\(^1\) or the Short-Form \(^2\) are included in the prospective behavioural/educational trial design, in order to calculate the cost-effectiveness of the health behaviour change intervention. In a model-based economic evaluation, data from a wide range of sources (behavioural/educational trials, observational studies, trial-based economic evaluations, etc.) are synthesized using an economic model. These two methods of economic analysis are seen as complementary\(^3\) and each has its own advantages and disadvantages.\(^4,6\)

As this special issue for *Psychology & Health* focuses on the risk of bias in randomized controlled trials (RCTs) of health behaviour change interventions,\(^7\) this article will focus on trial-based economic evaluations. A major strength of trial-based economic evaluations is that the data are collected alongside an already existing behavioural/educational trial, making it feasible to do an economic evaluation without investing a lot of additional research funding. Furthermore, both costs and outcomes are collected in the same population, making a head-to-head comparison more viable. The discipline of trial-based economic evaluation has a strong tradition in epidemiology, where the RCT is regarded as the gold standard.\(^8\) Although evidence from RCTs is very powerful, it is often said that the RCT design has limitations when it comes to the generalisability of economic evaluations,\(^9,10\) especially when looking at behavioural/educational interventions. In these public health interventions, costs and benefits are very broad,\(^10\) and the context and setting are of critical importance, due to variation and
uncertainty in diagnosis and treatment. As a result, relying solely on evidence from RCTs can mean that de facto cost-efficacy rather than cost-effectiveness studies are produced, see Evers 2007. In addition to trial-based and model-based economic evaluations, qualitative methods of evaluation may help to address some of these limitations.

In addition to the choices which researchers have to make when designing a behavioural/educational trial, a number of choices have to be made by the economic evaluators: i.e. the perspective of the study, the comparator, the costs included, the outcomes measured, the valuation methods used, and the threshold used. There is evidence, that based on the choices of the economic evaluator, certain biases are induced which might influence the cost-effectiveness of the health behaviour change intervention.

A bias or a systematic error occurs when there is a difference between the true value (in the population) and the observed value (in the study) from any cause other than sampling variability. In other words, a bias is an effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average. A bias can be unintentional or intentional and can have either substantial or little impact.

In our personal experience, we have often noticed that, although evaluators are often aware of the current methods of trial-based economic evaluations, awareness of biases induced by these methods is, in our opinion, limited. To date, no study has given an explicit overview of additional biases identified in trial-based economic evaluations and the way these biases can be reduced. However, several studies have highlighted the importance of rigid methodology in trial-based economic evaluations, by highlighting methodological challenges, assessing methodological quality, or by developing guidelines for economic evaluation, see Pharmacoeconomic Guidelines Around The World at ispor.org/peguidelines or Ramsey et al. As far as we know, the study of Eble is the only study explicitly looking at the risk of bias in trial-based economic evaluations. However, Ebles’ study is limited to bias that the medical literature already has identified, and doesn’t focus on the additional bias induced by performing an economic evaluation within a trial.

The objectives of this article are first to give an overview of the risks of bias in trial-based economic evaluation and, second, to identify how the key sources
Risk of bias in trial-based economic evaluations

for bias can be revealed and overcome (i.e., what bias-reducing strategies might be employed) in future trial-based economic evaluation in the field of health psychology. The article does not provide novel methodological approaches in the field of trial-based economic evaluations, but highlights sources of potential biases, as well as some bias-reducing strategies. As not all readers of Psychology & Health might have knowledge of economic evaluations, this article will start with a small paragraph on how health behaviour change interventions can become cost-effective.

How does a health behaviour change intervention become cost-effective?

Whether an intervention ultimately becomes cost-effective depends on a number of choices made by the economic evaluator. In a trial-based economic evaluation, a behavioural or educational intervention is compared with one or more interventions, mostly care as usual. During the follow-up period, called ‘time horizon’ in economic evaluation, the costs are measured continuously. This means that all the relevant cost items of the behavioural/educational intervention as well as of the comparator are assessed during the total follow-up period of the trial. At the end of the trial the difference in costs between the behavioural/educational intervention and the comparator are related to the differences in effects. The results of an economic evaluation are often presented in an aggregated way, i.e. an incremental cost-effectiveness ratio (ICER). This ICER represents how much it will cost to improve one point on an outcome scale (for instance, a psychological measure or quality-adjusted life years (QALYs)). The ICER is calculated as follows: ICER = (C_i – C_c) / (E_i – E_c), where C_i is the annual total cost of the behavioural/educational intervention group, C_c is the annual total cost of the comparator group, E_i is the effect at follow-up for the behavioural/educational intervention group and E_c is the effect at follow-up for the comparator group. Be aware that the numerator, the C, includes the resources saved by the behavioural/educational intervention or the comparator. So C stands for costs minus savings in both groups. In a cost-effectiveness analysis (CEA) the main endpoint used in the denominator of the ICER can be any ordinal (psychological) measure, i.e. any morbidity and mortality outcome;
psychological outcome; physical, social, and mental functioning; nutritional status, etc. In most cases, the primary outcome of behavioural/educational trial is chosen to be the denominator of the ICER of the cost-effectiveness analysis. In a cost-utility analysis (CUA), utilities or quality-adjusted life years (QALYs) are used in the denominator. A utility is a preference or a value attached to health in which 0 represents death and 1 represents full health. Utilities can be derived using questionnaires such as the EQ-5D, or the SF-12 or SF-36. To exemplify, the EuroQol is a self-administered questionnaire, containing 5 dimensions of health-related quality of life, i.e. mobility, self-care, daily activities, pain/discomfort and depression/anxiety. Each dimension can be rated at three levels (or five levels): ranging from no problems to major problems. The 5 dimensions can be summed into a health state. Utility values can be calculated for these health states, using preferences elicited from a general population, the so-called valuation sets; these are available for the EuroQol and Short-Form. The QALY is based on the number of years of life that would be added by the behavioural/education intervention corrected for their utilities.

The height of ICER is thus defined by the difference in costs and the differences in effects between the behavioural/educational intervention and the comparator. If the ICER, the costs per QALY, is below a certain threshold, the health behaviour change intervention is regarded as cost-effective. Each jurisdiction has its own threshold, which varies between €12,000 and €80,000 per QALY depending upon the disease burden.

A behavioural/educational intervention is thus not cost-effective if: 1) the costs of the behavioural/educational intervention are too high in relation to the comparator, 2) the effects of the behavioural/educational intervention are too low in relation to the comparator, or 3) the ICER is unfavourable in relation to the threshold of the jurisdiction.

Academic disagreement between economic evaluators, as well as creative ways of influencing ICER, may lead to methodological uncertainty. Consequently, several countries have developed guidelines on how to perform an economic evaluation study, especially for pharmaceuticals. In addition, as mentioned before, a number of choices have to be made at every stage of the process of an economic evaluation, and these might lead to uncertainty. As a consequence each trial-based economic evaluation study
should include a sensitivity analysis to address any items about which the evaluator is uncertain.\textsuperscript{23}

Design

This article, which looks at common biases in economic evaluation, is based on a scoping review complemented with experiences of the authors. For the scoping review, a literature search was performed in July 2013 using PubMed (search terms (Costs and Cost Analysis [MeSH Terms] cost-effective and cost-effectiveness, based on and bias)\textsuperscript{24,25} and the NHS Economic Evaluation Database (search term bias in “any field” in NHS EED, retrievable via www.crd.york.ac.uk). The NHS EED focuses primarily on the economic evaluation of health care interventions, for which each week extensive literature searches are undertaken to identify relevant economic evaluations. It is known that combining PubMed and NHS EED is an optimal search strategy for economic evaluations.\textsuperscript{24,25} Furthermore, additional information was found by citation tracking and own knowledge. For inclusion, the paper had to deal explicitly with any kind of bias (in the aim or research question), which could occur when performing a trial-based economic evaluation. The selection process was performed by the first author.

Biases can occur in several phases within the trial (a classification scheme for this was developed by Pannucci\textsuperscript{26}: i.e. I, in the pre-trial phase, during the planning of the trial and study design; II, during the trial itself, in the data collection and analysis, and III, after the trial, in the publication phase of the economic evaluation study. As shown in Table 3.1, the results of this scoping review are placed in this classification overview.

Results

The search strategy resulted in 2,370 hits in PubMed and 1,860 hits in NHS EED. However, only 12 articles were relevant. The most common reason for exclusion was that bias was mentioned as part of a cost-effectiveness model, rather than bias being the subject under investigation.
### Table 3.1 Tips for avoiding different types of bias in trial-based economic evaluation.

<table>
<thead>
<tr>
<th>Type of bias (i)</th>
<th>Influence of the bias (ii)</th>
<th>Recommendation/Bias reducing strategies (iii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Pre-trial bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrow perspective bias</td>
<td>Omission of relevant costs and outcomes</td>
<td>State perspective explicitly; societal is</td>
</tr>
<tr>
<td></td>
<td></td>
<td>preferred, justification of narrower</td>
</tr>
<tr>
<td></td>
<td></td>
<td>perspective</td>
</tr>
<tr>
<td>Inefficient comparator bias</td>
<td>Overestimation of costs and outcomes</td>
<td>3-armed trial, including best alternative as</td>
</tr>
<tr>
<td></td>
<td></td>
<td>comparator, as well as care as usual, and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>describe all comparators sufficiently</td>
</tr>
<tr>
<td>Cost measurement omission bias</td>
<td>Overestimation of the incremental costs</td>
<td>Explicitly identify items relevant to</td>
</tr>
<tr>
<td>Intermittent data collection bias</td>
<td>Under- or overestimation of costs</td>
<td>disease and intervention studied</td>
</tr>
<tr>
<td>II Bias during trial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invalid valuation bias</td>
<td>Under- or overestimation of costs</td>
<td>Use reference prices, or present price</td>
</tr>
<tr>
<td>Ordinal ICER bias</td>
<td>Uninformative ICER</td>
<td>calculation in detailed manner</td>
</tr>
<tr>
<td>Double-counting bias</td>
<td>Under- or overestimation of costs and effects</td>
<td>Consider explicitly if variables are not</td>
</tr>
<tr>
<td>Inappropriate discounting bias</td>
<td>ICER varying from cost-effective to not being cost-effective</td>
<td>double counted; items included in the</td>
</tr>
<tr>
<td>Limited sensitivity analysis bias</td>
<td>Level of certainty surrounding the ICER is wrongly regarded as low</td>
<td>effect side should be included in the costs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and vice versa</td>
</tr>
<tr>
<td>III Bias after trial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sponsor bias</td>
<td>Favourable ICER of sponsored studies</td>
<td>Sponsorship disclosure, stringent</td>
</tr>
<tr>
<td>Reporting and dissemination bias</td>
<td>Tendency to underreport non-significant findings in journals versus monographs</td>
<td>guidelines, study protocol, rigorous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>review</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monographs publicly available and trial-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>based economic evaluation listed in trial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>registers</td>
</tr>
</tbody>
</table>
As mentioned earlier, until now, no study has given an explicit overview of biases in trial-based economic evaluations. However, we have identified a number of publications, either tutorials (such as Ramsey et al., 2005) or articles which highlight one or more methodological issues in trial-based economic evaluation (such as Bell et al.; M. Drummond & Sculpher, 2005; Matthews, Dumville, Hewitt, & Torgerson, 2011), which are included in this scoping review. In particular, the publication of Drummond et al. gives a good overview of the common methodological flaws in trial-based economic evaluations. Based on the search results, citation tracking, and extensive experience of the research group in conducting cost-effectiveness studies, an overview is made of the potential biases and how these biases can be detected and overcome (see also Figure 3.1). A total of 11 potential biases were identified (see Table 3.1) of which four were related to pre-trial, five to bias during trial and two to bias after the trial. Each of them is described separately below. For each bias, we first explained and defined the type of bias (under i), illustrated the influence of bias (ii), and described bias-reducing strategies (iii). An overview of these points (i, ii and iii) is given in Table 3.1.

**Major Sources of Bias in Trial-based economic evaluation**

- Narrow perspective bias
- Inefficient comparator bias
- Cost-measurement omission bias
- Intermittent data collection bias
- Invalid valuation bias
- Ordinal ICER bias
- Double-counting bias
- Inappropriate discounting bias
- Limited-sensitivity analysis bias
- Sponsor bias
- Reporting/dissemination bias

**Figure 3.1**
Chapter 3

I Pre-trial bias

Narrow perspective bias

In general, textbooks and guidelines state that economic evaluators have to be explicit about the perspective which is taken in the analysis.\textsuperscript{29} The perspective is the viewpoint from which the trial-based economic evaluation is performed and it will have particular influence on the range of costs included in the analysis. In general, a societal perspective is seen as optimal and superior,\textsuperscript{30} and is recommended by most (pharmacoeconomic) guidelines. In a societal perspective, all costs and outcomes are considered important, regardless of who pays the costs and will have the effects. The remaining minority of guidelines recommend performing economic evaluations from the perspective of the payer or of the healthcare sector;\textsuperscript{31} this can be more suitable in certain limited cases.

The narrow perspective bias can be defined (i) as using a restricted perspective, which will lead to omissions of important costs and outcomes, as these costs and outcomes are regarded as irrelevant from a narrower perspective. This will lead to the inefficient allocation of resources in the short term as well as over the long run (ii).

The bias-reducing strategy in this case (iii) is quite straightforward. Economic evaluators should be explicit about the perspective, which is used for the trial-based economic evaluation, and the societal perspective is preferred.\textsuperscript{30,31} In addition, within a broader societal perspective, it is possible to differentiate the perspective into narrower perspectives (such as the payer and the healthcare sector perspectives) in order to address specific policy issues.\textsuperscript{30} Other narrower perspectives will include only certain components. If economic evaluators use a more narrow perspective, the evaluators should provide justification as to why this narrower perspective is valid. A recent systematic review looking at economic evaluations of behaviour change interventions revealed that this is not common practice;\textsuperscript{32} this review showed that the perspective of the analysis was explicitly stated by only 57% of the studies. Moreover, trial-based economic evaluations often state that a broader (societal) perspective is taken, while a review of the study reveals that in reality sometimes a more limited perspective has been used.

Although using the societal perspective is advised, we would like to comment that it is not completely without drawbacks. An issue is that, when
considering healthcare planning, there is a strong debate on whether the societal perspective is the most informative in comparison with, for instance, the healthcare perspective. Often costs outside the healthcare sector, such as labour, educational and judicial costs, are not of interest to health care planners. However, a societal perspective, which includes a disaggregated description of types of costs, could improve the use of economic evaluations in healthcare planning. However, use of the societal perspective in economic evaluation might lead to a greater likelihood of other biases, such as double counting (see double counting bias) and increased uncertainty (see limited sensitivity analysis bias).

**Inefficient comparator bias**

The selection of appropriate comparators for the intervention studied is critical for any trial-based economic evaluation.\(^{33}\) If, in the case of inefficient comparator bias (i), a less effective treatment option is used as a comparator, the possible gains for both costs and outcomes might be overestimated, leading to biased cost-effectiveness estimates. Therefore, the choice of comparator is critical when evaluating a behavioural or educational intervention for a trial-based economic evaluation.

At first glance the choice of a comparator in a trial-based economic evaluation seems to be no different from the choice of comparator for any other analytic design. However, guidelines for economic evaluations often state that care-as-usual should be used as the most relevant comparator (see www.ispor.org/peguidelines/index.asp). The motivation for using care-as-usual as a comparator is that economic evaluations serve to inform policymakers about the consequences of implementing a (behavioural or educational) intervention into their healthcare system. For policymakers facing this decision, care as usual is the best comparator possible for revealing the opportunity costs and outcomes when implementing a new treatment in comparison with prevailing practice. The problem is that care as usual is often not defined\(^ {34,35}\) and this comparator can differ from country to country. Furthermore, care as usual is a very diverse and often collective term for all the different interventions with which control subjects are confronted during the trial. Therefore, the care as usual comparator is unlikely to be the ‘best’ alternative and thus not the most cost-effective alternative intervention currently available.
As care-as-usual is somewhat vague it is likely to provoke comparator issues (ii) in a trial-based economic evaluation. This is illustrated by the review of Hill et al., looking at submission for reimbursement of pharmaceuticals in Australia, which revealed that in 6% of the economic evaluation studies, uncertainty about the choice of the comparator or an inappropriate comparator is an issue.

To avoid inefficient comparator bias (iii), a three-armed trial in which the behavioural or educational intervention is compared with care as usual and the best alternative might be an option. In any case, the behavioural or educational intervention, the comparator, (s), including usual care should be described in detail; i.e. the authors should tell who did what to whom, where, and how often. This will enable the reader to relate the information on costs and outcomes of behavioural or educational intervention to the comparators.

\textit{Cost measurement omission bias}

Within a trial-based economic evaluation looking at behavioural and educational interventions, it is crucial to identify and measure all relevant resource use, looking at the disease and at the intervention studied. Even if the societal perspective is chosen, this does not necessarily mean that all important costs and outcomes are included. The adaptation of the societal perspective raises the biggest measurement challenges, and as a result it may be tempting to omit items for consideration.

As mentioned by others, one way of biasing a trial-based economic evaluation would be the cost measurement omission bias (i), i.e. leaving out measurement items that might count against the behavioural or educational intervention. This might lead to overestimation of the cost differences between the intervention and the comparator (ii).

In choosing items for a resource-use measure (RUM) - often a diary or questionnaire - it is important to focus on the dominant cost drivers, and on those cost drivers that will be different between the intervention and the comparator. Accordingly, items included in the RUM need to be sensitive to the needs of the group studied and sensitive to the intervention under consideration. The RUM should highlight benefits and savings which are ultimately valued in monetary terms. This implies that an RUM looking at, for instance, ‘free heroin under medical supervision to people heavily addicted to the drug’ might include other items (including criminal and judicial costs),
Risk of bias in trial-based economic evaluations

than an RUM looking at ‘self-management for people with stroke’ which might include items such as the costs of informal care. Within that process it is advised, for standardisation, to base the items of the RUM on earlier resource use measurement instruments rather than developing specific items over and over. The Database of Instruments for Resource Use Measurement (DIRUM) (www.dirum.org) can be very helpful in establishing shared resource use measures and methods. Using an available RUM has several advantages; first of all, it is less costly and it will lead to consistency across studies. But even the most appropriate existing questionnaire may not represent the best way to capture the key cost drivers for any given study. As a result, each existing instrument has to be adapted or extended with new items. The process of developing valid and reliable items for an RUM encompasses a lot of steps in planning, development, and piloting, and might well take a lot of time. Due to time constraints, existing RUMs are often used without any adaptation when looking at behavioural and educational interventions in trial-based economic evaluation; this might in the end bias the results of the study. Accordingly, a strategy for overcoming cost measurement omission bias (iii) is to identify explicitly those items which are relevant for the disease and intervention studied. In order to get a grip on the big items when looking at resource use, information can be gathered from cost-of-illness studies, qualitative interviews with stakeholders, earlier economic evaluations in the same field, or from existing RUMs. In addition, several authors have made taxonomies of possible costs and benefits inside and outside the healthcare sector; these might be informative when identifying relevant items for the resource use measurement.

**Intermittent data collection bias**

Once a valid RUM is constructed, it is important that there is insight into all resource use (costs) during the entire follow-up period of the trial-based economic evaluation. From a societal perspective, all relevant resource use, formal and informal, reimbursed and not-reimbursed, has to be included in an RUM. Accordingly, the RUM has to include healthcare sector costs, patient and family costs, and costs in other sectors. Because there are no institutional records tracking information on the type, frequency, and magnitude of resource use, data must be obtained through self-reporting RUMs, such as questionnaires and diaries. In comparison with questionnaires, diaries have
been reported to provide information prospectively over a period of time, resulting in minimum recall error.\textsuperscript{41}

Whatever methods are chosen, the resource use has to be known continuously, meaning that insight has to be gained into the utilization of resources every day during the study period,\textsuperscript{42} as not doing this will lead to an intermittent data collection bias (i). Earlier studies\textsuperscript{41,43} revealed that the intermittent collection of data – for at least 3 months during a 1 year period - might provide good estimates of annual costs. However, overall little is known about the patterns in which resource use occurs during the follow-up period of a trial, and as a result limited evidence is available about exactly when - which three months of the year - the resource use should be collected. For instance, intermittent data collection could be performed during the first, middle, and last months of a year, during periods with supposedly higher costs, or randomly.\textsuperscript{42} When data are collected intermittently, total annual costs have to be estimated based on the obtained data. Several imputation techniques are available for this. A study of Hendriks et al.\textsuperscript{42} compared several intermittent data collection patterns combined with several imputation techniques with data collected continuously from a trial-based economic evaluation. The study revealed that the amount of bias for random source measurement combined with individual mean imputation - in which the sum of the observations available is multiplied to obtain annual costs - leads to the smallest bias (0.4%), whereas other data collection patterns and imputation techniques could lead to a bias of up to 20% (ii). Overall the choice for intermittent data collection instead of continuous data collection is often based on practical considerations, such as relieving the burden on participants of the trial. As shown by the study of Hendriks,\textsuperscript{42} this choice might lead to an unintentional bias. Intentional bias might be induced if in a trial-based economic evaluation, data collection patterns are chosen in such a way that they are in favour of one of the comparators. Although there is some evidence that intermittent data collection patterns with imputation might be valid,\textsuperscript{42} this evidence is based on only one trial on fall prevention for elderly people. As we have concerns regarding the generalisability of these results to other settings, we would like to advise (iii) measuring resource use data continuously in a trial-based economic evaluation in order to reduce the data collection bias.
II Bias during trial

Invalid valuation bias

In the valuation step, all the resource use (such as contacts with GPs, consultations with psychologists, and days in hospital), which is measured by the RUM, are translated into monetary values (Euros e.g.). For the valuation in economic evaluation studies, the resources used should be valued at their opportunity costs, i.e. the value of their best alternative use. As the healthcare market is not fully competitive, it is in general very difficult to use opportunity costs. To resemble these costs, cost prices are calculated optimally using a bottom-up approach, which is a very detailed method of costs calculation covering the full costs of the resources used. Overall tariffs or charges are not seen as a valid approximation of these costs. Tariffs can be used only if there is a clear indication that a tariff represents a reasonable approximation of the actual costs. As calculating the costs bottom-up is labour intensive alternative sources for cost prices, such as unit costs which are published annually, costs of individual facilities, or combined estimates of different reference setting, are used as well as tariffs. In the case of an invalid valuation bias (i), unit prices are not estimated correctly, leading to either an under- or overestimation of the costs (ii). The amount of a cost price of a certain resource use, for instance a consultation with a clinical psychologist, depends on a number of factors. Jacobs et al. addressed issues of bias in cost valuation, and illustrated that the height of the cost price depends on the scale on which the clinical psychologist is working (scale bias), the case-mix (case-mix bias), the sites selected for the trial (site selection bias) and the methodology used (methods bias). Accordingly, calculating bottom-up cost prices can be biased and is very time consuming. Comparability between studies can be at stake due to differences in the methodologies used by the economic evaluator. A special issue of cost valuation, which might be inducing biases, is the fact that profit making organisations, like the pharmaceutical industry, have to set a price for their drugs. Knowing that if the ICER, the costs per QALY, is below a certain threshold, the drug is regarded as cost-effective, might lead to price inflating mechanisms. In this situation the final price of a drug is set at such a level that this price is the maximum price acceptable for the decision maker (so that the final ICER is below the threshold), and at a minimum price
acceptable to the company (the highest price possible for an ICER being below the threshold).

To overcome the invalid valuation bias (iii), for every aspect of trial-based economic evaluation, country-specific data (cost prices and utility values) have to be used as much as possible, and cost prices should relate to one recent reference price year. Accordingly, to overcome systematic errors in the price valuation, a number of countries, including among others Australia, Canada and the Netherlands, have established guidelines with recommended unit costs. A reference price is an average unit cost estimated on the basis of large, diverse populations; the average unit cost estimate can be used directly to value resource utilization. Reference prices reduce the likelihood that biases will occur and increase the comparability of economic evaluations within one country. Furthermore, if the cost price of a particular resource use, such as the intervention studied, is based on own cost valuation, this calculation should be presented in a detailed and transparent manner.

**Ordinal ICER bias**

A crucial part of an economic evaluation is the incremental analysis of costs and outcomes of the behavioural/educational intervention versus the comparator, such as in an ICER. An issue that is highlighted in almost every systematic review of trial-based economic evaluations, also in field of public health, is that, without any motivation, this incremental analysis is often missing.

As mentioned earlier, the ICER reflects how much it will cost to improve one point on a certain outcome scale. For resource allocation decisions, cardinal (interval or ratio) scales are necessary. In cardinal scales, the difference between two points is quantifiable and interpretable. In addition, it is assumed that the difference between two points has the same meaning throughout the whole range of the measurement scale.

It is known that nominal scales are unsuitable for use in an ICER, and will thus induce a ordinal ICER bias (i). If the ICER is calculated based on nominal scales, in many cases the ICER may be misleading for resource allocation, as these scales do not allow for applying fundamental algebraic operations (ii). As a result, the ICER calculated using ordinal scales is not informative. The easiest way to circumvent an ordinal ICER bias (iii) is to avoid using ordinal scales when calculating an ICER. This means that in designing the trial
the economic evaluator should already be aware of this bias when choosing primary outcome measures for the cost-effectiveness study. In addition, economic evaluators should be aware that using previously mentioned thresholds to decide if a health behaviour change intervention is cost-effective or not is meaningful only if the outcomes used are QALYs, as to our knowledge the current threshold defines only how much a society is willing to pay for a QALY, but does not necessarily apply to other parameters.

Double-counting bias

Double-counting bias (i), is an error that can occur in economic evaluation, whereby a parameter is counted more than once. It occurs in cost-utility analyses, when consequences of a treatment get incorporated on the cost side (numerator) as well as in the estimation of quality weights, i.e. QALY (denominator). In addition, double counting may also occur when social security benefits and other transfer costs are valued monetarily, in additional to valuing, for instance, productivity losses monetarily.

As already noted above, a societal perspective is recommended. Accordingly, all relevant costs and outcomes should be included in the ICER. In other words, benefits which are excluded in the QALY estimation need to be included in the costs. As it is not always clear what has to be included in the QALY estimation, two problems can arise (ii). First, a cost or benefit item can get lost and not be included in the analysis and second, cost and benefit may be included in the QALYs as well as in the costs. Double-counting in general is an important issue, which is discussed in detail elsewhere. The extent to which double-counting becomes relevant depends on a number of factors, such as the healthcare system, reimbursement procedures and sick-leave payment. Looking at changes in productivity and healthcare costs, it's obvious that patients might incorporate the QALY assessment in their responses. Being aware of this issue (iii) might lead to a clear strategy to overcome or at least minimize this bias. Furthermore, a good phrasing of the assessment questions is crucial, for instance by stating explicitly what assumptions the patient should make when completing a questionnaire or diary.
Inappropriate discounting bias

In behavioural and educational trials the follow-up period of respondents is often longer than one year. In general, regarding time preference, people prefer to have positive outcomes right now (for instance health gains) and negative outcomes, like costs, later in time. This is especially the case for preventive activities, such as lifestyle programs which aim to avoid diseases later in life. Due to the fact that people want to pay their costs as late as possible, and to enjoy their benefits as soon as possible, discounting is necessary in trial-based economic evaluation studies. Discounting is rather technical;\textsuperscript{51} several rates (varying between 1.5\% for effects to 6\% for costs) and models are suggested.

Several authors have shown that the ICER, i.e. the final results of a trial-based economic evaluation, is very sensitive to the discount model and the discount rate used, thus inducing an inappropriate discounting bias (i).\textsuperscript{52,53} Consequently, discounting can have a strong influence on cost-effectiveness ratios (ii) for setting priority, especially in preventive interventions. As a result, the conclusions of a trial-based economic evaluation study could change from a behavioural/educational intervention being cost-effective to not being cost-effective, just by changing the discount rates.

As there is no unique correct way for dealing with discounting, variation in practice is defensible.\textsuperscript{44} However, as there is no agreement over the models and rates used for discounting, this might introduce biases in which economic evaluators choose the most favourable discount rate. To overcome the use of a possibly inappropriate discount rate (iii), we suggest presenting several scenarios of the results, using various discount models and rates as a kind of sensitivity analysis. This sensitivity analysis should include undiscounted results (0\%); this is also suggested by Hillman and Kim\textsuperscript{54} as a scenario, and if available, the discount models or rates that are suggested by the country-specific guidelines should be used.

Limited sensitivity analysis bias

The limited sensitivity analysis refers to the fact that the trial-based economic evaluation does not apportion uncertainty according to different sources (i). Uncertainty in trial-based economic evaluation might come from several sources; the methodology, sampling variation, extrapolation,
Risk of bias in trial-based economic evaluations

generalizability.\textsuperscript{23} Moreover, evaluators in trial-based economic evaluation can be uncertain about the calculation of one specific parameter. Methodological uncertainty derives from the situation that different techniques can be used to value costs and utilities. As there is no gold standard, several countries advise performing their trial-based economic evaluation alongside a reference case. A reference case is a standard set of methodological practices that an economic evaluator would seek to follow in an economic evaluation study. Often, even in a randomised trial, the evaluator might have doubts about the group selected or there might be outliers in one of the arms of the trial. This sampling variation will result in uncertainty around the ICER. In order to test for sampling uncertainty, statistical techniques such as bootstrapping can be used. Another source of uncertainty arises in the situation when a researcher would like to extrapolate the results of a trial-based economic evaluation beyond the time of the trial. Modelling techniques can be used for this. Finally, sensitivity analysis can be used to test whether the results of a trial-based economic evaluation still hold if the results are generalised to another setting or if one parameter is changed. In sensitivity analysis, the robustness of a trial-based economic evaluation is assessed by examining the changes in results of the analysis when key variables are varied over a specific range.

The bias (ii) induced from having no or a limited sensitivity analysis is that the level of certainty surrounding the ICER is wrongly regarded as low. Economic evaluators (iii) should make a list of the variables for which uncertainty is assumed. Evaluators should use a reference case and test the uncertainty regarding this variable by means of alternative assumptions. This will include analysing the effects of sampling uncertainty, the impact of methodological assumptions (such as discount rate, study perspective), and subgroup analysis.

III Bias after trial

Sponsor bias

The most discussed bias in research is probably the sponsor bias.\textsuperscript{55} It is often mentioned that profit-making organisations (first of all the pharmaceutical industry) have been considering economic evaluations to be a marketing device instead of academic exercises (M. Drummond, 1991\textsuperscript{56}; Hillman et al.,
and thus are more likely to induce bias (i). Accordingly, a number of studies (ii) (Azimi & Welch, 199858; Baker, Johnsrud, Crismon, Rosenheck, & Woods, 200359; Bell et al., 200627; Bero & Rennie, 199660; Friedberg, Saffran, Stinson, Nelson, & Bennett, 199961; Hill et al., 200036; Hillman et al., 199157; John-Baptiste & Bell, 201062; Lexchin, Bero, Djulbegovic, & Clark, 200363; Miners, Garau, Fidan, & Fischer, 200564 have been focusing on the fact that profit-making organisations are more likely to report favourable results from trial-based economic evaluation studies. The studies’ overall conclusion is that ICERs submitted by profit-making organisations are significantly favourable in comparison with submissions performed by non-profit-making organisations. Studies sponsored by the industry were more than twice as likely to report favourable ICERs.27 One of the explanations for this difference might be the lower methodological quality of studies submitted by the industry.27,36,63 An alternative explanation might be a publication bias,61,63 and the phase (timing) in which these economic evaluation studies are performed. Regarding publication, there seems to be a bias in profit-making organisations towards “positive” results.61 Furthermore, it can be argued that profit-making organisations are more likely to be involved in research in the early stage of the dissemination of an intervention, as the study population in this phase is more selective towards people for whom the intervention is optimally intended. Accordingly, early trial-based economic evaluations are more likely to yield positive outcomes.61 Notwithstanding the presence of a sponsorship bias, the reality is that the industry is still one of the major funding bodies for trial-based economic evaluation,65 especially for pharmaceuticals. As a result, solutions (iii) have to be sought to overcome this bias. Of course, the sources of sponsorship should be fully disclosed at all times. The idea behind this is that disclosure will place sufficient pressure on economic evaluators to be sure that their research and subsequent publication is as unbiased as possible.66 However, disclosure of the financial conflict of interest cannot eliminate all biases. In the long run, maintenance of good methodological standards is of major importance. This implies that there are stringent guidelines on how to perform trial-based economic evaluation, and that the study protocol (including the analysis) should be published beforehand. Finally, rigorous independent assessment and review of the studies might also be ways of limiting this bias.57,67
Reporting/dissemination bias

Within the field of economic evaluation there seems to be a specific form for reporting bias (i), in addition to the publication bias, which is known. Economic evaluations are often published both in a monograph produced for the funders, and in a corresponding professional journal. A review study by Matthews et al.\textsuperscript{28} revealed (ii) that there is a tendency to underreport non-significant outcomes in journal publications as opposed to monographs, with the consequence that significant statistical findings tend to be disseminated more widely.

To avoid the possibility of a reporting/dissemination bias (iii) the monographs should be included in decision-making processes.\textsuperscript{28} Furthermore, the monographs should be available publicly and (trial-based) economic evaluations should be listed in trial registers.

Conclusion

This article attempts to identify biases, which may occur before, during, and after a trial-based economic evaluation. It is important for various stakeholders to be aware of these biases as trial-based economic evaluations are being used increasingly in reimbursement decisions. Understanding biases also allows readers to review trial-based economic evaluations critically and independently. To gain insight into the potential biases of a study, a transparent and detailed description of trial-based economic evaluation is necessary.\textsuperscript{15} Journals can encourage greater transparency by requiring authors to provide appendices with details to be published on the web in order to overcome space limitations.\textsuperscript{62}

As we have shown, there are several ways of influencing the final results of trial-based economic evaluation with respect to biases. To our current knowledge, this is the first scoping review giving an overview of possible biases in trial-based economic evaluations. A review of the literature revealed that there are several types of systematic errors, which can influence the final results of trial-based economic evaluation. Overall, the number of studies dealing with biases in trial-based economic evaluation is limited, and the evidence included in this scoping review thus comes from only a few articles.
Accordingly, additional research is needed to identify (additional) biases, and to reveal appropriate strategies for addressing the risk of bias. It is important to understand that different bias types can be related to each other and that the elimination of one can introduce another. For instance, using the societal perspective (in order to reduce the limited perspective bias) might induce double counting (double counting bias) and uncertainty (limited sensitivity analysis bias). Accordingly, different stakeholders have to be aware of all types of biases to make an informed decision.

During a trial-based economic evaluation, evaluators have to decide among several choices. Several checklists\textsuperscript{29,68} have been designed to scrutinize the quality of a study. There is a thin line between limited methodological quality and biases. In this article, we have classified factors as a bias if it is likely that any manipulation will lead to a difference between the true value in the population and the observed value in the trial-based economic evaluation. As a consequence, we have limited ourselves to true biases, although evidence showed that studies with a high score on quality were also less likely to report favourable ICER.\textsuperscript{27}

This article is not a systematic review as such, but in writing this article we have tried to retrieve as much information as possible about biases in trial-based economic evaluations. This scoping review is based on a limited search in PubMed and the NHS EED only, which might have led us to miss some relevant publications, although it is generally accepted that using this search practice is an optimal search strategy for retrieving economic evaluations.\textsuperscript{24,25}

Furthermore, this article, due to the scope of this special issue of *Psychology & Health*, focuses only on trial-based economic evaluations. In general, all the biases discussed in Table 3.1 also hold for model-based economic evaluation studies.\textsuperscript{69} In a model-based economic evaluation there are a number of additional potential biases which are relevant;\textsuperscript{3} however, identifying these biases goes beyond the scope of this article.

Like Cain et al.\textsuperscript{70}, we believe that everyone is a little bit biased, including economic evaluators. As noted earlier, a number of guidelines have been developed for trial-based economic evaluations. However, these guidelines and the solutions suggested in Table 3.1 are not enough to overcome all biases. Economic evaluators, like any other researchers, will have to rationalize the choices they have made during every phase of the study.\textsuperscript{71} In order to avoid biases in trial-based economic evaluations, researchers, readers,
reviewers, editors, decision-makers etc. have to be aware of the possible biases and all stakeholders have to examine trial-based economic evaluations in a rigorous and stringent manner. This article can be helpful in this examination as it provides an overview of the possible biases, which researchers should take into account.
References

7. Bruijn De M XXX Editorial
8. Williams A. Cochrane Lecture. All cost effective treatments should be free ... or, how Archie Cochrane changed my life! J Epidemiol Community Health 1997;51:116-120.


35. Bishop FL, Fenge-Davies AL, Kirby S, Geraghty AW. Context effects and behaviour change techniques in randomized trials: A systematic review using the example of trials to increase adherence to physical activity in musculoskeletal pain. Psychol Health 2015;30,104–121.


Chapter 4

Risk of bias in model-based economic evaluations: The ECOBIAS checklist

Adarkwah CC
Van Gils PF
Hiligsmann M
Evers SMAA

Submitted to Expert Rev Pharmacoecon Outcomes Res
Abstract

**Background**
Economic evaluations are becoming increasingly important in helping policymakers to efficiently allocate scarce resources. Several biases can occur when performing economic evaluations. It is therefore important for policymakers to be able to assess potential biases and for researchers to minimize them. Earlier research revealed which biases exist in trial-based economic evaluations.

**Objectives**
This article aims first to identify biases that are specifically related to model-based economic evaluations and to illustrate their potential impact on economic outcomes using examples from the literature. Second, the article aims to present a checklist for assessing the risk of biases in economic evaluations (the ECOBIAS checklist), which can be used for trial- and model-based studies.

**Methods**
Several possible sources of bias in model-based economic evaluations were identified using the Philips guideline for good practices in modelling economic studies as a structuring framework. Biases were identified and illustrated using evidence from previously published model-based economic evaluation studies. Examples of biases were found through a scoping review, scrutinization of systematic reviews that used the Philips guideline, working-group meetings and discussions with the lowlands Health Economics Study Group (lolaHESG). By combining biases that can occur in trial-based with those that can occur in model-based studies, which were identified in a previous article by the author group, a checklist for assessing biases in economic evaluations was developed (ECOBIAS).

**Results**
Eleven additional model-specific biases were identified and classified by structure, data and consistency of the model. These biases are related to structural assumptions, model type, time horizon, data selection (such as treatment effects), assessment of uncertainty and internal validation. The impact of these biases could be massive, changing the outcomes from being highly cost-effective to not being cost-effective at all. The ECOBIAS checklist aims to help researchers identify biases and includes a total of 22 biases, of which eleven are specific for modelling economic studies.

**Conclusion**
In this study, we identified several biases that are related to model-based economic evaluations and developed the ECOBIAS checklist for identifying biases in economic evaluations. We hope that our results and the ECOBIAS checklist will help to reduce biases in future economic evaluations and will increase faith in model-based economic evaluations in particular.
Introduction

Economic evaluations are becoming increasingly important in providing policymakers with information for reimbursement decisions. However, in many cases, a significant difference is notable between study results and real-life observations. This can be due to confounding and many other factors. In this context, bias might play an important role.

There are basically two analytical frameworks used to perform economic evaluations: model-based and trial-based. In a model-based economic evaluation, data from a wide range of sources (randomized-controlled trials (RCTs), meta-analyses, observational studies) are combined using a mathematical model to represent the complexity of a healthcare process. Several studies and guidelines have already highlighted the importance of strong methodology by developing guidelines for model-based economic evaluations.1-7 Those guidelines provide information on the different steps for conducting a model-based economic evaluation. Overall, these guidelines instruct researchers on what to do, but not exactly on how to deal with specific biases. Some guidelines provide a checklist that includes several criteria that are assumed to reflect aspects of quality11-13 while others do not.11-13 The checklists provided are either based on a number of items that are given the same weight,8,10,14 or are based on a quantitative grading system (Chiou et al., 2003).8 Other guidelines focus on specific methodological aspects. Eddy et al. focused on transparency and validation.15 They evaluated methods and recommended best practices to make models more transparent and to validate them. Briggs et al. developed a guideline for examining and reporting uncertainty surrounding an outcome.16 In Caro et al. (2014), a questionnaire to assess relevance and credibility was developed.17 All of the abovementioned concepts should be critically considered by modelling researchers. However, none of these studies focused specifically on biases in economic evaluations. In the current article, biases are addressed. The Cochrane Handbook18 defines bias as “...a systematic error, or deviation from the truth, in results or inferences”.19 Bias can lead to over- or underestimation of the true intervention effect and is different from imprecision in the estimated treatment effect, which implies random rather than systematic error.20 A bias
can be unintentional or intentional and can have either substantial or insubstantial impact.
In other fields, such as epidemiology and health psychology, explicit attention has been paid to common sources of bias and strategies for reducing the risk of bias. This has resulted in widely-used instruments for assessing the risk of bias, such as the Cochrane risk of bias tool for randomized controlled trials.\textsuperscript{19}
Awareness of potential biases is also extremely important in the field of economic evaluation for both researchers and policymakers. Researchers should be aware of the forms of bias so they can reduce or at least minimize them, while policymakers should be aware of potential bias so they can accurately assess economic evaluations.
Several biases related to trial-based economic evaluation have already been identified and discussed previously.\textsuperscript{21} As model-based economic evaluation involves other methods, it also faces other challenges. Specific biases that could occur in modelling studies need to be identified when using and reviewing model-based studies.
This study aims to identify additional key biases that could occur in model-based economic evaluations and to illustrate their potential impact on cost-effectiveness results using examples from the literature. As part of this study, a checklist was also developed, called the Bias in Economic Evaluation (ECOBIAS) checklist, which provides a full overview of the biases that could occur in model-based and trial-based economic evaluations. The ECOBIAS checklist can be used to consistently examine any bias in an economic evaluation. Researchers can use it in planning, performing and analysing an economic evaluation.

Methods
The first step was to identify relevant biases in model-based economic evaluations. For this identification step, we combined several strategies including: (a) a scoping review to identify any biases, (b) an analysis of earlier systematic reviews and (c) input from experts to find additional biases. For the scoping review (a), a literature search was performed in March 2015 using PubMed (search terms: Costs and Cost Analysis [MeSH Terms] cost-effective
Bias in model-based economic evaluation

As we had the idea that systematic reviews (b) looking at model-based economic evaluation might reveal some additional biases, we also performed a literature review, searching for earlier systematic reviews focusing on model-based economic evaluations. To identify relevant model-based studies, a literature search was performed in April 2015, primarily focusing on reviews using the Philips guideline, which is a comprehensive assessment tool that is recommended by the Cochrane Handbook for Systematic Reviews of Interventions for assessing the methodological quality of model-based economic evaluations. We identified these systematic reviews using PubMed (search terms: philips[All Fields] AND (“checklist”[MeSH Terms] OR "checklist”[All Fields])). All articles were checked for valuable information on bias.

The findings were complemented with knowledge from the author group and input from external experts. Regular meetings of the author group were held, and two external experts were consulted for the identification and illustration of biases. The findings were also presented and intensively discussed during the lowlands Health Economics Study Group (lolaHESG) Annual Meeting (May 2015) to reveal additional biases.

A second step in the strategy was to classify all the biases that were identified in step 1 into a logical framework. In this step, we classified biases into three different sections (i.e. structure, data and consistency) and related subsections, following the Philips guideline.

Third, for every bias identified in published model-based economic evaluation studies, we first described the overall recommendations of the Philips guideline, defined the bias thereafter and then illustrated the influence of the bias using one or two examples.

As a final step, the ECOBIAS checklist for bias in economic evaluation was developed, which gives an overview of biases in trial-based and model-based economic evaluations. Here, the results of the current paper were merged with the results of our previous article. The ECOBIAS checklist encompasses a common part for model- and trial-based economic evaluations and a model-specific part. Part A of the checklist incorporates eleven biases that occur in both types of economic evaluation, which were identified and discussed in a
previous article by the author group,\textsuperscript{21} whereas Part B covers eleven model-specific aspects of biases presented in the following section.

Results

Model-specific biases

In total, 38 studies were identified that cited the Philips guideline, and among those were twelve reviews. Eleven model-specific biases (Table 4.1) were identified, which are presented in detail in the following sections. The results section is structured according to the items of the Philips guideline to allow for a good overview. The following types of bias were identified:

<table>
<thead>
<tr>
<th>Type of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>I  Bias related to structure</td>
</tr>
<tr>
<td>Structural assumptions bias</td>
</tr>
<tr>
<td>No treatment comparator bias</td>
</tr>
<tr>
<td>Wrong model bias</td>
</tr>
<tr>
<td>Limited time horizon bias</td>
</tr>
<tr>
<td>II Bias related to data</td>
</tr>
<tr>
<td>Bias related to data identification</td>
</tr>
<tr>
<td>Bias related to baseline data</td>
</tr>
<tr>
<td>Bias related to treatment effects</td>
</tr>
<tr>
<td>Bias related to quality-of-life weights (utilities)</td>
</tr>
<tr>
<td>Non-transparent data incorporation bias</td>
</tr>
<tr>
<td>Limited scope bias</td>
</tr>
<tr>
<td>III Bias related to consistency</td>
</tr>
<tr>
<td>Bias related to internal consistency</td>
</tr>
</tbody>
</table>

A. Bias related to structure

\textit{A 1. Structural assumptions bias}

\textbf{Overall recommendation of the guideline:} It is of great importance that the structure of a model is consistent with a coherent theory of the condition that is being investigated.\textsuperscript{6,24} For instance, treatment pathways should carefully reflect the underlying nature and process of the disease, which is emphasized
separately in the Philips guideline. In addition, all structural assumptions should be transparent and well justified, and the presentation should be tailored to policymakers’ needs. The cycle length chosen should also be dictated by the natural history of disease.

**Definition of the bias:** Structural assumption bias exists if the model structure is not consistent with a coherent theory of the condition that is being investigated. If the issues mentioned above are not critically considered, structural assumptions bias is likely to occur.

**Illustration of potential bias:** The structural assumptions bias is very well illustrated in the paper of Hoogendoorn et al. Here, seven cost-effectiveness models on COPD were compared using the same input parameters. The results of four hypothetical interventions were checked against each other and the models were in this way cross-validated by running the same treatment scenarios. There was a high variability of structural assumptions, such as the number of COPD severity states, the possibility of backward transition and the inclusion of exacerbation-related mortality. The ICER significantly differs according to the model used, ranging from €17,000 per QALY to €47,000 per QALY, depending on the model. Those differences could only be explained by structural uncertainty in the models.

A 2. No treatment comparator bias

**Overall recommendation of the guideline:** In line with the recommendation, there should be a clear definition of all options that are evaluated. The model should reflect all options relating to the decision problem and therefore should also reflect current practice.

**Definition of the bias:** Hence, in a sound economic evaluation, you are comparing the intervention under investigation with care as usual. In a model-based economic evaluation study, you can, as opposed to a trial-based economic evaluation, induce the comparison of nobody receiving any alternative intervention, which is unlikely in practice. This will lead to the no treatment bias (i.e. an overestimation of the cost-effectiveness of the intervention studied) by comparing it with an unrealistic alternative (i.e. that the comparator receives no treatment).

**Illustration of potential bias:** In a systematic review of cost-effectiveness analyses of drugs in osteoporosis published between 2008 and 2013, it was shown that relatively few economic evaluations included all potential relevant
interventions in their analysis.\textsuperscript{27} Although there has been an increased use of active comparators, some articles are still only comparing a drug treatment versus no treatment, which could lead to a more favourable cost-effective result.

A 3. Wrong model bias

Overall recommendation of the guideline: The decision problem should dictate the appropriate model. In addition, the choices made with regard to causal relationships within the model also need to be considered carefully.\textsuperscript{24}

Definition of the bias: For instance, different model types have different justifications and indications. For certain decision problems (i.e. more simple scenarios that occur over a short time period), decision trees can be more appropriate than Markov models, while the latter can be more suitable for other problems, especially when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. Neglecting this might lead to bias.

Illustration of potential bias: Although it is difficult to find a completely wrong model in the literature, the paper of Simpson et al. clearly shows that the choice of the right model makes a difference.\textsuperscript{28} The article presents a comparison of Markov model and discrete-event simulation (DES) techniques in HIV modelling. Making use of published data as model input, both approaches were compared regarding the clinical outcomes at five-year and lifetime cost-effectiveness estimates. It was shown that the DES model predicted more detailed outcomes and had better long-term predictive validity than the Markov model. So the authors concluded that the DES model is better than the Markov model in isolating long-term implications of small but important differences in crucial input data. In another article, Karnon and Brown\textsuperscript{29} reviewed decision trees, Markov models and DES, taking breast cancer as an illustrative example. They discussed pros and cons and concluded that the choice of the right model depends on a number of factors, which need to be critically considered before choosing the model type for an economic evaluation.
A 4. Limited time horizon bias

Overall recommendation of the guideline: According to Philips et al., the time horizon in the optimal case should cover a lifetime to make sure not to miss important events. In all other cases, justification for shorter time horizons should be stated appropriately. It is furthermore important to distinguish between the time horizon of the model, the duration of treatment and the duration of treatment effects.

Overall, the time horizon should be long enough to capture all relevant costs and outcomes relating to the disease and the duration of treatment and treatment effects.

Definition of the bias: Bias might occur where the choice of the time horizon is not appropriate (too short) for the decision problem or choosing a more limited time horizon (less than a lifetime horizon) is not justified appropriately. The bias which is induced is that not capturing relevant events (in terms of costs and effects) beyond the study’s time horizon may have a significant impact on the final results of the study.

Illustration of potential bias: The paper of Vijgen et al. illustrated how time horizons affect the ICER and can potentially lead to bias. This paper determined the cost-effectiveness of a Dutch school-based smoking education program. The ICER of the school-based smoking education program compared with current practice was highly dependent on the time horizon (25 years: €231,500 per QALY gained, 50 years: €29,300 per QALY gained and 75 years: €18,500 per QALY gained). So if the time horizon is not long enough, the ICERs are rather unfavorable. This is because the major health benefits are to be found beyond a time horizon of 25 years.

B. Bias related to the data

B 1. Data identification

Overall recommendation of the guideline: Researchers should make the methods used to identify data transparent, and those methods should be in line with the model’s objectives. There needs to be a justification for any choices that are made about specific data inputs, and data identification itself should be done in a systematic way. In addition, particular attention has to be
paid to data identification of those parameters to which the results of the model are extremely sensitive.

Data identification refers to the sources where data that are implemented into a model come from. Literature suggests that data should come from high-quality and well-designed studies.\textsuperscript{31,32} Rules for assessing the quality of input data have been suggested.\textsuperscript{32}

**Definition of the bias:** The identification might induce bias in several ways. Model-based economic evaluation normally combines data from different sources and integrates them into one model. The advantage is that data from different sources can be obtained and used. For example, models often use a combination of local observational data (i.e. registries) to reflect local treatment patterns and experimental data (from RCTs).

Bias can be induced if the data used for the model are not an adequate choice. This issue needs to be considered critically. So the question is whether the input data come from large meta-analysis or RCTs or only small trials, which might be of lower quality. Another bias can be induced if researchers only rely on a sole source of information. This could be data from only one RCT where important comparators or events might not even be considered. We are aware of the fact that meta-analysis or RCTs are not available in any case, but the justification of data sources needs to be transparent and stated clearly, especially when relying only on data from a single trial.

**Illustration of potential bias:** In a recent systematic review of economic evaluations, Hiligsmann et al.\textsuperscript{27} use the CHEERS checklist to assess the quality of reporting of 39 economic evaluations. The authors showed that the methods used for the identification and synthesis of clinical effectiveness data is often not reported. Although poor reporting does not necessarily lead to poor quality and results bias, high-quality and transparent data identification is needed.

Next to that, adherence to interventions has been shown to have a substantial impact on the ICER.\textsuperscript{33} When assumptions about the willingness to participate in a screening program are too optimistic, for example, it may lead to conclusions that are too optimistic, while with assumptions that are more realistic, the conclusions about the cost-effectiveness would be less favorable. Taking the case of colorectal cancer screening as an example, an economic evaluation showed that a decrease in adherence from 100% to the more
realistic figure of 30% resulted in a doubling of the ICER, from €21,000 to €42,000.\textsuperscript{34}

Especially when the implementation of large public health interventions is considered, it is necessary to have accurate information about adherence. The relevance of this can be illustrated using the case of chlamydia screening in the Netherlands.\textsuperscript{35} The results from a relatively small trial to determine Chlamydia trachomatis prevalence organized by the Municipal Public Health Services showed a participation rate of 41%.\textsuperscript{35} However, a large pilot implementation project showed an adherence of only 16% after the first invitation, which declined to 9.5% after the third invitation.\textsuperscript{36}

Following this limited participation of the target group in the intervention, cost-effectiveness could no longer be demonstrated. This example underlines the importance of using real-life data for important parameters, especially for the effectiveness of cost-effectiveness models. Other examples that illustrate the economic impact of non-adherence are available in the literature.\textsuperscript{37}

B 2. Bias related to pre-model data analysis

a. Bias related to baseline data

**Overall recommendation of the guideline:** The guideline suggests that the methodology of data analysis should be described adequately and that the methodology should be based on justifiable statistical and epidemiological methods.\textsuperscript{24} This should be done with regard to baseline data, treatment effects and quality-of-life weights. Regarding baseline data, probabilities, for example, should be based on natural history data derived from epidemiological or observational studies. Transformation of rates into transition probabilities needs to be done accurately. Furthermore, not applying a half-cycle correction on all transitions in a state transition model (costs as well as outcomes) needs to be justified.

**Definition of the bias:** Looking at the definition, it becomes clear that inducing a bias with regard to baseline data is not unlikely. Although it is hard to find a concrete example here, you could imagine that there might be over- or underestimation of certain risks, costs or benefits with regard to a certain intervention, which could lead to overestimating (or underestimating) the cost-effectiveness of an intervention.
b. Bias related to treatment effects

**Overall recommendation of the guideline:** The guideline suggests that relative treatment effects derived from trial data should be synthesised using appropriate meta-analytic techniques. The methods and assumptions that are used to extrapolate short-term results to final outcomes should be documented and justified, along with alternative assumptions regarding extrapolation, which should be explored through sensitivity analysis. Assumptions regarding the continuing effect of treatment once treatment is complete should be documented and justified. If evidence regarding the long-term effect of treatment is lacking, alternative assumptions should be explored through sensitivity analysis.

**Definition of the bias:** If assumptions about treatment effects are not realistic and not justified adequately, bias is likely to occur. Other scenarios besides the base case need to be shown in order to demonstrate the impact on the ICER.

**Illustration of potential bias:** Treatment effects can be both, over- and underestimated. In van Baal et al., the cost-effectiveness of a low-calorie diet intervention and a combination of both a low-calorie diet and Orlistat treatment over a one-year period were compared to no treatment at all. The sensitivity analysis performed (Table 4.4 in the paper) shows that the assumptions about weight loss maintenance in the long run dramatically influence the ICER. In detail, the base case weight loss maintenance was 23%. Assuming 100% maintenance after one year, ICERs decrease significantly, from €58,800 to €24,100 in the combination group.

Furthermore, the assumptions about treatment effects after stopping treatment are a potential source of bias. The article of Caulin et al. shows the impact of the effects of stopping treatment on the ICER, taking osteoporosis as an example. The ICER was shown to substantially differ when changing the assumption about the effect of therapy after stopping it.

c. Bias related to quality-of-life weights (utilities)

**Overall recommendation of the guideline:** The utilities incorporated into a model should be appropriate for the specified decision problem. Utilities should be assessed using adequate methods and should be derived from the selected population in a formal way.
Definition of the bias: Hence, it is important to have a correct and transparent description of the derivation and incorporation of utility values. QALYs are often the main and preferred outcome in decision models. Utility values need to be consistently used within a model and derived in the target population using established methods. Even if the utility weights were derived by the same method (e.g. time trade-off or standard gamble, multi-attribute utilities scores), consistency needs to be assured in order to avoid bias.

Illustration of potential bias: Unruh et al. assessed the health-related quality of life (HRQOL) in hemodialysis patients, comparing self-administered and interviewer-administered surveys. The Kidney Disease Quality of Life–Long Form questionnaire (KDQOL-LF) was used as a measurement tool. The authors demonstrated that patients in the interviewer-administered group had higher scores on scales that measured role-physical, role-emotional and effects of kidney disease. They concluded that dialysis studies that restrict HRQOL measurement to patients who are able to complete surveys without assistance (i.e. the self-administered group) will not accurately represent the health of the overall hemodialysis population.

B 3. Bias related to data incorporation

Overall recommendation of the guideline: All data that are incorporated into the model should be described, and the sources of data should be reported in detail. All in all, the process of data incorporation should be transparent, and authors need to state clearly if data are considered as point estimates or as distributions.

Definition of the bias: Not having a clear and transparent way of data incorporation might induce bias. For example, the methods used in a particular study should be stated in such a way that the study could be replicated. As previously mentioned, non-transparent data incorporation does not necessarily lead to poor quality and result in bias, but at least hinders the assessment of bias.

Illustration of potential bias: Hiligsmann et al. reported that quality of reporting of recent economic evaluations is still largely insufficient for several articles. Several items were partially or not at all reported in most articles. Poorly reported items included methods used for the identification and synthesis of clinical effectiveness data, the description of the population and
methods used to value preference-based outcomes, the reporting of the
dates of the estimated resource quantities and unit costs.

B 4. Bias related to the assessment of uncertainty

Limited scope bias

**Overall recommendation of the guideline:** Regarding the assessment of uncertainty, the four principles of uncertainty, i.e. methodological, structural, heterogeneity and parameter, need to be considered.

**Definition of the bias:** Event though the standards of certain model-based economic evaluations can be quite high, studies can still lack in some parts of uncertainty. The Philips guideline distinguishes between the different types of uncertainty mentioned above, and there is a large variation in the models published. This starts with models that do no analysis at all and also includes models that do not look at all four types of uncertainty in particular (e.g. missing one type). We feel that, in many cases, not all possible and reasonable scenarios could be presented. Probabilistic sensitivity analyses are not always reported. Furthermore, the presentation of a best-case and worst-case scenario is often missing.

C. Bias related to consistency

C 1. Bias related to internal consistency

**Overall recommendation of the guideline:** According to the guideline, there should be evidence that internal consistency of the model, in terms of its mathematical logic, has been evaluated.

**Definition of the bias:** Internal consistency seems to be underreported, which hinders an adequate assessment of bias.

**Illustration of potential bias:** In a systematic review of model-based economic evaluations of diagnostic and therapeutic strategies for lower extremity artery disease using the Philips guideline, Vaidya et al. pointed out that internal consistency was not addressed in any of the sixteen identified studies. Next to that, in another study that used the Philips guideline, none of the four studies investigated addressed internal consistency.
The ECOBIAS checklist

The ECOBIAS checklist, which aims to help researchers identify bias, includes a total of 22 biases (Part A + Part B), of which eleven are specific for modelling economic studies (Part B).

Two of the biases identified here ("no treatment comparator bias" and "limited scope bias") have some overlap with biases mentioned in Part A; this overlap is indicated in Table 4.2. Authors can note if a bias is relevant to the actual study and describe in detail how the bias was dealt with, explaining rationale and strategy.

Table 4.2 The ECOBIAS checklist for bias in economic evaluation.

<table>
<thead>
<tr>
<th>Type of bias</th>
<th>Issues addressed (questions to consider)</th>
<th>Relevant to study Yes/ No / Partly / Unclear / NA</th>
<th>How did you deal with this bias? (description of strategy and rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrow perspective bias</td>
<td>Was a societal perspective adopted? If not, has a different perspective been justified?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inefficient comparator bias*</td>
<td>Was the best alternative chosen as a comparator? Was current practice chosen as a comparator? Have all comparators been described in sufficient detail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost measurement omission bias</td>
<td>Were all costs relevant to the disease and intervention identified and considered?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermittent data collection bias</td>
<td>Was the resource use measured continuously?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invalid valuation bias</td>
<td>Is the price calculation presented in a detailed manner? Have reference prices been used?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinal ICER bias</td>
<td>Have cardinal scales for the outcomes measure in a CEA been used?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-counting bias</td>
<td>Have the variables been adequately checked for double-counting?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inappropriate discounting bias</td>
<td>Have discounting rates from guidelines been applied?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited sensitivity analysis bias</td>
<td>Have the four principles of uncertainty (methodological, structural, heterogeneity, parameter) been considered in sufficient detail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sponsor bias</td>
<td>Have sponsorships been disclosed? Is the study protocol freely accessible?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting and dissemination bias</td>
<td>Has the study / trial been listed in a trial register? Have all results been reported according to the study protocol?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2 (continued)

<table>
<thead>
<tr>
<th>Type of bias</th>
<th>Issues addressed (questions to consider)</th>
<th>Relevant to study</th>
<th>How did you deal with this bias? (description of strategy and rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PART B. Model-specific aspects of bias in economic evaluation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I Bias related to structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural assumptions bias</td>
<td>Is the model structure in line with coherent theory? Do treatment pathways reflect the nature of disease?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment comparator bias*</td>
<td>Is there an adequate comparator, i.e. care as usual?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrong model bias</td>
<td>Is the model chosen adequate regarding the decision problem?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited time horizon bias</td>
<td>Was a lifetime horizon chosen? Were shorter time horizons adequately justified?</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>II Bias related to data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to data identification</td>
<td>Are the methods of data identification transparent? Are all choices justified adequately? Do the input parameters come from high-quality and well-designed studies?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to baseline data</td>
<td>Are probabilities, for example based on natural history data? Is transformation of rates into transition probabilities done accurately?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to treatment effects</td>
<td>Are relative treatment effects synthesized using appropriate meta-analytic techniques? Are extrapolations documented and well justified? Are alternative assumptions explored regarding extrapolation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to quality-of-life weights (utilities)</td>
<td>Are the utilities incorporated appropriate for the specific decision problem?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-transparent data incorporation bias</td>
<td>Is the process of data incorporation transparent? Are all data and their sources described in details?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited scope bias§</td>
<td>Have the four principles of uncertainty (methodological, structural, heterogeneity, parameter uncertainty) been considered?</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>III Bias related to consistency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to internal consistency</td>
<td>Has internal consistency in terms of its mathematical logic been evaluated?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* These biases are overlapping regarding their content. § These biases are overlapping regarding their content.
The ECOBIAS checklist

The ECOBIAS checklist, which aims to help researchers identify bias includes a total of 22 biases (Part A + Part B), of which eleven are specific for modelling economic studies (Part B).

Two of the biases identified here (“no treatment comparator bias” and “limited scope bias”) have some overlap with biases mentioned in Part A, this overlap is indicated in Table 4.2. Authors can note if a bias is relevant to the actual study and describe in detail how the bias was dealt with, explaining rationale and strategy.

Discussion

This study identified several additional biases related to model-based economic evaluation and showed that the impact of these biases could be massive, changing the outcomes from being highly cost-effective to not being cost-effective at all. The majority of the biases discussed in this paper might, however, be unintentional. The aim of the paper was not to criticize researchers, but rather to raise awareness of biases in economic evaluation. As research in the field of economic evaluation is developing, an unintentional bias today might be an intentional bias ten years from now, also due to the fact that one is able to test and correct for this.15,44 Looking at the specific issues mentioned in the results, it becomes apparent that overcoming some biases is possible, at least to some extent.

In order to be able to identify biases, being transparent in reporting is extremely important.45 Hence, when looking at models, policymakers are often insecure because there is a high deficit in understanding of what researchers exactly do.46 Open access to input data, assumptions, data identification and model structure, for example, should be facilitated by researchers,47 as is already standard in some other fields of research, such as computational neuroscience. Open access to the model itself should be provided in an optimal way. Using the ECOBIAS checklist would facilitate this process.

Our findings suggest that a wide range of alternative research results should be presented in model-based economic evaluation. This means that more results should be presented in different scenarios and also that sensitivity
analyses should be extended to strengthen results that will form the basis for decisions regarding reimbursements.

The ECOBIAS checklist presented in Table 4.2 can be useful both for researchers and for those who use and monitor the results of model-based economic evaluations, such as reviewers and policymakers. The checklist can be used to check if the work is performed systematically looking at bias and to demonstrate that the risk of bias is dealt with in an adequate way. To allow for more transparency, the completed checklist should be made available to reviewers and audiences, for example as an online appendix with the supplementary material. Furthermore, policymakers are enabled to check for biases and make more valid assessments of economic evaluations.

The ECOBIAS checklist can be seen as a complementary tool next to the guidelines and checklists mentioned previously.\textsuperscript{1-17} The establishment of an ISPOR task force on bias or a consensus panel meeting could be the next steps forward.

Still, the current research on bias in economic evaluation has some limitations. We cannot present a complete overview of all possible biases in economic evaluations, also because unawareness of bias is an issue. The reader should first be aware that this study focused on specific biases in model-based economic evaluation, which are important and should be critically considered. In a previous study,\textsuperscript{21} an overview of biases in trial-based economic evaluation was provided. Many of those biases also apply to model-based economic evaluation and were not discussed again in this article but were included in the ECOBIAS checklist (Table 4.2). Second, we highlighted biases and elucidated the impact on the final results of the economic evaluation. Our aim was not to describe strategies for overcoming those biases. Further research is therefore needed to look explicitly at possible ways to manage biases in more detail. Third, we decided on the Philips guideline as a starting point. We are aware that there are many other approaches that could be used, such as focus groups or other guidelines. Fourth, our checklist is most likely not exhaustive. We are open to any additional feedback, which can be sent to the first author. Although ECOBIAS already covers a large number of items, further research beyond the scope of our article is needed to establish all bias-related items relevant to economic evaluation studies. In this further research, we might also learn from other fields that already have a
longstanding tradition of looking at bias, such as environmental research, health psychology and epidemiology. Notwithstanding these limitations, we think that this article, combined with earlier work, gives a good overview of the potential biases in economic evaluation, although further research is needed. We recommend that editors and reviewers consistently use a checklist, such as the ECOBIAS checklist presented in Table 4.2. To ensure more transparency, the completed checklist could be made compulsory with submission and could afterwards be made available to the community online with supplementary material.

Expert commentary

Although there is some overlap between biases in trial- and model-based economic evaluation, this research identified several biases specific for modelling studies. Being aware of these biases could help researchers to minimize them, and the ECOBIAS checklist could also help decision makers to assess economic evaluations. This would contribute to bridging the gap between researchers and policymakers and would most likely increase faith in model-based economic evaluations in particular.

Five-year view

We expect that minimizing biases in model-based economic evaluation will become increasingly important. This is first because modelling studies are becoming more popular due to several advantages they have over trial-based evaluations. Second, output of health economic models is increasingly needed in order to provide policymakers with reliable information to allocate the scarce resources in healthcare systems. More awareness about bias needs to be raised, and further research should therefore be done in order to continuously improve and further develop the checklist. To sum up, we expect that a correct handling of biases will be a crucial indicator for quality in and acceptance of model-based economic evaluation in future health economic research.
Key issues

- Several studies have highlighted the importance of strong methodology by developing guidelines for model-based economic evaluations.
- Bias can lead to over- or underestimation of the true intervention effect and is different from imprecision in the estimated treatment effect, which implies random rather than systematic error.
- Eleven key biases that could occur in model-based economic evaluations were identified and their potential impact on cost-effectiveness results was illustrated using examples from the literature.
- Bringing the results from the previous and present article together, a comprehensive checklist on bias in economic evaluation (ECOBIAS checklist) is presented.
- The ECOBIAS checklist can be used by researchers to document the quality of their work and can thereby increase the acceptance of economic evaluations.
References


Part II

Economic evaluation studies
Chapter 5

Cost-effectiveness of ACE inhibitors and ARBs in newly diagnosed type 2 diabetes in Germany

Adarkwah CC
Gandjour A

Int J Technol Assess Health Care 2010;26:62-70
Abstract

Objective
Type 2 diabetes is the main cause of end-stage renal disease (ESRD) in Europe and the USA. Angiotensin-converting enzyme (ACE) inhibitors slow down the progression of renal disease and therefore provide a renal-protective effect. The aim of this study was to assess the most cost-effective time to start an ACE inhibitor (or an angiotensin II receptor blocker (ARB) in the event of cough) in patients with type 2 diabetes in Germany.

Methods
Three strategies were compared: treating all patients at the time of diagnosing type 2 diabetes, screening for microalbuminuria, and screening for macroalbuminuria. A lifetime Markov decision model with simulated 50-year-old patients with newly diagnosed diabetes mellitus was developed using published data on costs and health outcomes and simulating the progression of renal disease. A statutory health insurance (SHI) perspective was adopted.

Results
In the base-case analysis, the treat-all strategy is associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. A multivariate sensitivity analysis shows that the probability of savings is 89%.

Conclusions
Patients with type 2 diabetes should receive an ACE inhibitor immediately after diagnosis if they do not have contraindications. The potential for cost savings would be even larger if the prevention of cardiovascular events were considered.
Introduction

Type 2 diabetes is the main cause of end-stage renal disease (ESRD) in Germany and the United States. Diabetic nephropathy leads to a gradual decline of the renal function and is initially characterized by micro- or macroalbuminuria. Diabetic nephropathy may progress to ESRD. In Germany, the cost of ESRD treatment is about €40 000 per patient year. Hence, prevention of ESRD is also important from an economic viewpoint.

ACE inhibitors slow down the progression of diabetic nephropathy independent of an elevated blood pressure. ARBs have similar effects on renal outcomes in diabetic patients, but are more expensive. Evidence suggests that the only major clinical difference between these classes of drugs is a higher risk of dry cough associated with ACE inhibitors.

Several clinical practice guidelines recommend starting ACE inhibitor therapy in diabetic patients with (micro)albuminuria. However, physician compliance in Germany is rather low. Cost-effectiveness models conducted in the United States by Golan et al. and Rosen et al. suggest that the best starting point is immediately after diagnosis of diabetes. However, results may not be transferable to the German setting, particularly due to differences in prices. Furthermore, the model by Golan et al. was criticized for presenting limited evidence on the transitions between the different renal disease stages (this criticism equally applies to Rosen et al.). Moreover, both models do not consider that 9.9% of patients develop cough as a side effect of ACE-inhibitor therapy and thus are eligible for more expensive ARB therapy as a well-tolerated alternative.

The goal of this paper was to present a cost-effectiveness model which determines the best time to start an ACE inhibitor in newly diagnosed patients with type 2 diabetes and without hypertension or heart failure in Germany. The analysis was conducted from the perspective of the German SHI, which covers approximately 90% of the German population. In our model we included ARBs as an alternative for patients who experience ACE-inhibitor-induced cough. In the base case we used 50 years as the age of diagnosing type 2 diabetes. In contrast to Golan et al. and Rosen et al., we
considered broader evidence for the transition between the different renal disease stages, based on two Cochrane reviews.\textsuperscript{20,21}

Methods

Overview and model design

We conducted a cost-utility analysis and measured health outcomes in terms of quality-adjusted life years (QALYs). We developed a Markov model in order to simulate the course of a cohort of 1000 patients at the age of 50 years as it progresses to microalbuminuria, macroalbuminuria, ESRD, and death. A Markov model is an iterative process where patients are assumed to stay in one cycle (i.e., a defined health state) for a certain time and then make a transition to another cycle. Markov models are useful when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. We chose a cycle length of one year for the health states defined by the Markov model because all transition probabilities gathered from the literature referred to a duration of at least one year (see Table 5.1). Our Markov model contains the following 5 health states (Figure 5.1), which represent the occurrence of events after model entry:

1. type 2 diabetes, with normoalbuminuria (excretion < 30 mg/d)
2. type 2 diabetes, with microalbuminuria (excretion 30-300 mg/d)
3. type 2 diabetes, with macroalbuminuria (excretion >300 mg/d)
4. ESRD (treated with dialysis or renal transplantation)
5. death

We assumed that diabetic nephropathy progresses without skipping any stage. Further, patients may transit to death at any time. During each cycle, patients accumulate utility (measured by QALYs) and costs. A half-cycle correction was applied to both costs and outcomes to allow for transition events occurring mid-way through each 12-month cycle.

The simulation was done until the age of 100. Hence, the time horizon is 50 years. The age of 100 was chosen as a cutting point as there are no mortality
data available beyond this age. Regardless, 99.4 % of patients in the simulation are dead at this age.

Table 5.1 Data used to determine the cost-effectiveness of ACE inhibitors (ARBs) in newly diagnosed type 2 diabetes.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-case estimate</th>
<th>Range tested</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial disease prevalence, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normoalbuminuria</td>
<td>79</td>
<td>66.5 – 100</td>
<td>22</td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>18</td>
<td>0 – 27.6</td>
<td></td>
</tr>
<tr>
<td>Macroalbuminuria</td>
<td>3</td>
<td>0 – 5.9</td>
<td></td>
</tr>
<tr>
<td>Annual transition probabilities (without ACE inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normoalbuminuria to microalbuminuria</td>
<td>0.056</td>
<td>0.03 – 0.08</td>
<td>21</td>
</tr>
<tr>
<td>Microalbuminuria to macroalbuminuria</td>
<td>0.094</td>
<td>-0.02 – 0.20</td>
<td>20</td>
</tr>
<tr>
<td>Macroalbuminuria to ESRD</td>
<td>0.056</td>
<td>0.025 – 0.08</td>
<td>8</td>
</tr>
<tr>
<td>Normo-/micro-/macro-albuminuria to death</td>
<td>Age-dependent</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>ESRD to death</td>
<td>0.13</td>
<td>---</td>
<td>1</td>
</tr>
<tr>
<td>Relative risk for progression with ACE inhibitors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normoalbuminuria to microalbuminuria</td>
<td>0.60</td>
<td>0.43 – 0.84</td>
<td>21</td>
</tr>
<tr>
<td>Microalbuminuria to macroalbuminuria</td>
<td>0.45</td>
<td>0.29 – 0.69</td>
<td>20</td>
</tr>
<tr>
<td>Macroalbuminuria to ESRD</td>
<td>0.61</td>
<td>0.50 – 0.75</td>
<td>8</td>
</tr>
<tr>
<td>Utilities (health states)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes (baseline health)</td>
<td>0.88</td>
<td>0.86 – 0.90</td>
<td>23</td>
</tr>
<tr>
<td>ESRD</td>
<td>0.62</td>
<td>0.39 – 0.84</td>
<td>24</td>
</tr>
<tr>
<td>ACE inhibitor / ARB treatment</td>
<td>1.00</td>
<td>0.95 – 1.00</td>
<td>25</td>
</tr>
<tr>
<td>Annual costs, €</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per-patient cost of diabetes compared to non-diabetic population (across all ages)</td>
<td>2507</td>
<td>---</td>
<td>26</td>
</tr>
<tr>
<td>ACE inhibitor (10 mg ramipril daily)</td>
<td>48.22</td>
<td>---</td>
<td>27</td>
</tr>
<tr>
<td>ARB (300 mg irbesartan daily)</td>
<td>363.25</td>
<td>---</td>
<td>28,27</td>
</tr>
<tr>
<td>Mixed drug therapy costs (10% treated with ARBs)</td>
<td>79.72</td>
<td>---</td>
<td>27</td>
</tr>
<tr>
<td>Creatinine and potassium testing</td>
<td>1.00</td>
<td>---</td>
<td>3,29</td>
</tr>
</tbody>
</table>
Table 5.1 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-case estimate</th>
<th>Range tested</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening for microalbuminuria</td>
<td>6.90</td>
<td>3; applied to patients with normoalbuminuria only</td>
<td>3</td>
</tr>
<tr>
<td>Screening for macroalbuminuria</td>
<td>0.50</td>
<td>3; applied to patients with microalbuminuria only</td>
<td>3</td>
</tr>
<tr>
<td>ESRD</td>
<td>40 493</td>
<td>32 394 – 48 591</td>
<td>1,3–5,10</td>
</tr>
<tr>
<td>Transplantation</td>
<td>19 958</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Dialysis</td>
<td>48 088</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Home/in-center hemodialysis</td>
<td>41 463</td>
<td>---</td>
<td>1,3,4</td>
</tr>
<tr>
<td>Continuous ambulatory peritoneal dialysis</td>
<td>39 627</td>
<td>---</td>
<td>1</td>
</tr>
<tr>
<td>Continuous cycling peritoneal dialysis</td>
<td>49 510</td>
<td>---</td>
<td>1,4</td>
</tr>
<tr>
<td>Dialysis on partly in-patient treatment</td>
<td>98 282</td>
<td>---</td>
<td>5</td>
</tr>
<tr>
<td>Inpatient dialysis</td>
<td>147 021</td>
<td>---</td>
<td>5</td>
</tr>
<tr>
<td>SMR</td>
<td>1.41</td>
<td>1.39 – 1.43</td>
<td>31</td>
</tr>
<tr>
<td>Rate of ARB use, %</td>
<td>9.9</td>
<td>9.6 – 10.2</td>
<td>10</td>
</tr>
<tr>
<td>Specificity of HPLC (microalbuminuria screening procedure)</td>
<td>1.00</td>
<td>0.81 – 1.00</td>
<td>32</td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.03</td>
<td>0.00 – 0.07</td>
<td>33</td>
</tr>
</tbody>
</table>

ACE = angiotensin-converting enzyme; ARB = angiotensin II receptor blocker; ESRD = end-stage renal disease; SMR = standardized mortality ratio; HPLC = high performance liquid chromatography.

Figure 5.1  Schematic representation of the Markov decision model.
Clinical strategies

Three starting points for ACE inhibitors were considered.\textsuperscript{15,16} In the “screen for microalbuminuria” strategy patients are screened for microalbuminuria once a year and treatment is started if the test result is positive. In the “screen for macroalbuminuria” strategy patients are screened for macroalbuminuria once a year and treatment is also started if the test result is positive. In the “treat all” strategy no screening is performed at all and patients start on ACE inhibitor therapy at the time of diagnosing type 2 diabetes. The states of albuminuria were defined according to the recommendations of the American diabetes Association.\textsuperscript{11} In addition, we considered a no-screening and no-treatment alternative representing the status quo. In order to find information on the distribution of health states at the time of diagnosis we used the following search strategy in PubMed (1966-2007): (newly diagnosed[All Fields]) AND macroalbuminuria[All Fields] AND microalbuminuria[All Fields] AND prevalence[All Fields] AND (albumin excretion [All Fields]) NOT (type 1 diabetes [All Fields]). We found two studies. One study was excluded because it was conducted among Pima Indians. The other one is a Finnish prospective observational study\textsuperscript{22} which was conducted from 1982 to 1992. In this study, the distribution of health states at the time of diagnosis (average age: 58 years) was as follows: 79% normoalbuminuria, 18% microalbuminuria, and 3% macroalbuminuria. We tested the impact of the distribution on results in a sensitivity analysis.

Transition probabilities

In order to identify studies on the effectiveness of ACE inhibitor or ARB therapy on the prevention of diabetic kidney disease we searched in the Cochrane Database of Systematic Reviews using the search strategy normoalbuminuria OR microalbuminuria OR macroalbuminuria. We found two meta-analyses proving evidence that ACE inhibitors halt the transition from normo- to microalbuminuria and micro- to macroalbuminuria.\textsuperscript{20,21} These meta-analyses pooled studies on patients with type 1 and type 2 diabetes, as heterogeneity did not appear to be significant. Compared to placebo, ACE inhibitors significantly reduced the development of microalbuminuria (six trials, 3840 patients: relative risk (RR) 0.60, 95% confidence interval (CI) 0.43 to
and the progression from microalbuminuria to macroalbuminuria (17 trials, 2036 patients: RR 0.45, 95% CI 0.29 to 0.69).

In order to identify studies on the effectiveness of ACE inhibitors or ARBs on the transition from macroalbuminuria to ESRD, we checked evidence-based clinical practice guidelines on the prevention of diabetic nephropathy. We found one randomized clinical trial which was rated as a well-designed randomized controlled trial, providing high grade evidence. In this trial captopril significantly reduced the development of ESRD compared to placebo (409 patients, RR 0.61, 95% CI 0.50 to 0.75).

All input parameters including transition probabilities and the effectiveness of ACE inhibitor therapy are shown in Table 5.1. To determine annual transition probabilities we first calculated a total probability for each arm, by dividing the number of events (ESRD) during the trial period by the number of patients. Next, we determined annual transition probabilities by assuming a constant annual hazard rate over the study time horizon. A constant hazard rate yields an exponential survival curve.

In patients with normo-, micro-, and macroalbuminuria mortality is a function of age and was calculated by multiplying age-specific mortality rates of the general population with a standardized mortality ratio for patients with diabetes compared to the general population. For patients with normo-, micro-, and macroalbuminuria we assumed that mortality is stage-independent as there are no valid data showing that a significant difference exists. For ESRD patients we calculated the annual mortality rate based on 87151 patients in Germany, by dividing the annual number of decedents by the total number of patients.

Preference weights

We took preference weights of diabetic patients (Table 5.1) from a published cross-sectional study. Adult diabetic patients (n=292) with a disease duration of at least one year and a mean age of 62 years (range 21–85) were interviewed by the time tradeoff (TTO) method. We assumed that patients with normo-, micro-, or macroalbuminuria do not suffer from an additional
reduction in health-related quality of life.23 The preference weight for ESRD was taken from a systematic review of empirical studies in which TTO weights were provided by patients.24 The TTO is the most commonly used method to elicit quality-of-life weights for QALYs.39,40

Costs
Given the perspective of the German SHI, only direct costs relevant to the SHI were considered. Furthermore, a portion of drug expenditures needs to be paid out of pocket. Costs were inflated to year 2006 Euros using data on the consumer price index.30 Costs of ACE inhibitors, ARBs, annual screening procedures, and treatment for ESRD as well as health care expenditures related and unrelated to diabetes were taken into account. For ACE inhibitor therapy we used a daily dose of 10 mg of the most frequently prescribed ACE inhibitor, which is ramipril.41 In particular, we used the cheapest generic version of ramipril.27 A daily dose of 10 mg is at the high end of the therapeutic range.42-44. In a sensitivity analysis we applied the reference price of ramipril. For ARBs we applied 300 mg irbesartan daily,27 which is more effective in renal protection than a dose of 150 mg.28 As irbesartan is not a generic drug and still protected by patent, patients have to pay €37.24 per year out of pocket.45 In addition, we included costs for biannual monitoring of serum creatinine and potassium levels due to ACE inhibitor or ARB therapy.29,3 For microalbuminuria screening we preferred a quantitative screening test (high performance liquid chromatography) over a semi quantitative one (e.g., Micral-Test®)46 because it has a higher sensitivity (100%46) and specificity (81-98%32) and is also recommended by a recently published health technology assessment report.47 In the base-case analysis we assumed a specificity of 100%, which is conservative because treating false positives (i.e., patients with normoalbuminuria) leads to cost savings. In the sensitivity analysis we applied a specificity of 81%.2 To screen for macroalbuminuria we used a dipstick test applied in a primary care physician’s office.47 For patients without renal failure we included age-specific health care expenditures unrelated to diabetes48 as well as health care expenditures related to diabetes as an average across all ages.26 We validated the use of average-expenditure data against the use of age-specific expenditures data in patients diagnosed at the age of 60 years (these age-specific data were not available below the
age of 60 years). We also assessed whether explicitly modeling the costs of survivors and decedents (as opposed to using average-cost data) changes the cost-effectiveness of ACE inhibitors. This approach accounts for the fact that the high costs of those who die dominate age-specific costs. Hence, when the last year of life is postponed, average age-specific costs fall. Age-related expenditure data of diabetic patients were available from a large German statutory health insurer (unpublished data).

Annual costs of patients with ESRD were calculated as a weighted average of the costs of patients with different types of dialysis as well as renal transplantation and include costs unrelated to dialysis and renal transplantation. In detail, the following calculations were made:

1. cost of dialysis = $\beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 = X$, where $x_n =$ annual cost of dialysis treatment $n = 1,2,\ldots,5$, $\beta_n =$ prevalence weight of the dialysis treatment, and $\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 = 1$


Prevalence data served as weights. Data on the cost of continuous ambulatory peritoneal dialysis and continuous cycling peritoneal dialysis were obtained from an empirical survey by Nebel. To determine the cost of home/in-center hemodialysis we used the German fee schedule “Uniform value scale” to calculate costs of the dialysis itself and the survey by Nebel to determine all other expenditures for dialysis patients. Health care expenditures related and unrelated to diabetes were both included. See Table 5.2 for parameters used in the analysis.

Costs and benefits were discounted at an annual rate of 3%.33
Table 5.2 Parameters used for calculating the cost of end-stage renal disease (see cost section under “Methods”).

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Cost (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home/center hemodialysis (1)</td>
<td>41,463</td>
</tr>
<tr>
<td>Hemodialysis on inpatient treatment (2)</td>
<td>147,021</td>
</tr>
<tr>
<td>Hemodialysis on partly inpatient treatment (3)</td>
<td>98,282</td>
</tr>
<tr>
<td>Continuous ambulatory peritoneal dialysis (CAPD) (4)</td>
<td>39,627</td>
</tr>
<tr>
<td>Continuous cycling peritoneal dialysis (CCPD) (5)</td>
<td>49,510</td>
</tr>
</tbody>
</table>

Sensitivity analysis

To address uncertainty around mean incremental costs and effectiveness, we conducted univariate sensitivity analyses. Whenever possible, we ran the analysis using the upper and lower bound of the 95% CI of the mean. For the annual cost of ESRD we did not have information on the 95% CI and thus varied costs by 20% (see Table 5.1 for uncertainty ranges).

In order to assess how a simultaneous change of several variables affects the incremental cost-effectiveness ratio (ICER), we performed a Monte Carlo simulation, a type of multivariate sensitivity analysis. This technique runs a large number of simulations (here: 1000) by repeatedly drawing samples from probability distributions of input variables. Thus, it provides a probability distribution for the output variables, i.e., incremental costs and effectiveness. Probabilities and relative risks were assumed to follow a beta distribution Beta(α, β) because they are restricted to take on values between 0 and 1. Because the distribution of health states at the time of diagnosis had more than 2 outcomes, we assumed a Dirichlet distribution Dirichlet(α1, α2, . . . , αn).
Cost data were assumed to follow a gamma distribution Gamma(α, β) because they are normally distributed but restricted to take on values between 0 and 1. The standard deviation of probabilities and relative risks was calculated according to the following formula:

\[ \sigma = \sqrt{\frac{p(1 - p)}{n}} \]

Given the ambiguous interpretation of negative ICERs, we transformed ICERs into net monetary benefits (NMBs). We generated a cost-effectiveness acceptability curve based on the distribution of NMBs for each value of the willingness to pay per QALY gained. A cost-effectiveness acceptability curve allows a decision maker to consider whether a prevention strategy is cost-effective in relation to the maximum amount a decision-maker is willing to pay for a QALY. At each ceiling value for the willingness to pay for a QALY, the cost-effectiveness acceptability curve shows the probability that treatment is cost-effective.

Results

Base-case analysis

The base-case analysis, which applies to 50-year-old patients, shows that “no screening and treatment”, “screening for macroalbuminuria”, and “screening for microalbuminuria” are all dominated by the “treat all” strategy, which is associated with the lowest costs and highest benefit (Table 5.3). Treating all patients with ACE inhibitors at the age of 60 years also leads to savings. Comparing the use of average-expenditure data with age-specific expenditure data and separate expenditure data for survivors and decedents yields little difference in cost savings compared to screening for microalbuminuria (-€1 852, -€1 875, and -€1 982, respectively).
Cost-effectiveness of ACE inhibitors and ARBs in newly diagnosed type 2 diabetes in Germany

Table 5.3 Results of the base-case analysis.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Costs (€)</th>
<th>Undiscounted LYs</th>
<th>Discounted QALYs</th>
<th>ICER (€/QALY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No screening and treatment</td>
<td>151 579</td>
<td>25.96</td>
<td>14.46</td>
<td>Dominated</td>
</tr>
<tr>
<td>Screening for macroalbuminuria</td>
<td>144 059</td>
<td>26.64</td>
<td>14.83</td>
<td>Dominated</td>
</tr>
<tr>
<td>Screening for microalbuminuria</td>
<td>137 406</td>
<td>27.19</td>
<td>15.14</td>
<td>Dominated</td>
</tr>
<tr>
<td>Treating all patients with ACEIs/ARBs</td>
<td>135 555</td>
<td>27.32</td>
<td>15.21</td>
<td>Dominant</td>
</tr>
</tbody>
</table>

ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin II receptor blocker; LYs = life years; QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio.

Sensitivity analysis

In the univariate sensitivity analysis, variables with the largest impact on incremental costs and effectiveness are the discount rate, the absolute risk for progression from micro- to macroalbuminuria without ACE inhibition as well as the relative risk for progression from normo- to microalbuminuria with ACE inhibitor therapy (Table 5.4). When assuming a low progression rate from micro- to macroalbuminuria without ACE inhibition, screening for microalbuminuria dominates the treat-all strategy. A threshold sensitivity analysis shows that at an annual drug cost of €309.40 (base case: €75.72) the breakeven point is reached.

Table 5.4 Univariate sensitivity analyses: effects of varying base-case estimates on the incremental cost-effectiveness ratio of treating all patients with ACE inhibitors vs screening for microalbuminuria (reference strategy). “Lower bound” and “higher bound” refer to the limits of the 95% confidence interval.

<table>
<thead>
<tr>
<th>Incremental costs</th>
<th>Incremental QALYs</th>
<th>Incremental cost-effectiveness ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial disease prevalence: Proportion of normoalbuminuric patients, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-€1852</td>
<td>0.08</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-€2344</td>
<td>0.33</td>
</tr>
<tr>
<td>Annual transition probabilities (without ACE inhibitors) from normo- to microalbuminuria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-€1044</td>
<td>0.05</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-€2309</td>
<td>0.10</td>
</tr>
<tr>
<td>from micro- to macroalbuminuria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>€1253</td>
<td>-0.02</td>
</tr>
<tr>
<td>Higher bound</td>
<td>€3202</td>
<td>0.13</td>
</tr>
<tr>
<td>from macroalbuminuria to ESRD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-€693</td>
<td>0.04</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-€2522</td>
<td>0.10</td>
</tr>
<tr>
<td>Relative risk for progression with ACE inhibitors</td>
<td>Incremental costs</td>
<td>Incremental QALYs</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------</td>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>from normo- to microalbuminuria</td>
<td>Lower bound - €3149 0.11 Dominance</td>
<td>Higher bound - €300 0.03 Dominance</td>
</tr>
<tr>
<td>from micro- to macroalbuminuria</td>
<td>Lower bound - €1186 0.05 Dominance</td>
<td>Higher bound - €2590 0.11 Dominance</td>
</tr>
<tr>
<td>from macroalbuminuria to ESRD</td>
<td>Lower bound - €1516 0.06 Dominance</td>
<td>Higher bound - €2231 0.09 Dominance</td>
</tr>
<tr>
<td>Utilities (health states)</td>
<td>Diabetes (baseline health) Lower bound - €1852 0.07 Dominance</td>
<td>Higher bound - €1852 0.08 Dominance</td>
</tr>
<tr>
<td></td>
<td>ESRD Lower bound - €1852 0.10 Dominance</td>
<td>Higher bound - €1852 0.05 Dominance</td>
</tr>
<tr>
<td>Disutility of ACE inhibitor treatment</td>
<td>Lower bound - €1852 0.04 Dominance</td>
<td>Higher bound - €1852 0.08 Dominance</td>
</tr>
<tr>
<td>Costs</td>
<td>ACE inhibitor Lower bound -€1852 0.08 Dominance</td>
<td>Higher bound -€1783 0.08 Dominance</td>
</tr>
<tr>
<td></td>
<td>ESRD Lower bound - €1050 0.08 Dominance</td>
<td>Higher bound - €2643 0.08 Dominance</td>
</tr>
<tr>
<td></td>
<td>SMR Lower bound - €1855 0.08 Dominance</td>
<td>Higher bound - €1848 0.08 Dominance</td>
</tr>
<tr>
<td>Rate of ARB use</td>
<td>Lower bound - €1862 0.08 Dominance</td>
<td>Higher bound - €1847 0.08 Dominance</td>
</tr>
<tr>
<td>Specificity of HPLC (microalbuminuria screening procedure)</td>
<td>81% - €1549 0.06 Dominance</td>
<td></td>
</tr>
<tr>
<td>Discount rate of costs and effects</td>
<td>0% - €4543 0.17 Dominance</td>
<td>3% (base case) - €1852 0.08 Dominance</td>
</tr>
<tr>
<td></td>
<td>5% - €1035 0.05 Dominance</td>
<td>7% - €582 0.03 Dominance</td>
</tr>
</tbody>
</table>

QALYs = quality-adjusted life years; ACE = angiotensin-converting enzyme; ESRD = end-stage renal disease; SMR = standardized mortality ratio; HPLC = high performance liquid chromatography.
Changing the distribution of health states at the time of diagnosis to all patients being normoalbuminuric yields little difference in cost savings (-€2 344 vs. -€1 852) and QALYs gained (0.10 vs. 0.08).

Figure 5.2 shows the cost-effectiveness acceptability curve, which considers uncertainty in cost-effectiveness. The probability of savings of the “treat all” strategy compared to screening for microalbuminuria is 89% (see also Figure 5.3 for the scatterplot).
Discussion

This modeling study shows that treating all patients with type 2 diabetes with ACE inhibitors (and more expensive ARBs in the event of cough) immediately after diagnosis is cost-effective and even reduces health care expenditures. Our results were robust to a variety of different assumptions of uncertainty. Although a significant number of newly diagnosed type 2 patients may receive blood pressure medications, it has not been shown that these patients are primarily prescribed an ACE inhibitor, which underlines the significance of this analysis.

Our model is far from being perfect, but in modeling studies this is rarely the case due to constraints of resources, time, and information availability. Savings by treating all diabetic patients with ACE inhibitors may be even underestimated for several reasons. First, we did not model that ACE inhibitors and ARBs reduce the risk for cardiovascular events which leads to additional savings. Second, we did not consider that some patients discontinue ACE inhibitor therapy and thus incur no drug costs. The reason is that information on compliance was unavailable from the trials on ACE inhibitors. On the other hand, the rate of compliance is incorporated in the effectiveness data (thus reducing effectiveness). Third, patients may also be sufficiently treated by a smaller dose of ramipril than in our analysis. We used a dose of 10 mg, which is at the high end of the therapeutic range in studies investigating the effect of ACE inhibitors on renal outcomes. Forth, for patients who receive dialysis on an inpatient basis (2% of all patients with renal replacement therapy) we were not able to consider health care costs unrelated to dialysis due to a lack of available data. This underestimates the savings from preventing ESRD. Fifth, costs of dialysis treatment will likely continue to rise in the future, thus increasing the potential for savings by preventing ESRD. In the last 10 years dialysis costs have increased by approximately 25 percent and we expect this trend to continue due to stricter regulations concerning dialysis safety, technological advancement of dialysis machines, and better-tolerated dialysis solutions. Alleviating this trend may be lower prices of erythropoietin and lower reimbursement of physicians. Finally, we calculated annual transition rates based on trial duration instead of follow-up time as data concerning follow-up time were
unavailable for most trials. As follow-up time is generally shorter than trial duration this might have underestimated transition probabilities as well as cost savings.

Further limitations of the model relate to the data sources. First, when calculating the weighted average cost of treating ESRD we used prevalence data on intermittent and automated peritoneal dialysis (IPD/APD) to weigh the cost of continuous cycling peritoneal dialysis (CCPD). Second, the model uses some epidemiological data from Western countries other than Germany. For example, we used a Finnish study as the source of the distribution of health states at the time of diagnosis. However, changing the initial distribution of health states had little impact on the results. Third, transition rates from macroalbuminuria to ESRD with and without ACE inhibitors were not available for patients with type 2 diabetes. Therefore, we used a randomized controlled trial in patients with type 1 diabetes as the source. Fourth, the standardized mortality ratio (SMR) we applied to diabetic patients without ESRD includes patients with ESRD. Excluding these patients would lower the SMR, but only to a minor degree, as less than 1% of the German diabetic population receives renal replacement therapy. Fifth, we assumed that the SMR is the same for patients with normo-, micro-, and macroalbuminuria as there are no valid data showing that a significant difference exists. The slightly higher mortality ratio in microalbuminuric patients in the HOPE study was most likely the result of prior cardiovascular events. There is no evidence in the literature that mortality rates increase only on the basis of the level of albumin in the urine. Sixth, having microalbuminuria or macroalbuminuria might cause disutility from anxiety. However, standard preference measures such as the SG or the TTO method are not able to capture anxiety over future events as both evaluation methods assume a constant (!) health state over the remaining period of life. Seventh, the definition of hypertension was not consistent in the studies we incorporated in our model. Cut-off values varied between 130/80 and 160/90 mm Hg. Finally, it is unclear whether a societal perspective leads to larger or smaller savings than a SHI perspective: on the one hand, ACE inhibitors and ARBs avoid productivity loss and copayments due to renal failure. On the other hand, drug copayments lead to additional costs.
Compared to previous cost-effectiveness models, which were conducted by Golan et al.\textsuperscript{15} and Rosen et al.\textsuperscript{16} based on U. S. data, we considered a much broader evidence base for the transition between normo- to microalbuminuria and micro- to macroalbuminuria. In addition, we considered that patients who are noncompliant with ACE inhibitors due to cough may receive more expensive ARBs. We did not consider that a small proportion of patients on ARBs (3.2\%) also develop cough\textsuperscript{10} and thus may discontinue treatment. The reason is that noncompliance with treatment is already incorporated in the relative risk of treatment (thus lowering the relative risk), as in RCTs a certain proportion of patients discontinued treatment. We assumed equal effectiveness of all ACE inhibitors and ARBs, as meta-analyses do not suggest any independent effect of single renin-angiotensin-system agents.\textsuperscript{52,20} Still, similar to the model by Rosen et al.\textsuperscript{16}, our model shows that treating all newly diagnosed type 2 patients with ACE inhibitors saves costs. In contrast to Rosen et al., we did not consider the preventive effect of ACE inhibitors on cardiovascular outcomes, which would have increased savings. An important reason for the large savings potential in Germany is the low price of ramipril, which has significantly gone down during the last few years.\textsuperscript{27,53} ARBs are still protected by patent for a few additional years. Irbesartan, which we included in our study, is protected until March 2012. Assuming that prices of ARBs will decline after expiration of the patent protection would further strengthen our conclusion.

Policy Considerations

We recommend that all patients with type 2 diabetes covered by the SHI receive an ACE inhibitor at the higher end of the therapeutic dose range immediately after diagnosis. In case of cough, patients should be prescribed an ARB. This is a major step from current practice, which is not to screen for microalbuminuria. Finally, the transferability of the conclusion of this study to other countries is limited given differences in costs, clinical management, and epidemiology (for example, with regard to the incidence of unrelated diseases).
References

44. The EUCLID study group. Randomised placebo-controlled trial of lisinopril in normotensive patients with insulin-dependent diabetes and normoalbuminuria or microalbuminuria. Lancet 1997;349:1787-1792.

Chapter 6

Cost-effectiveness of angiotensin-converting enzyme inhibitors for the prevention of diabetic nephropathy in the Netherlands – A Markov model

Adarkwah CC
Gandjour A
Akkerman M
Evers S

PLoS ONE 2011;6:e26139
Abstract

Objective
Type 2 diabetes is the main cause of end-stage renal disease (ESRD) in Europe and the USA. Angiotensin-converting enzyme (ACE) inhibitors have a potential to slow down the progression of renal disease and therefore provide a renal-protective effect. The aim of our study was to assess the most cost-effective time to start an ACE inhibitor (or an angiotensin II receptor blocker (ARB) if coughing as a side effect occurs) in patients with newly diagnosed type 2 diabetes in the Netherlands.

Methods
A lifetime Markov decision model with simulated 50-year-old patients with newly diagnosed diabetes mellitus was developed using published data on costs and health outcomes and simulating the progression of renal disease. A health insurance perspective was adopted.
Three strategies were compared: treating all patients at the time of diagnosing type 2 diabetes, screening for microalbuminuria, and screening for macroalbuminuria.

Results
In the base-case analysis, the treat-all strategy is associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. A multivariate sensitivity analysis shows that the probability of savings is 70%.

Conclusion
In the Netherlands for patients with type 2 diabetes prescription of an ACE inhibitor immediately after diagnosis should be considered if they do not have contraindications An ARB should be considered for those patients developing a dry cough under ACE inhibitor therapy. The potential for cost savings would be even larger if the prevention of cardiovascular events were considered.
Introduction

The prevalence of type 2 diabetes and its secondary complications will rise due to ageing population and growing obesity. This type of diabetes represents the most common form of carbohydrate disorders affecting at least 5% of the population in the industrialized world. As a result higher costs for diabetes treatment in general and especially treatment of secondary complications will be a huge burden for health care systems. Type 2 diabetes is the main cause of end-stage renal disease (ESRD) in the Netherlands as well as in other European countries and the United States. Diabetic nephropathy leads to a gradual decline of the renal function and is initially characterized by micro- or macroalbuminuria. Diabetic nephropathy may progress to ESRD, which is defined by the need for either long-term dialysis or renal transplantation. The prevalence of patients in renal replacement therapy in the Netherlands doubled within the last 15 years. In 2010, about 15,000 patients underwent renal-replacement therapy. In the last five years, the proportion of transplanted patients has been continuously increasing and represents about 57% of all patients requiring renal replacement therapy. The costs of ESRD treatment are rather high, with a share of the national expenditures in European countries ranging from 0.7% in the UK to 1.8% in Belgium, with a share in the Netherlands of about 1.3%. In the Netherlands, the costs of ESRD treatment amount to €42,000 per patient per year. Hence, prevention of ESRD is not only important from a medical, but also from an economic viewpoint.

Angiotensin converting enzyme (ACE) inhibitors slow down the progression of diabetic nephropathy independent of an elevated blood pressure. Angiotensin receptor blockers (ARBs) have similar effects on renal outcomes in diabetic patients but are more expensive, mostly due to patent protection. Evidence suggests that the only major clinical difference between these classes of drugs is a higher risk of dry cough associated with ACE inhibitors. Several national and international clinical practice guidelines recommend starting ACE inhibitor therapy in diabetic patients with (micro)albuminuria. However, physician compliance in the Netherlands as well as in many other European countries is rather low. Cost-effectiveness models conducted in the United States by Golan et al. (1999), Rosen et al. (2005) and in Germany...
by Adarkwah et al. (2010) suggest that the best starting point for ACE inhibitor therapy is immediately after diagnosis of diabetes. For the Netherlands no data are available on the cost-effectiveness of ACE inhibitor therapy in diabetic patients with (micro)albuminuria. However, results of the non-Dutch studies may not be transferable to the Netherlands. Transferability of economic evaluation studies between countries is hindered by a number of factors such as demography, the epidemiology of the disease, availability of health care resources and differences in reimbursement systems between countries, in particularly due to variances in absolute and relative costs/prices. The goal of this study is to present a cost-effectiveness model, which determines the best time to start an ACE inhibitor in newly diagnosed patients with type 2 diabetes and without hypertension or heart failure in the Netherlands. The analysis is conducted from a health care perspective in order to increase comparability to other models on this topic. In our model we included ARBs as an alternative for patients who experience ACE-inhibitor-induced cough. In the base case the age of 50 years was assumed as the mean age of diagnosing type 2 diabetes.

Methods

Overview and model design

Is it cost-effective to treat all newly diagnosed type 2 diabetic patients in the Netherlands with an ACE inhibitor to prevent renal disease? We conducted a cost-utility analysis and measured health outcomes in terms of quality-adjusted life years (QALYs). We adapted a Markov decision model previously developed for the German setting and also proven applicable for non-diabetic advanced renal disease in order to simulate the course of a cohort of 1 000 patients at the age of 50 years as it progresses to microalbuminuria, macroalbuminuria, ESRD, and death. A Markov model is an iterative process where patients are assumed to stay in one cycle (i.e., a defined health state) for a certain time and then make a transition to another cycle. Markov models are useful when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. The model was built in Microsoft Excel® 2007. We chose a cycle length of one year for the health states defined by the Markov
model because all transition probabilities gathered from the literature referred to a duration of at least one year. All input data included in the model can be found in Table 6.1. Our Markov model contains the following five health states (Figure 6.1), which represent the occurrence of events after model entry:

1. type 2 diabetes, with normoalbuminuria (excretion <30 mg/d)
2. type 2 diabetes, with microalbuminuria (excretion 30–300 mg/d)
3. type 2 diabetes, with macroalbuminuria (excretion >300 mg/d)
4. ESRD (treated with dialysis or renal transplantation)
5. Death

We assumed that diabetic nephropathy progresses without skipping any stage. Further, patients may die at any time (stage 5). The states of albuminuria were defined according to the recommendations of the American diabetes Association.20 During each cycle, patients accumulate utility (measured by QALYs) and costs. A half-cycle correction will be applied to both costs and outcomes to allow for transition events occurring mid-way through each 12-month cycle.

The simulation was done until the age of 99. Hence, the time horizon is 50 years. The age of 99 was chosen as a cutting point as there are no mortality data available beyond this age. Regardless, more than 99% of patients in the simulation are dead at this age.

Table 6.1 Data used to determine the cost-effectiveness of ACE inhibitors and ARBs in newly diagnosed type 2 diabetes.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-case estimate</th>
<th>Range tested*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial disease prevalence, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normoalbuminuria</td>
<td>79</td>
<td>66.5 – 100</td>
<td>[28]</td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>18</td>
<td>0 – 27.6</td>
<td></td>
</tr>
<tr>
<td>Macroalbuminuria</td>
<td>3</td>
<td>0 – 5.9</td>
<td></td>
</tr>
<tr>
<td>Annual transition probabilities (without ACE inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normoalbuminuria to microalbuminuria</td>
<td>0.056</td>
<td>0.03 – 0.08</td>
<td>[29]</td>
</tr>
<tr>
<td>Microalbuminuria to macroalbuminuria</td>
<td>0.094</td>
<td>-0.02 – 0.20</td>
<td>[30]</td>
</tr>
<tr>
<td>Macroalbuminuria to ESRD</td>
<td>0.056</td>
<td>0.025 – 0.08</td>
<td>[15]</td>
</tr>
<tr>
<td>Normo-/micro-/macro-albuminuria to death</td>
<td>Age-dependent</td>
<td>---</td>
<td>[35]</td>
</tr>
<tr>
<td>ESRD to death</td>
<td>0.09</td>
<td>---</td>
<td>[9]</td>
</tr>
</tbody>
</table>

Relative risk for progression with ACE inhibitors
Table 6.1 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-case estimate</th>
<th>Range tested*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normoalbuminuria to microalbuminuria</td>
<td>0.60</td>
<td>0.43 – 0.84</td>
<td>[29]</td>
</tr>
<tr>
<td>Microalbuminuria to macroalbuminuria</td>
<td>0.45</td>
<td>0.29 – 0.69</td>
<td>[30]</td>
</tr>
<tr>
<td>Macroalbuminuria to ESRD</td>
<td>0.61</td>
<td>0.50 – 0.75</td>
<td>[15]</td>
</tr>
<tr>
<td>Utilities (health states)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes (baseline health)</td>
<td>0.88</td>
<td>0.86 – 0.90</td>
<td>[38]</td>
</tr>
<tr>
<td>ESRD</td>
<td>0.62</td>
<td>0.39 – 0.84</td>
<td>[39]</td>
</tr>
<tr>
<td>ACE inhibitor / ARB treatment</td>
<td>1.00</td>
<td>0.95 – 1.00</td>
<td>[63]</td>
</tr>
<tr>
<td>Annual costs, €</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General health care expenditures</td>
<td>3.310,23 - 23.626,23</td>
<td>---</td>
<td>[43,44,55]</td>
</tr>
<tr>
<td>Per-patient cost of diabetes compared to non-diabetic population</td>
<td>547</td>
<td>---</td>
<td>[43,44,55] applied to all health states except for ESRD</td>
</tr>
<tr>
<td>ACE inhibitor (20mg enalapril daily)</td>
<td>6.96</td>
<td>---</td>
<td>[46] applied to all health states except for ESRD</td>
</tr>
<tr>
<td>ARB (300mg irbesartan daily)</td>
<td>298.68</td>
<td>---</td>
<td>[46,47] applied to all health states except for ESRD</td>
</tr>
<tr>
<td>Mixed drug therapy costs (9.9% treated with ARBs)</td>
<td>62.70</td>
<td>62.70 – 83.78</td>
<td>[46] applied to all health states except for ESRD</td>
</tr>
<tr>
<td>Screening for microalbuminuria</td>
<td>7.00</td>
<td>---</td>
<td>[58,59] applied to patients with normoalbuminuria only</td>
</tr>
<tr>
<td>Screening for macroalbuminuria</td>
<td>1.12</td>
<td>---</td>
<td>[58,59] applied to patients with microalbuminuria only</td>
</tr>
<tr>
<td>ESRD</td>
<td>42 110</td>
<td>33 688 – 50 532</td>
<td>[9,13]</td>
</tr>
<tr>
<td>Transplantation</td>
<td>14 387</td>
<td>---</td>
<td>[9,13]</td>
</tr>
<tr>
<td>Dialysis</td>
<td>79 112</td>
<td>---</td>
<td>[9,13]</td>
</tr>
<tr>
<td>Home/in-center hemodialysis</td>
<td>83 217</td>
<td>---</td>
<td>[9,13]</td>
</tr>
<tr>
<td>Continuous ambulatory peritoneal dialysis</td>
<td>54 067</td>
<td>---</td>
<td>[9,13]</td>
</tr>
<tr>
<td>Continuous cycling peritoneal dialysis</td>
<td>69 546</td>
<td>---</td>
<td>[9,13]</td>
</tr>
<tr>
<td>SMR</td>
<td>1.41</td>
<td>1.39 – 1.43</td>
<td>[1]</td>
</tr>
<tr>
<td>Rate of ARB use, %</td>
<td>9.9</td>
<td>9.6 – 10.2</td>
<td>[17]</td>
</tr>
<tr>
<td>Specificity of HPLC (microalbuminuria screening procedure)</td>
<td>1.00</td>
<td>0.81 – 1.00</td>
<td>[53]</td>
</tr>
<tr>
<td>Discount rate of costs</td>
<td>0.04</td>
<td>0.00 – 0.10</td>
<td>[44,54,55];</td>
</tr>
<tr>
<td>Discount rate of benefits</td>
<td>0.015</td>
<td>0.00 – 0.10</td>
<td>[44,54,55];</td>
</tr>
</tbody>
</table>

ACE = angiotensin-converting enzyme; ARB = angiotensin II receptor blocker; ESRD = end-stage renal disease; HPLC = high performance liquid chromatography; SMR = standardized mortality ratio
Clinical strategies

Three starting points for ACE inhibitors were considered. In the “screen for microalbuminuria” strategy patients are screened for microalbuminuria once a year and treatment is started if the test result is positive. In the “screen for macroalbuminuria” strategy patients are screened for macroalbuminuria once a year and treatment is also started if the test result is positive. In the “treat all” strategy no screening is performed at all and patients start on ACE inhibitor therapy at the time of diagnosing type 2 diabetes. In addition, the analysis performed included the ARB option for the entire patient population in all three strategies reflecting a more expensive treatment. To find information on the distribution of health states at the time of diagnosis, we used the following search strategy in the PubMed database (date: February 08, 2011): (newly diagnosed[All Fields]) AND macroalbuminuria[All Fields] AND microalbuminuria[All Fields] AND prevalence[All Fields] AND (albumin excretion [All Fields]) NOT (type 1 diabetes [All Fields]). We obtained 2 hits. Thereof one study was excluded because it was conducted among Pima Indians. The other one is a Finnish prospective observational study, which was conducted from 1982 to 1992. In this study, the distribution of health states at the time of diagnosis (average age: 58 years) was as follows: 79% normoalbuminuria, 18% microalbuminuria, and 3% macroalbuminuria. We tested the impact of the initial distribution on results in a sensitivity analysis.
Transition probabilities

In order to identify studies on the effectiveness of ACE inhibitor or ARB therapy on the prevention of diabetic kidney disease we searched in the Cochrane Database of Systematic Reviews using the search strategy *normoalbuminuria OR microalbuminuria OR macroalbuminuria*. We found two meta-analyses proving evidence that ACE inhibitors halt the transition from normo- to microalbuminuria and micro- to macroalbuminuria.29,30 These meta-analyses pooled studies on patients with type 1 and type 2 diabetes, as heterogeneity did not appear to be an issue. Compared to placebo, ACE inhibitors significantly reduced the development of microalbuminuria (six trials, 3 840 patients: relative risk (RR) 0.60, 95% confidence interval (CI) 0.43 to 0.84), and the progression from microalbuminuria to macroalbuminuria (17 trials, 2 036 patients: RR 0.45, 95% CI 0.29 to 0.69).

In order to identify studies on the effectiveness of ACE inhibitors or ARBs on the transition from macroalbuminuria to ESRD, evidence-based clinical practice guidelines were checked on the prevention of diabetic nephropathy.19,31-33 One randomized clinical trial15 was identified that was rated as well-designed randomized controlled trial (RCT)19,33 providing high-grade evidence. In this trial captopril significantly reduced the development of ESRD compared to placebo (409 patients, RR 0.61, 95% CI 0.50 to 0.75).

To determine annual transition probabilities we first calculated a total probability for each arm, by dividing the number of events (ESRD) during the trial period by the number of patients. Next, we determined annual transition probabilities by assuming a constant annual hazard rate over the study time horizon.34 A constant hazard rate yields an exponential survival curve.

In patients with normo-, micro-, and macroalbuminuria mortality is a function of age and was calculated by multiplying age-specific mortality rates of the Dutch general population35 with a standardized mortality ratio for patients with diabetes compared to the general population.1,36 For patients with normo-, micro-, and macroalbuminuria we assumed that mortality is stage-independent as there are no valid data showing that a significant difference exists. For patients with ESRD, we calculated the annual mortality rate based on 13 905 patients in the Netherlands,9 by dividing the annual number of decedents by the total number of patients. While the annual number of decedents treated with dialysis could be derived from the website, the
number of decedents with a transplant was obtained by personal communication (A. Hemke, Dutch End-Stage Renal Disease Registry, March 17, 2011).

Preference weights
We included preference weights of diabetic patients (Table 6.1) from a published cross-sectional study. Adult diabetic patients (n=292) with a disease duration of at least one year and a mean age of 62 years (range 21-85) were interviewed by the time trade-off (TTO) method. We assumed that patients with normo-, micro-, or macroalbuminuria do not suffer from an additional reduction in health-related quality of life. There is no convincing evidence in the literature that confirms a utility decrease merely due to albuminuria. The preference weight for ESRD was taken from a systematic review of empirical studies in which TTO weights were provided by patients. The TTO is the most commonly used method to elicit quality-of-life weights for QALYs. The TTO technique determines the proportion of remaining life years in poor health patients are willing to give up or trade in exchange for perfect health. Based on patient responses utility scores are calculated. Utility measures in economic evaluations are becoming increasingly important given the fact that decision makers are asked to optimize the allocation of scarce health care resources across disease areas and patient groups. Values are similar to EQ-5D scores (baseline value 0.61) reported by de Wit et al. (1998).

Costs
As stated, the analysis is conducted from the health care perspective. Hence, only direct costs and direct health effects – defined as life years gained – were considered. Costs were inflated to year 2010 euros using data on the consumer price index. Costs of ACE inhibitors, ARBs, annual screening procedures, and treatment for ESRD as well as health care expenditures related and unrelated to diabetes were taken into account. The recommendations of the Dutch guidelines for pharmacoeconomic research were followed. For ACE inhibitor therapy the most frequently prescribed ACE inhibitor in the Netherlands, enalapril, was taken into consideration. In the base case, the cheapest generic of enalapril 10 mg daily was used, whereas
the most expensive one was applied in the sensitivity analysis [46]. For ARBs we considered a dose of 300 mg irbesartan daily, which is more effective in renal protection than a dose of 150 mg. The costs of enalapril and irbesartan treatment were based on 2011 Dutch prices and include 6% value-added tax as well as a 3-monthly pharmacists’ prescription fee of €7.50. As recommended by a published health technology assessment (HTA) report and a national clinical chemistry report, a quantitative screening test for microalbuminuria (high performance liquid chromatography or immunoturbidimetric) was preferred over a semi quantitative one (e.g., Micral-Test®) because it demonstrates higher sensitivity (100%) and specificity (81-98%). Bakker et al. clearly state that a simple dipstick test is not sufficient to detect microalbuminuria at an early stage. In the base-case analysis we assumed a specificity of 100% which is conservative because treating false positives (i.e., patients with normoalbuminuria) leads to cost savings. In the sensitivity analysis we applied a specificity of 81%. To screen for macroalbuminuria we used a dipstick test applied in a general practitioner’s office recommended by the Dutch Kidney Check Campaign.

The annual costs of patients with ESRD were calculated as a weighted average of the costs of different types of dialysis as well as renal transplantation based on a Dutch study and prevalence data available from the national register.

In detail, the following calculations were made (see Table 6.2 in the appendix for details):

1. cost of dialysis = $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = X$, where $x_n =$ annual cost of dialysis treatment $n=1,2,3; \beta_n =$ prevalence weight of the dialysis treatment, and $\beta_1 + \beta_2 + \beta_3 = 1$


A transplant survival of 10 years was assumed and a distinction made between the first year of transplantation and the years following. Costs were inflated to 2010 Dutch prices. Health care expenditures related and unrelated to diabetes were both included. Costs were discounted at an annual rate of 4% whereas benefits
were discounted at an annual rate of 1.5% in accordance with the CVZ recommendations.44,54,55

Table 6.2 Parameters used for calculating the cost of end-stage renal disease (see cost section under “Methods”).

<table>
<thead>
<tr>
<th>variable</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>home/center hemodialysis</td>
</tr>
<tr>
<td>2</td>
<td>continuous ambulatory peritoneal dialysis (CAPD)</td>
</tr>
<tr>
<td>3</td>
<td>continuous cycling peritoneal dialysis (CCPD)</td>
</tr>
<tr>
<td>β1</td>
<td>0.82</td>
</tr>
<tr>
<td>β2</td>
<td>0.106</td>
</tr>
<tr>
<td>β3</td>
<td>0.074</td>
</tr>
<tr>
<td>x1</td>
<td>€ 83 217</td>
</tr>
<tr>
<td>x2</td>
<td>€ 54 067</td>
</tr>
<tr>
<td>x3</td>
<td>€ 69 546</td>
</tr>
<tr>
<td>p</td>
<td>0.43</td>
</tr>
<tr>
<td>X</td>
<td>€ 79 112</td>
</tr>
<tr>
<td>Y</td>
<td>€ 14 387</td>
</tr>
</tbody>
</table>

Sensitivity analyses
To address uncertainty around mean incremental costs and effectiveness, univariate sensitivity analyses were conducted. Whenever possible, we run the analysis using the upper and lower bound of the 95% CI of the mean. In order to assess how a simultaneous change of several variables affects the incremental cost-effectiveness ratio (ICER), we performed a Monte Carlo simulation, a type of multivariate sensitivity analysis. This technique runs a large number of simulations (here: 1 000) by repeatedly drawing samples from probability distributions of input variables. Thus, it provides a probability distribution for the output variables, i.e., incremental costs and effectiveness. Probabilities and relative risks were assumed to follow a beta distribution Beta(α, β) because they are restricted to take on values between 0 and 1. Because the distribution of health states at the time of diagnosis had more than 2 outcomes, we assumed a Dirichlet distribution Dirichlet (α1, α2, . . . , αk).37 Cost data were assumed to follow a gamma distribution Gamma(a, b) because they are normally distributed but restricted to take on values between 0 and 1. The standard deviation of probabilities and relative risks was calculated according to the following formula56:
Given the ambiguous interpretation of negative ICERs, we transformed ICERs into net monetary benefits (NMBs). We generated a cost-effectiveness acceptability curve based on the distribution of NMBs for each value of the willingness to pay per QALY gained. A cost-effectiveness acceptability curve allows a decision maker to consider whether a prevention strategy is cost-effective in relation to the maximum amount a decision-maker is willing to pay for a QALY. At each ceiling value for the willingness to pay for a QALY, the cost-effectiveness acceptability curve shows the probability that treatment is cost-effective. The input data for the model are summarized in Table 6.1.

Results

Base-case analysis

The base-case analysis, which applies to 50-year-old patients, shows that “no screening and treatment”, “screening for macroalbuminuria”, and “screening for micro-albuminuria” are all dominated by the “treat all” strategy, which is associated with the lowest costs and highest benefit (Table 6.3). Again, the “treat all” strategy implies that all patients are treated with an ACE inhibitor (or an ARB in the event of cough).

Table 6.3 Results of the base-case analysis, based on mean estimates of input variables.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Costs (€)</th>
<th>Undiscounted LYs</th>
<th>Discounted QALYs</th>
<th>ICER (€/QALY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening for macroalbuminuria</td>
<td>110 777</td>
<td>28.52</td>
<td>19.15</td>
<td>Dominated</td>
</tr>
<tr>
<td>Screening for microalbuminuria</td>
<td>101 140</td>
<td>28.88</td>
<td>19.54</td>
<td>dominated</td>
</tr>
<tr>
<td>Treating all patients with ACEIs/ARBs</td>
<td>98 421</td>
<td>28.94</td>
<td>19.63</td>
<td>dominant</td>
</tr>
</tbody>
</table>

ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin II receptor blocker; LYs = life years; QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio

Sensitivity analysis

In the univariate sensitivity analysis, variables with the largest impact on incremental costs and effectiveness are the absolute risk for progression from
micro- to macroalbuminuria without ACE inhibition as well as the relative risk for progression from normo- to microalbuminuria with ACE inhibitor therapy and the discount rate (see Table 6.4 for details). When assuming a low progression rate from micro- to macroalbuminuria without ACE inhibition, screening for microalbuminuria dominates the “treat all” strategy. A threshold sensitivity analysis shows that at an annual drug cost of €426.70 (base case: €62.70) the breakeven point is reached. The probability of savings is 70%.

Table 6.4  Univariate sensitivity analyses: effects of varying base-case estimates on the incremental cost-effectiveness ratio of treating all patients with ACE inhibitors vs screening for microalbuminuria (reference strategy). “Lower bound” and “higher bound” refer to the limits of the 95% confidence interval.

<table>
<thead>
<tr>
<th></th>
<th>Incremental costs</th>
<th>Incremental QALYs</th>
<th>Incremental cost-effectiveness ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial disease prevalence:</strong> Proportion of normoalbuminuric patients, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,289</td>
<td>0.080</td>
<td>-28,647</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,442</td>
<td>0.120</td>
<td>-28,647</td>
</tr>
<tr>
<td><strong>Annual transition probabilities (without ACE inhibitors)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from normo- to microalbuminuria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-1,712</td>
<td>0.062</td>
<td>-27,659</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,348</td>
<td>0.123</td>
<td>-27,214</td>
</tr>
<tr>
<td>from micro- to macroalbuminuria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>1,238</td>
<td>-0.22</td>
<td>-57,155</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-4,604</td>
<td>0.166</td>
<td>-27,736</td>
</tr>
<tr>
<td>from macroalbuminuria to ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-1,202</td>
<td>0.047</td>
<td>-25,823</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,625</td>
<td>0.126</td>
<td>-28,661</td>
</tr>
<tr>
<td><strong>Relative risk for progression with ACE inhibitors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from normo- to microalbuminuria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-4,352</td>
<td>0.141</td>
<td>-30,831</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-734</td>
<td>0.036</td>
<td>-20,510</td>
</tr>
<tr>
<td>from micro- to macroalbuminuria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-1,836</td>
<td>0.066</td>
<td>-27,921</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,730</td>
<td>0.131</td>
<td>-28,403</td>
</tr>
<tr>
<td>from macroalbuminuria to ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,274</td>
<td>0.080</td>
<td>-28,358</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,229</td>
<td>0.112</td>
<td>-28,727</td>
</tr>
<tr>
<td><strong>Utilities (health states)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Diabetes (baseline health)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,719</td>
<td>0.090</td>
<td>-30,264</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,719</td>
<td>0.100</td>
<td>-27,194</td>
</tr>
</tbody>
</table>

127
Table 6.4 (continued)

<table>
<thead>
<tr>
<th></th>
<th>Incremental costs</th>
<th>Incremental QALYs</th>
<th>Incremental cost-effectiveness ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ESRD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,719</td>
<td>0.142</td>
<td>-19,081</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,719</td>
<td>0.049</td>
<td>-55,041</td>
</tr>
<tr>
<td><strong>Disutility of ACE inhibitor treatment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,719</td>
<td>0.092</td>
<td>-29,554</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,719</td>
<td>0.095</td>
<td>-28,647</td>
</tr>
<tr>
<td><strong>Costs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACE inhibitor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,719</td>
<td>0.095</td>
<td>-28,647</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,569</td>
<td>0.095</td>
<td>-27,062</td>
</tr>
<tr>
<td><strong>ESRD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-1,858</td>
<td>0.095</td>
<td>-19,581</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-3,579</td>
<td>0.095</td>
<td>-37,713</td>
</tr>
<tr>
<td><strong>SMR</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,723</td>
<td>0.096</td>
<td>-28,249</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,715</td>
<td>0.093</td>
<td>-29,046</td>
</tr>
<tr>
<td><strong>Rate of ARB use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>-2,419</td>
<td>0.095</td>
<td>-25,463</td>
</tr>
<tr>
<td>Higher bound</td>
<td>-2,854</td>
<td>0.095</td>
<td>-30,042</td>
</tr>
<tr>
<td><strong>Specificity of HPLC (microalbuminuria screening procedure)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81%</td>
<td>-1,853</td>
<td>0.039</td>
<td>-47,513</td>
</tr>
<tr>
<td><strong>Discount rate of costs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>-9,179</td>
<td>0.095</td>
<td>-96,710</td>
</tr>
<tr>
<td>1.5%</td>
<td>-5,708</td>
<td>0.095</td>
<td>-60,140</td>
</tr>
<tr>
<td>4%</td>
<td>-2,719</td>
<td>0.095</td>
<td>-28,647</td>
</tr>
<tr>
<td>7%</td>
<td>-1,189</td>
<td>0.095</td>
<td>-12,523</td>
</tr>
<tr>
<td>10%</td>
<td>-537</td>
<td>0.095</td>
<td>-5,655</td>
</tr>
<tr>
<td><strong>Discount rate of effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>-2,719</td>
<td>0.139</td>
<td>-19,592</td>
</tr>
<tr>
<td>1.5%</td>
<td>-2,719</td>
<td>0.095</td>
<td>-28,647</td>
</tr>
<tr>
<td>4%</td>
<td>-2,719</td>
<td>0.051</td>
<td>-52,850</td>
</tr>
<tr>
<td>7%</td>
<td>-2,719</td>
<td>0.026</td>
<td>-105,670</td>
</tr>
<tr>
<td>10%</td>
<td>-2,719</td>
<td>0.014</td>
<td>-200,909</td>
</tr>
</tbody>
</table>

QALYs = quality-adjusted life years; ACE = angiotensin-converting enzyme; ESRD = end-stage renal disease; SMR = standardized mortality ratio; HPLC = high performance liquid chromatography
Figure 6.2 shows the cost-effectiveness acceptability curve, which considers uncertainty in cost-effectiveness. The probability of savings of the “treat all” strategy compared to screening for microalbuminuria is 70% (see also Figure 6.3 for the scatterplot).

Figure 6.2  Cost-effectiveness acceptability curve.

Figure 6.3  Cost-effectiveness plane showing 1000 replications from a distribution of cost and quality-adjusted life year (QALY) differences (angiotensin converting enzyme inhibitor vs microalbuminuria screening).
Discussion

This modeling study shows that treating all patients with type 2 diabetes with ACE inhibitors (and more expensive ARBs in the event of cough) immediately after diagnosis is cost-effective and even reduces health care expenditures in the Dutch setting. The results were robust to a variety of different assumptions of uncertainty.

Although a significant number of newly diagnosed type 2 patients may receive blood pressure medications, there is no evidence to date that these patients are primarily prescribed an ACE inhibitor, which underlines the significance of this analysis. Still, our model is far from being perfect, but in modeling studies this is rarely the case due to constraints of resources, time, and information availability.

In the present study, savings by treating all diabetic patients with ACE inhibitors may even be underestimated for several reasons. First, we did not model that ACE inhibitors and ARBs reduce the risk for cardiovascular events, which would lead to additional savings. Second, we did not consider real-world compliance with ACE inhibitor therapy due to a lack of data. In the real world some patients discontinue ACE inhibitor therapy and thus do not incur any drug cost. On the other hand, the model considered trial-based compliance on the effect side, as the rate of compliance is implicitly incorporated in clinical trial results, i.e. efficacy data refer both to adherers and non-adherers. For this reason the Markov model includes patients who discontinue ACE inhibitor treatment in the ACE inhibitor arm.

Third, the screening costs considered for microalbuminuria screening are based on one annual test only. In contrast, considering recommended screening procedures from the PREVEND IT study as a basis would lead to a fundamental increase of screening costs as a spot-urine sample (either the first-morning void or at the time of the visit to the medical office) was used as a pre-screening. Patients whose urine is tested positive should have their 24-h urine samples tested repeatedly afterwards.

Forth, as this study is based on a cohort simulation it uses data on the population mean. In contrast, a patient-level simulation would account for the fact that some individuals may stay in more than 2 stages in a year, although this is rarely the case. In any case, if patients progressed more rapidly (had
higher risk), then ACE inhibitor treatment could lead to an even larger absolute risk reduction and therefore larger savings.

Finally, costs of dialysis treatment will likely continue to rise in the future, thus increasing the potential for savings by preventing ESRD. Dialysis costs have increased within the last years\textsuperscript{12} and we expect this trend to continue due to stricter regulations concerning dialysis safety, technological advancement of dialysis machines, and better-tolerated dialysis solutions. Further limitations of the model relate to the data sources.

First, the model uses some epidemiological data from Western countries other than the Netherlands. For example, we used a Finnish study\textsuperscript{28} as the source of the distribution of health states at the time of diagnosis. However, changing the initial distribution of health states had little impact on the outcome.

Second, transition rates from macroalbuminuria to ESRD with and without ACE inhibitors were not available for patients with type 2 diabetes. Therefore, we used a randomized controlled trial in patients with type 1 diabetes as the source.\textsuperscript{15}

Third, the standardized mortality ratio (SMR) we applied to diabetic patients without ESRD\textsuperscript{1} includes patients with ESRD. Excluding these patients would lower the SMR to a minor degree as less than 2% of the Dutch diabetic population receives renal replacement therapy.\textsuperscript{35}

Fourth, we assumed that the SMR is the same for patients with normo-, micro-, and macroalbuminuria as there are no valid data showing that a significant difference exists. The slightly higher mortality ratio in microalbuminuric patients in the HOPE study (2000)\textsuperscript{57} was most likely the result of prior cardiovascular events. There is no evidence in the literature that mortality rates increase only on the basis of the level of albumin in the urine. This is the same with the utilities, which are assumed to do not differ between different stages of albuminuria.

Fifth, having microalbuminuria or macroalbuminuria might cause disutility due to anxiety. However, standard preference measures such as the SG or the TTO method are not able to capture anxiety over future events as both evaluation methods assume a constant health state over the remaining period of life.

Finally, it is unclear whether a societal perspective leads to larger or smaller savings than a health insurance perspective: on the one hand, ACE inhibitors
and ARBs avoid productivity loss and copayments due to renal failure. On the other hand, drug copayments lead to additional costs. Compared to previous cost-effectiveness models, which were conducted by Golan et al. (1999)\textsuperscript{22} and Rosen et al. (2005)\textsuperscript{23} based on U.S. data, a much broader evidence base for the transition between normo- to microalbuminuria and micro- to macroalbuminuria was included in the present study. In addition, we considered that patients who are noncompliant with ACE inhibitors due to cough may receive more expensive ARBs, as similarly done for the German setting.\textsuperscript{24} The fact that a small proportion of patients on ARBs (3.2\%) also develop cough\textsuperscript{17} and thus may discontinue treatment was disregarded. The reason for the exclusion is that noncompliance with treatment is already incorporated in the relative risk of treatment (thus lowering the relative risk), as in RCTs a certain proportion of patients discontinued treatment. In contrast to the previous models mentioned above we additionally conducted the analysis including an ARB for the entire patient population in need of treatment. This was done as some studies question that ARBs are not only a more expensive, but also a more effective alternative compared to ACE inhibitors. As the breakeven point is higher than the annual treatment costs of the ARB therapy this strategy must be considered cost-effective. However, we assumed equal effectiveness of all ACE inhibitors and ARBs, as meta-analyses do not suggest any independent effect of single renin-angiotensin-system agents.\textsuperscript{30,61} For instance, an ARB as an equivalent but more expensive alternative should only be prescribed in case of a contraindication (e.g. dry cough associated with ACE inhibitor treatment).

Still, similar to Adarkwah et al. (2010)\textsuperscript{24} our model shows that treating all newly diagnosed type 2 patients with ACE inhibitors saves costs. The probability of savings is higher in Germany than in the Netherlands (89\% vs. 70\%). Reasons for this difference are not obvious as, e.g. costs of screening, ACE inhibitor treatment, and ESRD are quite similar. In contrast to Rosen et al. (2005),\textsuperscript{23} we did not consider the preventive effect of ACE inhibitors on cardiovascular outcomes, which would have increased savings. An important reason for the large savings potential in the Netherlands is the low price of enalapril, which has substantially decreased during the last few years.\textsuperscript{46,62} The most ARBs are still protected by patent. Irbesartan, which we included in our study, is protected until March 2012. Assuming that prices of ARBs will decline
after expiration of the patent protection would further strengthen our conclusion.

Conclusion and policy considerations

For patients with diabetic kidney disease treatment with an ACE inhibitor is highly cost-effective. Current national guidelines which do not even consistently recommend an ACE inhibitor for patients with microalbuminuria need to be reconsidered. Still, it is unclear whether a societal perspective leads to smaller or larger savings than a health care perspective. For instance, ACE inhibitor treatment avoids productivity loss due to renal failure and copayments for the treatment of renal failure. But drug copayments lead to additional costs.
References


46. Farmacotherapeutisch kompas 2010, Medisch Farmaceutische Voorlichting, Uitgave van de Commissie Farmaceutische Hulp (CFH) van het College voor zorgverzekering (CVZ) 2010.


Chapter 7

To treat or not to treat? Cost effectiveness of ACE inhibitors in non-diabetic advanced renal disease – a Dutch perspective

Adarkwah CC
Gandjour A
Akkerman M
Evers S

Abstract

Background
Treating non-diabetic proteinuric patients with advanced renal disease with an angiotensin-converting enzyme (ACE) inhibitor is still subject to discussion. This study aims to determine the cost-effectiveness of ACE inhibitor therapy in this patient population in the Netherlands.

Methods
We compared two strategies: first, treating patients with advanced renal disease with an ACE inhibitor and no-treatment. A lifetime Markov decision model was developed simulating the progression of renal disease and using published data on costs and health outcomes. A health care perspective was adopted.

Results
In the base-case analysis, treatment with ACE inhibitors leads to higher benefits and lower costs and dominates the no-treatment strategy. Sensitivity analysis shows that the probability of savings is 83%.

Conclusion
ACE inhibitor treatment for non-diabetic patients with advanced renal disease in the Netherlands is highly cost-effective and should therefore be considered.
Introduction

End stage renal disease (ESRD) can be defined by the need for renal transplantation or long-term dialysis. It is a final outcome of advanced renal disease, which is defined as a serum creatinine ≥3.0 mg per deciliter. An increasing prevalence of patients in renal replacement therapy in the Netherlands is notable within the last 15 years as it almost doubled within this period. Currently, more than 15,000 patients require renal-replacement therapy (RRT) and it is remarkable that the proportion of transplanted patients has been continuously increasing and now represents almost 60% of all patients requiring RRT.

ESRD costs represent a significant burden for the health care system, with a proportion of the national expenditures in European countries ranging from 0.7% in the UK to 1.8% in Belgium, and with a proportion in the Netherlands of about 1.3%. To state that prevention of ESRD is important from a medical perspective, it becomes clear that this also accounts for an economic viewpoint as the costs of ESRD treatment amount to €42,000 per patient per year in the Netherlands.

Angiotensin-converting enzyme (ACE) inhibitors decelerate the progression from mild-to-moderate chronic kidney disease (CKD, with a serum creatinine level of 1.5 to 3.0 mg per deciliter) to ESRD in non-diabetic patients with hypertension. ACE inhibitors have also shown to be cost-effective in this patient population. In the U. S., e.g., benazepril treatment improves health and lowers costs as modelled over a 7-year period. The cost-effectiveness of ACE inhibitors in mild-to-moderate renal disease has also been shown for the Dutch setting.

In patients with advanced renal disease (serum creatinine levels of 3.0 mg per deciliter or more) physicians might be reluctant to use ACE inhibitors because of concern that potassium levels or serum creatinine will rise. Several side effects and risks of ACE inhibitor therapy are well described in patients with chronic renal disease. Those risks include cough, hyperkalemia, mild reduction in glomerular filtration rate in patients with parenchymal renal disease, and acute renal failure in patients with bilateral renal artery stenosis.
or volume depletion.\textsuperscript{15-18} A placebo-controlled randomized controlled trial (RCT) by Hou et al.\textsuperscript{1} showed that ACE inhibitors can be used safely in this patient population: in the 8-week run-in period only 4 out of 281 patients with a serum creatinine above 3 mg/ml had to be excluded because of hyperkalemia; in the 224 patients who underwent randomization the incidence of hyperkalemia was similar among patients who received an ACE inhibitor and those who received placebo.

The RCT by Hou et al.\textsuperscript{1} also showed that ACE inhibitor therapy has a renoprotective effect in this patient population, thus confirming the result of a small prior RCT.\textsuperscript{19} Hou et al. demonstrated that the renal protective effect of ACE inhibitors in patients with advanced chronic kidney disease is not only due to their antihypertensive action as the reduction in the rate of progression changed little after correction for blood pressure.\textsuperscript{1} Although more than 90% of patients in this study population suffered from hypertension (baseline blood pressure of \textgreater 150/85 mmHg), these patients did not suffer from hypertensive kidney disease as the main cause for their renal failure. In about 60% glomerular disease was found to be the main cause.\textsuperscript{1}

A cost-effectiveness model conducted in Germany by Adarkwah et al.\textsuperscript{20} suggests that ACE inhibitor treatment in this patient population is cost-effective, at least in the German setting. Currently, no data are available on the cost-effectiveness of ACE inhibitor therapy in non-diabetic patients with advanced nephropathy in the Dutch setting. However, results of the German study may not be transferable to the Netherlands. Different factors such as epidemiology of the disease, demography, the availability of health care resources and differences in reimbursement systems between countries hinder transferability of economic evaluation studies between countries, in particularly due to variances in absolute and relative costs/prices. Other cost effectiveness models on renal disease even in the Dutch setting did include patients with less severe stages or renal disease, i.e. a serum creatinine level of 1.5 to 3.0 mg per decilitre.\textsuperscript{3,12} This also applies to other European cost-effectiveness models.\textsuperscript{21,22}

The goal of this paper was to determine the cost-effectiveness of ACE inhibitor therapy in non-diabetic patients with advanced renal disease (serum
creatinine above 3.0 mg/dl) and proteinuria in the Netherlands, based on a meta-analysis of RCTs by Ihle et al.\textsuperscript{19} and Hou et al.\textsuperscript{1}

Materials and methods

Overview and model design

A Markov decision model previously developed for the German setting\textsuperscript{20} was adapted and also proven suitable for the prevention of diabetic nephropathy\textsuperscript{23,24} in the Dutch and German setting. Based on the results of the RCTs by Ihle et al.\textsuperscript{19} and Hou et al.\textsuperscript{1}, the progression of a cohort of 1000 patients aged 44 years with advanced renal insufficiency (serum creatinine: >3.0 mg/dl, glomerular filtration rate (GFR): 15-26 ml/min/1.73 m\textsuperscript{2}), proteinuria, and hypertension (>150/85 mm Hg), but without severe heart failure (New York Heart Association III or IV) was simulated.

In order to simulate costs and effectiveness over lifetime, a Markov model was constructed. A Markov model is an iterative process where patients stay in one cycle (i.e., a defined health state) for a certain time (here: 1 year) and then make a transition to another cycle. Markov models are useful when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. A half-cycle correction was used to allow for transition events occurring midway through each 12-month cycle. A cycle length of one year was chosen because most data were given in units per year. The age of 100 was chosen as a cut-off as there are no mortality data available beyond this age. Regardless, more than 99% of patients in the simulation are dead at this age so that the model represents a lifetime horizon. The model was built in Microsoft Excel\textsuperscript{\textregistered} 2010.

Our model contains the following 3 health states (Figure 7.1), which represent the occurrence of events after model entry:

1. advanced renal disease (serum creatinine: >3.0 mg/dl, GFR: 15-26 ml/min/1.73 m\textsuperscript{2}, CKD stage 4)
2. ESRD and
3. death.
We performed a cost-utility analysis and health outcomes were measured in terms of quality-adjusted life years (QALYs). QALYs are the product of life years and a measure of health-related quality of life (preference weight or score) and allow comparing cost-effectiveness across diseases. Preference weights are anchored on a scale from 0 to 1, where 0 and 1 represent death and full health, respectively. Economic outcomes were set in relation to clinical outcomes by dividing the incremental (i.e., additional) costs of providing ACE inhibitors compared to no therapy by the incremental QALYs gained. The analysis is conducted from a health care perspective instead of choosing a societal perspective in order to increase comparability our previous model on this topic.

Doubling of the serum creatinine level (an outcome in the RCT by Hou) was not included as a separate Markov state, as this is a surrogate for and shortly followed by ESRD. After each cycle, a specified proportion of patients move from advanced renal insufficiency to ESRD. This transition is delayed by ACE inhibitors. Simultaneously, a specified proportion of patients die. During each cycle, patients accumulate utilities (measured by QALYs) and costs. Combining both states – dialysis and transplant – to ESRD is a common course of action when modelling renal disease. We pooled individual data for
transplantation and dialysis as well as prevalences according to the national register in order to assess valid costs for the stage ESRD.

Clinical strategies

Two treatment strategies were considered. In the “ACE inhibitor treatment” strategy, patients are treated with benazepril 10 mg twice a day. Patients in the control group receive no ACE inhibitor. In addition, both groups receive other antihypertensive agents (diuretics, alpha- or beta-blockers, calcium-channel antagonists, or some combination of these medications), but no other renin-angiotensin-system agents.

Transition probabilities

In order to identify placebo-controlled RCTs on the effect of ACE inhibitors on the progression of advanced renal insufficiency to ESRD we used a literature review until July 2001 by Terajima et al. as a starting point. This paper found one placebo-controlled RCT. In order to identify additional placebo-controlled RCTs published after July 2001 we searched in the PubMed database (date: September 05, 2012) using the following search strategy:

Renal Insufficiency, Chronic AND creatinine AND antihypertensives AND end-stage renal disease AND angiotensin-converting-enzyme NOT(type-2-diabetes), Limits: Randomized Controlled Trial, All Adult: 19+ years. We obtained 43 hits. Thereof, 41 papers were excluded because they did not compare ACE inhibitors to placebo, did not consider advanced renal insufficiency, or did not include ESRD as an outcome. The above mentioned REIN study including amongst other Dutch nephropathy patients was excluded because the overall patient population suffered from mild-to-moderate renal disease and therefore also did not meet our inclusion criteria. The only 2 studies left were by Ihle et al. and Hou et al. and therefore used as the source of our effectiveness data. Both studies were double-blinded (see Table 7.1 for baseline characteristics). To combine effectiveness data from both studies and estimate the common odds ratio, we used the Mantel-Haenszel test.

In order to determine the annual transition probability from advanced renal disease to ESRD without ACE inhibitor therapy, we first calculated a total probability for each of the two trials, by dividing the number of events (ESRD)
during the trial period in the control arm by the number of patients. Second, we determined the annual transition probability by assuming a constant annual hazard rate over the study time horizon. And third, we determined a weighted average rate, with the number of individuals included in each study as weights. In the study by Hou et al. the number of events was obtained by directly contacting the authors and does not include death.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Hou et al.</th>
<th>Ihle et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACE inhibitor (n=112)</td>
<td>Placebo (n=112)</td>
</tr>
<tr>
<td>Age – years</td>
<td>44</td>
<td>45</td>
</tr>
<tr>
<td>Male sex – no. (%)</td>
<td>56 (50)</td>
<td>55 (49)</td>
</tr>
<tr>
<td>Blood pressure – mmHg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>153</td>
<td>152</td>
</tr>
<tr>
<td>Diastolic</td>
<td>87</td>
<td>85</td>
</tr>
<tr>
<td>Renal function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum creatinine – mg/dl</td>
<td>4.0</td>
<td>3.9</td>
</tr>
<tr>
<td>GFR – ml/min/1.73m²</td>
<td>26.3</td>
<td>25.8</td>
</tr>
<tr>
<td>Urine protein excretion – g/day</td>
<td>1.6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Transition probabilities with and without an ACE inhibitor therapy are shown in Table 7.2. In patients without ESRD mortality was regarded as a function of age. Specifically, we used age-specific mortality rates of non-diabetic patients. To calculate the latter, age-specific mortality rates of the general population were multiplied with a standardized mortality ratio for patients with advanced renal disease compared to the general population, which was derived from a large community-based cohort study. For patients with ESRD we assumed that mortality was age-independent.

Preference weights

For patients in the health state “advanced renal disease” we considered a utility loss according to a survey using the time tradeoff (TTO) method in 65 patients. This population had a mean age of 66 years and a GFR between 15 and 30 ml/min/1.73 m². Furthermore, we considered an age-dependent loss of utility. For patients treated with an ACE inhibitor a disutility of 0.05 from all side effects was considered in a sensitivity analysis. The preference
weight for ESRD was derived from a systematic review of empirical studies in which TTO weights were provided by patients.\textsuperscript{35} The TTO can be regarded as the most commonly used method to derive quality-of-life weights for QALYs.\textsuperscript{36,37}

Table 7.2 Data used to determine the cost-effectiveness of angiotensin-converting enzyme inhibitor treatment.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-case estimate</th>
<th>Range tested in sensitivity analysis*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual transition probabilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced renal insufficiency to ESRD without ACE inhibitor</td>
<td>0.14</td>
<td>0.08 – 0.19</td>
<td>1,19</td>
</tr>
<tr>
<td>Advanced renal insufficiency to death</td>
<td>Age- and gender dependent</td>
<td>0.003 – 0.54</td>
<td>61</td>
</tr>
<tr>
<td>ESRD to death</td>
<td>0.09</td>
<td>---</td>
<td>2</td>
</tr>
<tr>
<td>Effectiveness of ACE inhibitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative risk of advanced renal insufficiency to ESRD</td>
<td>0.61</td>
<td>0.39 – 0.95</td>
<td>1,19</td>
</tr>
<tr>
<td>Utilities (health states)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General population (age-dependent)</td>
<td>0.83 – 0.94</td>
<td>---</td>
<td>33</td>
</tr>
<tr>
<td>Disutility of chronic kidney disease</td>
<td>0.056 – 0.156</td>
<td>---</td>
<td>32</td>
</tr>
<tr>
<td>Disutility of ACE inhibitor treatment</td>
<td>0.00 – 0.05</td>
<td>---</td>
<td>34</td>
</tr>
<tr>
<td>ESRD</td>
<td>0.62</td>
<td>0.39 – 0.84</td>
<td>35</td>
</tr>
<tr>
<td>Annual costs, €</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General health care expenditures by age group</td>
<td>3.140,58 – 23.456,58</td>
<td>---</td>
<td>39,42,47</td>
</tr>
<tr>
<td>Per-patient cost of chronic kidney disease compared to non-CKD population</td>
<td>Age-dependent</td>
<td>---</td>
<td>36,39-42,47</td>
</tr>
<tr>
<td>ACE inhibitor (20 mg benazepril daily; 5 mg enalapril daily in the SA)</td>
<td>240,92</td>
<td>28,70 – 240,92</td>
<td>43</td>
</tr>
<tr>
<td>ESRD</td>
<td>42 110</td>
<td>33 688 – 50</td>
<td>26</td>
</tr>
<tr>
<td>Transplantation</td>
<td>14 387</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dialysis</td>
<td>79 122</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home/in-center hemodialysis</td>
<td>83 217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPD</td>
<td>54 067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCPD</td>
<td>69 546</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMR of chronic kidney disease</td>
<td>3.42</td>
<td>2.84 – 4.11</td>
<td>31</td>
</tr>
<tr>
<td>Discount rate of costs</td>
<td>0.04</td>
<td>0.00 – 0.10</td>
<td>42,47,48</td>
</tr>
<tr>
<td>Discount rate of benefits</td>
<td>0.015</td>
<td>0.00 – 0.10</td>
<td>42,47,48</td>
</tr>
</tbody>
</table>

ESRD = end-stage renal disease; ACE = angiotensin-converting enzyme; CKD = chronic kidney disease; SA = sensitivity analysis; CAPD = Continuous ambulatory peritoneal dialysis; CCPD = Continuous cycling peritoneal dialysis; SMR = standardized mortality ratio. * Based on 95%
confidence intervals of the mean except for ESRD (estimate)
The TTO technique determines the proportion of remaining life years in poor health which patients are willing to give up or trade in exchange for perfect health. Based on patient responses utility scores are calculated. Those utility measures in economic evaluations are becoming more and more important given the fact that decision makers are asked to optimize the allocation of scarce health care resources across disease areas and patient groups.\textsuperscript{38} Notably, the values are similar to EQ-5D scores (baseline value 0.61) reported by de Wit et al.\textsuperscript{6} We tested a rather broad range of TTO weights in sensitivity analysis.

Costs
The analysis was conducted from the health care perspective. Therefore, only direct health effects – defined as life years gained – and direct costs and were considered. Making use of data on the consumer price index, costs were inflated to year 2010 euros.\textsuperscript{39} Costs of health care expenditures unrelated and related to CKD as well as costs for ACE inhibitors and treatment of ESRD were accounted for. To determine the health care expenditures related and unrelated to CKD, we took resource consumption into account, based on international guidelines for the treatment of CKD.\textsuperscript{40,41} The recommendations of the Dutch guidelines for pharmacoeconomic research were followed.\textsuperscript{42}

In the base case, the cheapest generic of benazepril 10 mg daily was used twice daily.\textsuperscript{43} We included costs of ACE inhibitor therapy for patients who had to be excluded in the 8-week run-in period as these patients would be treated in the real world. In a sensitivity analysis we applied the price of the cheapest generic of enalapril 5 mg once daily as used in the RCT by Ihle et al.\textsuperscript{19} This is, for instance, the most frequently prescribed ACE inhibitor in the Netherlands.\textsuperscript{44}

The costs of benazepril and enalapril were based on 2010 Dutch prices and include 6\% value-added tax as well as a 3-monthly pharmacists’ prescription fee of €7.50.\textsuperscript{45}

The annual costs of patients with ESRD were calculated as a weighted average of the costs of renal transplantation as well as different types of dialysis as
based on a Dutch study and prevalence data available from the national register.

In detail, the following calculations were made (see Table 7.3 for details):

1) cost of dialysis = $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = X$, where $x_n =$ annual cost of dialysis treatment $n = 1, 2, 3$; $\beta_n =$ prevalence weight of the dialysis treatment, and $\beta_1 + \beta_2 + \beta_3 = 1$


We assumed a transplant survival of 10 years and a distinction was made between the first year of transplantation and the years following. Costs were inflated to 2010 Dutch prices.

As our study population is mainly hypertensive (91.5%) and no valid data on CKD treatment costs per se are available, we included the costs for hypertension treatment as an approximation. Costs were discounted at an annual rate of 4% whereas benefits were discounted at an annual rate of 1.5% in accordance with the CVZ recommendations.

Table 7.3 Parameters used for calculating the cost of end-stage renal disease (see cost section under "Methods").

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>home/center hemodialysis</td>
</tr>
<tr>
<td>2</td>
<td>continuous ambulatory peritoneal dialysis (CAPD)</td>
</tr>
<tr>
<td>3</td>
<td>continuous cycling peritoneal dialysis (CCPD)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.82</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.106</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.074</td>
</tr>
<tr>
<td>$x_1$</td>
<td>€ 83 217</td>
</tr>
<tr>
<td>$x_2$</td>
<td>€ 54 067</td>
</tr>
<tr>
<td>$x_3$</td>
<td>€ 69 546</td>
</tr>
<tr>
<td>$p$</td>
<td>0.43</td>
</tr>
<tr>
<td>$X$</td>
<td>€ 79 112</td>
</tr>
<tr>
<td>$Y$</td>
<td>€ 14 387</td>
</tr>
</tbody>
</table>
Sensitivity analysis

In order to address uncertainty around mean incremental costs and effectiveness, we conducted univariate sensitivity analyses, where we varied one variable at a time while keeping all other variables constant at their mean or base-case value. We ran analyses using the upper and lower bound of the 95% confidence intervals of the mean whenever possible. For transplantation costs we did not have information on the 95% CI and thus varied costs by 20% (see Table 7.2 for uncertainty ranges).

To assess how a simultaneous change of several variables affects the incremental cost-effectiveness ratio (ICER), we performed a Monte Carlo simulation, a type of multivariate sensitivity analysis. This method runs a large number of simulations (here: 1000) by repeatedly drawing samples from probability distributions of input variables. Therefore, it provides a probability distribution for the output variables (incremental costs, incremental effectiveness, and ICERs). Cost data were assumed to follow a gamma distribution Gamma(a, b), which reflects the long right tail and restriction to positive values. Probabilities were assumed to follow a beta distribution Beta(α, β) because they are normally distributed but restricted to take on values between 0 and 1. The standard error of a probability or proportion was calculated according to the following formula:

$$
\sigma = \sqrt{\frac{p(1-p)}{n}}
$$

Where \( p \) = probability and \( n \) = sample size. Relative risks and odds ratios were assumed to follow a lognormal distribution.

Based on the Monte-Carlo simulation we generated a cost-effectiveness acceptability curve (CEAC). A CEAC allows a decision-maker to consider whether a prevention strategy is cost-effective in relation to the maximum amount a decision-maker is willing to pay for a QALY. At each ceiling value for the willingness to pay for a QALY, the cost-effectiveness acceptability curve shows the probability that the treatment is cost-effective.
Results

Base case

The base-case analysis, which had a lifetime horizon, shows domination of the ACE inhibitor treatment strategy over the no treatment strategy (Table 7.4). Treatment was therefore associated with lower costs and larger health benefit. Even shorter-term simulations over 3 and 10 years revealed the same finding, i.e., lower costs and larger benefits.

Table 7.4 Results of the base-case analysis, based on mean estimates of input variables.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Costs (€)</th>
<th>Undiscounted LYs</th>
<th>Discounted QALYs</th>
<th>ICER (€/QALY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ACE inhibitor treatment</td>
<td>220,942</td>
<td>13.38</td>
<td>9.32</td>
<td></td>
</tr>
<tr>
<td>ACE inhibitor treatment</td>
<td>183,535</td>
<td>14.66</td>
<td>11.11</td>
<td>Dominates</td>
</tr>
</tbody>
</table>

ACE = angiotensin-converting enzyme; LYs = life-years; QALY = quality-adjusted life-years; ICER = incremental cost-effectiveness ratio

Sensitivity analysis

Variables with the largest impact on incremental costs and effectiveness in the univariate sensitivity analysis are the effectiveness of ACE inhibitor treatment, the costs of ESRD, and the discount rates of costs and effects (Table 7.5). Figure 7.2 shows the cost-effectiveness acceptability curve, considering uncertainty in cost-effectiveness. Replications with lower effectiveness of the “ACE inhibitor treatment” strategy compared to the “no treatment” strategy (1.6%) were categorized as not cost-effective (regardless of concurrent savings). This is because it would be unethical to accept treatment that is less effective than no treatment. Hence, the maximum probability of cost-effectiveness independent of the willingness to pay is 98.4%. The probability of savings is 83%. A scatter plot is presented in Figure 7.3.
Table 7.5  Univariate sensitivity analyses from a health care perspective: effects of varying base-case estimates on the incremental cost-effectiveness ratio of treating all patients with advanced renal insufficiency with an ACE inhibitor vs no ACE inhibitor treatment (reference strategy). “Lower bound” and “higher bound” refer to the limits of the 95% confidence interval or estimate thereof.

<table>
<thead>
<tr>
<th></th>
<th>Incremental costs</th>
<th>Incremental QALYs</th>
<th>Incremental cost-effectiveness ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case</td>
<td>- €37 407</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>Annual transition probabilities from advanced renal insufficiency to ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without ACE inhibitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €36 685</td>
<td>1.76</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €32 842</td>
<td>1.62</td>
<td>Dominance</td>
</tr>
<tr>
<td>Effectiveness of ACE inhibitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative risk of advanced renal insufficiency to ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €72 134</td>
<td>3.39</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €2 673</td>
<td>0.17</td>
<td>Dominance</td>
</tr>
<tr>
<td>Utilities (health states)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €37 407</td>
<td>1.63</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €37 407</td>
<td>1.86</td>
<td>Dominance</td>
</tr>
<tr>
<td>ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €37 407</td>
<td>2.01</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €37 407</td>
<td>1.58</td>
<td>Dominance</td>
</tr>
<tr>
<td>Disutility of ACE inhibitor treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €37 407</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €37 407</td>
<td>1.57</td>
<td>Dominance</td>
</tr>
<tr>
<td>Costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE inhibitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €36 680</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €37 407</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>ESRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €28 400</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €46 415</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>SMR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>- €37 307</td>
<td>1.88</td>
<td>Dominance</td>
</tr>
<tr>
<td>Higher bound</td>
<td>- €37 521</td>
<td>1.69</td>
<td>Dominance</td>
</tr>
<tr>
<td>Discount rate of costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>- €49 072</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>5%</td>
<td>- €35 105</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>7%</td>
<td>- €31 066</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>10%</td>
<td>- €26 158</td>
<td>1.79</td>
<td>Dominance</td>
</tr>
<tr>
<td>Discount rate of effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>- €37 407</td>
<td>2.31</td>
<td>Dominance</td>
</tr>
<tr>
<td>5%</td>
<td>- €37 407</td>
<td>2.07</td>
<td>Dominance</td>
</tr>
<tr>
<td>7%</td>
<td>- €37 407</td>
<td>0.84</td>
<td>Dominance</td>
</tr>
<tr>
<td>10%</td>
<td>- €37 407</td>
<td>0.60</td>
<td>Dominance</td>
</tr>
</tbody>
</table>

QALYs = quality-adjusted life years; ESRD = end-stage renal disease; SMR = standardized mortality ratio.
Figure 7.2 Cost-effectiveness acceptability curve.

Figure 7.3 Cost-effectiveness plane showing 1000 replications from a distribution of cost and quality-adjusted life year (QALY) differences (angiotensin-converting enzyme inhibitor vs no treatment).
Discussion

The study shows that treating non-diabetic patients with advanced renal insufficiency in the Netherlands with ACE inhibitors saves both money and lives. Despite the considerable costs of dialysis treatment, preventing ESRD does not necessarily lead to savings,\textsuperscript{50} thus necessitating a detailed calculation of costs and benefits. This result is consistent with a previous cost-effectiveness analysis conducted in the German setting.\textsuperscript{20} Treating patients with more advanced renal disease (creatinine $\geq 3.0$ mg per deciliter) with an ACE inhibitor to delay disease progression is still subject to discussion and no clear guideline statement regarding this aspect is available to date.\textsuperscript{51-53} Therefore results could be helpful to adjust current guideline recommendations.

Still, our model is subject to some limitations. In modeling studies this is usually the case due to constraints of resources, time, and information availability. The cost-effectiveness by treating non-diabetic progressed chronic kidney disease with ACE inhibitors may be even underestimated for several reasons. First, we did not model that ACE inhibitors reduce the risk for cardiovascular events,\textsuperscript{54,55} which leads to additional savings. As our study does not consider savings from avoiding cardiovascular events, it does, however, consider that patients with chronic kidney disease have a higher risk of to die because of cardiovascular reasons. We captured this by the mortality ratio for patients with chronic kidney disease compared to the general population. Second, effectiveness (and thus savings) of benazepril may have been underestimated as ESRD treatment may have been initiated earlier than necessary in the ACE inhibitor arms of the RCTs by Hou et al.\textsuperscript{1} and Ihle et al.\textsuperscript{19} due to a drug-induced GFR decrease. Third, we did not consider on the cost side that some patients may discontinue ACE inhibitor therapy and therefore incur no drug costs. The reason is that information on compliance was unavailable in the two RCTs.\textsuperscript{1,19} Treatment discontinuation was only considered on the effect side, as the rate of compliance is implicitly incorporated in clinical trial results, i.e., efficacy data refer both to adherers and non-adherers. Therefore the Markov model includes patients who discontinue ACE inhibitor treatment in the ACE inhibitor arm. And forth, for
patients who receive dialysis we were not able to consider health care costs unrelated to dialysis due to a lack of available data.

On the other hand, the savings might be smaller than estimated by our model. This is because some elderly patients may have a preference against replacement therapy and hence do not incur the associated costs. The bias seems small, however, given that the percentage of elderly patients requiring RRT is small (e.g., less than 3% at the age of 80 years and above) according to the register. Further limitations of the model relate to its data sources.

First, the two trials considered as the source of our effectiveness data used for ACE inhibition enalapril and benazepril, respectively. Whether other types of ACE inhibitors can achieve the results reported in these studies is not clear. ACE inhibitors have potentially important differences in plasma protein binding, their affinity for ACE, and pharmacokinetics.\textsuperscript{56} To date, there is no convincing evidence that generic long-acting ACE inhibitors (such as lisinopril and enalapril) are inferior to brand-name ACE inhibitors.\textsuperscript{56,57}

Second, our preference weights for patients with advanced renal disease were taken from a U. S. study\textsuperscript{32} as no similar data for Dutch patients were available. Third, the model also uses some epidemiological data from Western countries other than the Netherlands. One of the two studies we used as the source of our effectiveness data\textsuperscript{1} was conducted in China. There is a debate on whether a significant difference between white and Asian populations exists due to ACE gene polymorphism, which might lead to a difference in the efficacy of ACE inhibitors. In fact, studies examining the association between differences in ACE gene polymorphism and the efficacy of ACE inhibitors have had contradictory results.\textsuperscript{58} In a controlled trial performed among Japanese patients with non-diabetic chronic kidney disease whose blood pressure was maintained at an average of 126/72 mm Hg with ACE inhibitors, no significant correlation was observed between the type ACE gene polymorphism (DD, ID, or II) and the efficacy of the intervention or renal outcome.\textsuperscript{59} On the other hand, a controlled trial among 212 whites showed that ACE inhibitor therapy was protective against progression to ESRD only in patients with the DD genotype.\textsuperscript{60} But even if this small trial is confirmed, treating whites regardless of ACE genotype may not have a smaller impact than treating Asians as whites with the DD genotype may have a stronger response than the average
Asian patient. Furthermore, the study by Ihle et al., 19 which was conducted in Australia, shows a similar result as the study by Hou et al. 1

Forth, quality of life values were derived from non-Dutch populations due to a lack of availability. Hence, we tested a broad range of TTO values in sensitivity analysis.

And fifth, longitudinal studies that have examined long-term outcomes of ACE inhibitor treatment in patients with advanced chronic kidney disease are not available for model validation. This is a common limitation of models on the prevention of renal failure.

In addition, the costs of dialysis treatment will likely continue to rise in the future, and therefore increase the potential for savings by preventing ESRD. Dialysis costs have significantly increased within the last years. 5 We expect this trend to continue due to stricter regulations concerning dialysis safety, better-tolerated dialysis solutions and technological advancement of dialysis machines. On the other hand, prices of erythropoietin and reimbursement of physicians may drop.

Conclusion

For patients with non-diabetic proteinuric advanced renal disease, treatment with an ACE inhibitor is highly cost-effective and should therefore be considered. A societal perspective will likely lead to even larger savings than a health care perspective as ACE inhibitor treatment avoids productivity loss due to renal failure. Yet, the transferability of the conclusion of this study to other countries is limited given differences in costs, clinical management (eg, treatment of ESRD), and epidemiology.
References


43. Farmacotherapeutisch kompas 2010, Medisch Farmaceutische Voorlichting, Uitgave van de Commissie Farmaceutische Hulp (CFH) van het College voor zorgverzekering (CVZ).


Chapter 8

General discussion
Introduction

As introduced in Chapter 1, the aim of the dissertation is threefold: first, to assess whether costs for consumption and leisure activities need to be considered in the incremental cost-effectiveness ratio (ICER) in cost-effectiveness analyses (Part Ia – methodological aspects, utility measurement), second, to review evidence and highlight important bias types in trial-based as well as in model-based economic evaluations (Part Ib – methodological aspects, bias in economic evaluation) and third, to assess the cost-effectiveness of angiotensin-converting enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs) in diabetic as well as non-diabetic renal disease in different countries (Part II – economic evaluation studies).

In this chapter, we will first summarize the main findings of our research and then discuss the implications for clinical practice, research and policy.

Summary of main findings

Ia. Methodological work, utility measurement

Chapter 2 of the thesis presents the results of a randomized study on patients with inflammatory bowel disease in Germany (n=104). The purpose of the study was to conduct an empirical survey on whether the effects of ill health on consumptive activities are spontaneously considered in a health state valuation exercise and how much this matters. Patients were randomly provided or not provided with explicit instructions regarding the consideration of consumption and leisure effects in a time trade-off (TTO) exercise. The study showed that explicit instructions to consider non-health-related utility in TTO exercises did not influence TTO scores. However, spontaneous consideration of non-health-related utility in patients without explicit instruction (60% of respondents) led to significantly lower TTO scores.

Based on these findings, we recommend that consumption costs be included in the numerator of the ICER, at least for those respondents who spontaneously consider non-health-related utility from treatment. However, double-counting always has to be avoided.

To the best of our knowledge, for the effects of ill health on consumptive activities, no such empirical evidence has been gathered so far, although this
study shows that it plays an important role in the valuation of utilities. Results also showed that exercises eliciting health valuations from the general public may include a description of the impact of disease on consumptive activities. Overall, this research might contribute to a more accurate ICER by providing further insights into the TTO understanding of patients.

1b. Methodological work, bias in economic evaluation

Chapters 3 and 4 complete the methodological part of the thesis, both dealing with bias types and addressing the second aim of the thesis. Chapter 3 aims to give an overview of the potential risks of bias in trial-based economic evaluations and to identify how key sources for bias can be revealed and overcome. In total, eleven biases were identified and assigned to a particular trial phase. A distinction was made between pre-trial biases, biases during the trial and biases that are relevant after the actual trial. We discussed these bias forms in detail and presented strategies to detect and overcome them in a systematic way. The biases identified also account for model-based studies. Chapter 4 is directly related to Chapter 3. Here, we focused on bias types in model-based economic evaluation. It is obvious that bias types relevant for trial-based economic evaluation are usually relevant for model-based studies as well. So in this chapter, we focused on selected model-specific issues, using the Philips checklist as a tool to structure our results.1

Eleven specific biases for model-based economic evaluations were identified and classified related to structure, data and consistency of the model, including structural assumptions, model type, time horizon, data selection (such as treatment effects), assessment of uncertainty and internal validation. In the end, together with biases which account for trial-based as well as for model-based studies identified in Chapter 3, a checklist for assessing bias in economic evaluations was developed (ECOBIAS). The ECOBIAS checklist encompasses a common part for model- and trial-based economic evaluation and a model-specific part. Part A of the checklist incorporates eleven biases that account for both types of economic evaluation, whereas Part B covers eleven model-specific aspects of bias.

Both studies demonstrated that there are several ways that biases can influence the final results of trial-based as well as model-based economic evaluations, which could change a result from being cost-effective to not
being cost-effective at all. Combining the biases that can occur in trial-based with those in model-based economic evaluations identified in a previous article by the author group, the ECOBIAS checklist for identifying and avoiding bias in economic evaluation is introduced. To our current knowledge, we provide the first checklist (ECOBIAS checklist) on biases in trial-based and model-based economic evaluations. It is important to understand that different bias types can be related to each other and that the elimination of one can introduce another. The checklist can be used by researchers and can be completed with submission of an article, but it can also be used by policymakers to check for bias in a systematic way.

II. Economic evaluation studies
In Chapter 5, a cost-utility analysis on the cost effectiveness of ACE inhibitors in newly diagnosed type 2 diabetes was performed. The aim of this study was to assess the most cost-effective time to start an ACE inhibitor in the event of cough in patients with type 2 diabetes in Germany. Three strategies were compared: treating all patients at the time of diagnosing type 2 diabetes, screening for microalbuminuria and screening for macroalbuminuria. A lifetime Markov decision model with simulated 50-year-old patients with newly diagnosed diabetes mellitus was developed using published data on costs and health outcomes and simulating the progression of renal disease. We showed that the treat-all strategy is associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. Our results were robust, even when considering a variety of different assumptions of uncertainty. It was concluded that patients with type 2 diabetes should be treated with an ACE inhibitor immediately after diagnosis if they do not have contraindications. Although a significant number of newly diagnosed type 2 patients may receive blood pressure medications, it has not been shown that these patients are primarily prescribed an ACE inhibitor, which underlines the significance of this analysis. Treatment guidelines in Germany and the Netherlands do not strictly recommend prescribing an ACE inhibitor in this patient population, so our analysis should lead to a correction of current treatment guidelines in a way that newly diagnosed type 2 diabetic patients should receive a low dose ACE inhibitor, even in the absence of hypertension or heart failure, which are
currently the standard indications for this drug class. It is therefore clear that an ACE inhibitor should be considered a first-line treatment for hypertensive diabetic patients as well. As the topic investigated in Chapter 5 is of great importance not only from patient’s perspective but also from an economic viewpoint since diabetes is a huge burden for healthcare systems, at least in the Western world, Chapter 6 investigated this research question in the Dutch setting. The majority of data taken into consideration for the model in Chapter 5 were country-specific or at least adapted to the German setting. So the generalizability of the results is not given. We therefore built a new model for examining the cost-effectiveness of ACE inhibitors (or angiotensin II receptor blockers (ARBs) if coughing occurs as a side effect) in patients with newly diagnosed type 2 diabetes in the Netherlands. We showed a similar trend as in the German setting: In the base-case analysis, the treat-all strategy was associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. A multivariate sensitivity analysis shows that the probability of savings was 70%. It was concluded that also in the Netherlands, patients with type 2 diabetes should be prescribed an ACE inhibitor immediately after diagnosis if they do not have contraindications. An ARB should be considered for patients who develop a dry cough when treated with an ACE inhibitor. In addition, accounting for both models, the potential for cost savings would be even larger if the prevention of cardiovascular events due to the ACE inhibitor treatment were considered.

Although effectiveness was shown in different trials,4-7 cost-effectiveness models did not yet exist in this area for the Dutch and German settings. A systematic review on ACE inhibitors and ARBs in type 2 diabetic nephropathy was published recently.8 Here, 30 articles were evaluated, including Chapters 5 and 6 of this thesis. Our articles were the only two studies that discussed that dry cough, as a possible side effect of ACE inhibitors, could potentially have an impact on the cost-effectiveness of the latter.

With respect to diabetes, diabetic nephropathy, with its final state of end-stage renal disease (ESRD), is a sequela that causes an enormous economic burden for healthcare systems due to high costs for either dialysis or renal transplantation. Considering this burden, other causes for ESRD are worth investigating in more detail. Diabetes itself accounts for about 40% of patients in renal replacement therapy and therefore represents the main cause for
ESRD. Contrariwise, about 60% of patients with ESRD suffer from non-diabetic renal disease, which is focused on in Chapter 7. There is evidence that ACE inhibitors need to be handled with care in this patient population. Therefore, national as well as international clinical guidelines do not consistently recommend treating non-diabetic proteinuric patients with advanced renal disease with an angiotensin-converting enzyme (ACE) inhibitor. This study aimed to determine the cost-effectiveness of ACE inhibitor therapy in this patient population in the Netherlands by comparing two strategies: treating patients with advanced renal disease with an ACE inhibitor versus no treatment. A lifetime Markov decision model was developed to simulate the progression of renal disease using country-specific published data on costs and health outcomes. We found that treatment with ACE inhibitors leads to higher benefits and lower costs and dominates the no-treatment strategy. So it could be concluded that ACE inhibitor treatment for non-diabetic patients in the Netherlands, even those with advanced renal disease, is highly cost-effective and should therefore be considered. This should have a strong impact on current treatment guidelines. Our study was the first cost-effectiveness study to investigate this patient population with respect to ACE inhibitor treatment.

Assessing cost-effectiveness in diabetic as well as non-diabetic patients with renal disease represented the second part of the thesis, addressing the third objective as mentioned above. Even though most guidelines recommend using a social perspective, where all costs and outcomes are considered important regardless of who pays for the costs and who will experience the effects, our analyses were performed from a healthcare perspective, which was considered more suitable and was stated clearly as well as justified in detail. Here, only costs and outcomes relevant to the healthcare system were considered. Not stating the perspective correctly or stating a societal view but not including all costs and outcomes might induce bias, which was also systematically described and investigated in this thesis as mentioned above.

Implications for clinical practice, research and policy

This thesis has several implications for clinical practice, further research and policy. All implications will be discussed in the following section.
Ia. Methodological work, utility measurement

Looking at the consideration of consumption and leisure activities, we could show that explicit instructions to consider non-health-related utility in TTO exercises did not influence TTO scores. However, spontaneous consideration of non-health-related utility in the group that did not receive explicit instructions (60% of respondents of this group) led to significantly lower scores. Hence, the first finding could be explained by spontaneous consideration of non-health-related utility even without explicit instructions (thus offsetting the effect of explicit instructions). Spontaneous consideration of non-health-related utility thus leads to a utility gain from treatment compared to no spontaneous consideration. This would justify an inclusion of consumption costs in the numerator of the ICER, not only for life extending but also for quality-of-life enhancing interventions in future research, at least for those respondents who spontaneously consider non-health-related utility from treatment. Results also suggest that exercises eliciting health valuations from the general public for the purpose of conducting economic evaluations should be reconsidered. Researchers should include a description of the impact of disease on consumptive and leisure activities if those exercises are performed from a societal perspective.

The limitations of this piece of research directly lead to further implications for research. For instance, the study population represents a well-defined patient population: patients who suffered from an inflammatory bowel disease, i.e. either Crohn’s disease or ulcerative colitis. Therefore, other diseases or disease groups need to be investigated in future research. Next, the population sample included relatively young patients, and while we adjusted for age in our analysis, another sample with more elderly patients may provide a different result. So studies on elderly patient samples are needed to see if our results hold for this age group as well. Last, results may not be transferable from country to country as a matter of course. Hence, studies that are conducted outside Germany are also needed.

Given the importance of this question for the conduct of cost-effectiveness analysis in healthcare, policymakers should improve the research setting since confirmation is urgently needed in additional studies that consider other health-state valuation techniques. In a setting where resources are scarce, the
reliance on valid data, including valid incremental cost-effectiveness ratios, becomes increasingly important.

1b. Methodological work, bias in economic evaluation

Investigating bias in economic evaluation explicitly shows that there is a thin line between limited methodological quality and biases. A number of guidelines have been developed for economic evaluations. However, these guidelines are not enough to overcome all biases. Guidelines overall only tell you what to do but not how to do it in a sound and transparent way. Economic evaluators, like any other researchers, will have to rationalize the choices they have made during every phase of each study. In order to avoid biases in economic evaluations, researchers, readers, reviewers, editors and policymakers have to be aware of the possible biases and all stakeholders have to examine trial-based economic evaluations in a rigorous and stringent manner. Our research can therefore be helpful in this examination for researchers as well as for policymakers, as it provides an overview of the possible biases that should be taken into account in order to better perform an evaluation or to rate an evaluation more adequately. The ECOBIAS checklist can be seen as a complementary tool next to existing guidelines and checklists. The establishment of an ISPOR task force on bias or a consensus panel meeting could be the next steps forward. Although ECOBIAS already covers a large number of items, further research beyond the scope of our article is needed to establish all bias-related items relevant to economic evaluation studies.

We might learn from other fields where something comparable does exist, e.g. the Cochrane risk of bias instrument for systematic reviews of interventions which is used to assess the risk of bias in five domains. Furthermore, a tool for health behavior change trials was developed recently, i.e. the risk of bias justification table (RATIONALE). A tool like the latter, tailor-made for economic evaluations, could be made obligatory for submission of an article.

Our work on bias gives a good overview of the potential biases in economic evaluation, although further research is needed. We recommend that editors and reviewers consistently use a checklist, e.g. the ECOBIAS checklist presented in Table 8.1. To ensure more transparency, the completed checklist
could be made obligatory with submission and afterwards made available online to the community with supplementary material.

Table 8.1  The ECOBIAS checklist for bias in economic evaluation.

<table>
<thead>
<tr>
<th>Type of bias</th>
<th>Issues addressed (questions to consider)</th>
<th>Relevant to study</th>
<th>How did you deal with this bias? (description of strategy and rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PART A. Overall checklist for bias in economic evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrow perspective bias</td>
<td>Was a societal perspective adopted? If not, has a different perspective been justified?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inefficient comparator bias*</td>
<td>Was the best alternative chosen as a comparator? Was current practice chosen as a comparator? Have all comparators been described in sufficient detail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost measurement omission bias</td>
<td>Were all costs relevant to the disease and intervention identified and considered?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermittent data collection bias</td>
<td>Was the resource use measured continuously?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invalid valuation bias</td>
<td>Is the price calculation presented in a detailed manner? Have reference prices been used?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinal ICER bias</td>
<td>Have cardinal scales for the outcomes measure in a CEA been used?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-counting bias</td>
<td>Have the variables been adequately checked for double-counting?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inappropriate discounting bias</td>
<td>Have discounting rates from guidelines been applied?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited sensitivity analysis bias§</td>
<td>Have the four principles of uncertainty (methodological, structural, heterogeneity, parameter) been considered in sufficient detail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sponsor bias</td>
<td>Have sponsorships been disclosed? Is the study protocol freely accessible?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting and dissemination bias</td>
<td>Has the study / trial been listed in a trial register? Have all results been reported according to the study protocol?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8.1 (continued)

<table>
<thead>
<tr>
<th>Type of bias</th>
<th>Issues addressed (questions to consider)</th>
<th>Relevant to study</th>
<th>How did you deal with this bias? (description of strategy and rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PART B. Model-specific aspects of bias in economic evaluation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I Bias related to structure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural assumptions bias</td>
<td>Is the model structure in line with coherent theory? Do treatment pathways reflect the nature of disease?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment comparator bias*</td>
<td>Is there an adequate comparator, i.e. care as usual?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrong model bias</td>
<td>Is the model chosen adequate regarding the decision problem?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited time horizon bias</td>
<td>Was a lifetime horizon chosen? Were shorter time horizons adequately justified?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II Bias related to data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to data identification</td>
<td>Are the methods of data identification transparent? Are all choices justified adequately? Do the input parameters come from high-quality and well-designed studies?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to baseline data</td>
<td>Are probabilities, for example, based on natural history data? Is transformation of rates into transition probabilities done accurately?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to treatment effects</td>
<td>Are relative treatment effects synthesized using appropriate meta-analytic techniques? Are extrapolations documented and well justified? Are alternative assumptions explored regarding extrapolation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to quality-of-life weights (utilities)</td>
<td>Are the utilities incorporated appropriate for the specific decision problem?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-transparent data incorporation bias</td>
<td>Is the process of data incorporation transparent? Are all data and their sources described in detail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited scope bias§</td>
<td>Have the four principles of uncertainty (methodological, structural, heterogeneity, parameter) been considered?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III Bias related to consistency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias related to internal consistency</td>
<td>Has internal consistency in terms of its mathematical logic been evaluated?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* These biases are overlapping regarding their content.

§ These biases are overlapping regarding their content.
In regard to trustworthiness, transparency is also a big issue. Hence, especially when looking at models, policymakers are often insecure because there is a high deficit in understanding about what researchers, especially modelers, exactly do.\textsuperscript{29,30} To overcome this problem or at least to minimize it, more transparency is needed. Peters et al.\textsuperscript{31} highlighted the importance of introducing a full disclosure policy and showed how even author’s reservations can easily be diminished. Peters et al.\textsuperscript{31} formed some recommendations that challenged both authors and editors. Taking those recommendations into account might lead to significant benefits. Transferring this to model-based economic evaluation would involve providing open access to all data and material used in a particular model (variables, questionnaires, etc.) or even allowing access to the model itself. As the number of studies dealing with biases in economic evaluation is limited, additional research is needed to identify biases and to reveal or develop appropriate strategies for addressing the risk of bias.

II. Economic evaluation studies

With regard to the cost-utility analyses of ACE inhibitors and ARBs in renal disease, our findings show the need for a change in the current treatment guidelines.\textsuperscript{2,3,32-34} Consistent treatment with ACEI or ARB in the patient populations investigated was the dominant strategy, meaning not only the best strategy from patient perspective, but also from a healthcare system perspective, as it was associated with the lowest costs. Therefore, clinicians should reconsider guidelines by adopting the results and treating all newly diagnosed type 2 diabetes patients in the Netherlands and in Germany with an ACE inhibitor, independent of elevated blood pressure—the earlier, the better. Of course, the monitoring of renal function and potassium should be mandatory, but nevertheless, treatment should take place. Second, this drug class should strictly and carefully also be considered for patients with non-diabetic advanced renal disease. Both might lead to better long-term outcomes, as they delay advancement towards ESRD, which has a poor prognosis and is associated with very low quality of life. Long-term costs and outcomes of a very widespread disease would be influenced in a positive way, as also a life-prolonging effect was shown. As mentioned above, our research shows the need for a change in the current treatment guidelines. As both
drug classes investigated, ACE inhibitors as well as ARBs, are no longer patent-protected, it is not expected that pharmaceutical companies would contribute to supporting further indications for the latter. In the light of this research and clinical evidence published on this topic, they should not be withheld from the patient groups who would – in the majority of cases – benefit from early treatment. Further research in this field should be performed to confirm our findings, also in other countries as well as in other patient populations, e.g. in non-diabetic patients with mild to moderate renal failure.

We are aware of the fact that the development of treatment guidelines is a long process. Here, policymakers might interfere and help to create a setting where new scientific findings could be turned into practice more rapidly.
References

Valorisation
Valorisation

This dissertation aimed to contribute to the clarification of methodological aspects of economic evaluation in terms of bias and quality of life, as measured using the TTO method. It further aimed to assess the cost-effectiveness of ACE inhibitors and ARBs in renal disease, i.e. nephropathy, in various patient populations. All of these studies have clinical, economic and societal implications, as they might help to improve treatment and to more efficiently allocate scarce healthcare resources.

This chapter shows how the main findings of this dissertation could be applied in practice. In addition, possible dissemination strategies and implications for future research and practice are shown.

With regard to the first part of the thesis about methodological aspects, studies on bias in economic evaluations can be helpful for researchers as well as for policymakers, as they provide an overview of the possible biases that should be taken into account in order to perform a good evaluation or to rate an evaluation more adequately. This accounts for trial-based as well as for model-based studies. The ECOBIAS checklist that was developed in this thesis can be seen as a complementary tool next to existing guidelines and checklists to increase the quality of economic evaluations. We discussed the ECOBIAS checklist at this year’s annual conference of the Lowland Health Economists’ Study Group (LolaHESG). Although a lot of disciplines such as epidemiology and psychology have longstanding traditions of discussing bias, this tradition is not yet established in the field of health economics. One way to do that could be to create an ISPOR taskforce on bias. This could be the next step forward in raising further awareness of the importance of a systematic approach to bias in economic studies. The checklist could be used or recommended by editors and reviewers to assess bias in a systemic way. Submitting a checklist like ECOBIAS together with the research article could be made a requirement for authors, and journals could adjust their author instructions. This would make things more transparent, as literature further suggests that barriers to the use of economic evaluations include a lack of transparency, a limited understanding and sometimes also a lack of quality.\textsuperscript{1-3} Researchers could agree upon an open storage of models, which would allow for open access to the model structures and data. This would be a significant step towards achieving more transparency. An increased use of economic
evaluations will consequently lead to a more efficient way of allocating healthcare resources, which society will benefit from.

Limiting bias is becoming extremely important, as it increases confidence in and the usability of economic evaluations. Our research on bias has limitations, as discussed in detail in Chapters 2 and 3, and of course further research on bias is needed. We hope that our articles will stimulate other researchers to work further on bias in economic evaluation, which we would consider to be a success.

Our work on the time trade-off method showed that the inclusion of consumption costs in the numerator of the ICER should be considered, not only for life-extending but also for quality-of-life enhancing interventions, at least for those respondents who spontaneously consider non-health-related utility from treatment. Further research should surely be done with other patient populations, diseases and countries, but researchers could adapt those results to create a valid ICER, at least for the disease population investigated. Our results also show that researchers should provide a good description of the impact of disease on consumptive and leisure activities when doing exercises eliciting health valuations from the general public for the purpose of conducting economic evaluations from a societal perspective. Currently, this is rarely the case. This is something which clinical researchers as well as economists should consider and which can be transferred into practice immediately.

The second part of the thesis, i.e. cost-effectiveness analyses of ACE inhibitors and ARBs in renal disease in different patient populations, also has important implications. As there has been an increasing incidence and prevalence of patients undergoing renal replacement therapy in Western countries in recent years, ESRD represents a challenge for healthcare systems. The costs of ESRD treatment amount to €42,000 per patient per year in the Netherlands,\textsuperscript{4,5} and up to €45,000 in Germany\textsuperscript{6} (from a healthcare perspective, respectively), with the proportion of the national expenditures devoted to ESRD in European countries ranging from 0.7% in the UK to 1.8% in Belgium,\textsuperscript{4,7} as mentioned previously.

Hence, ESRD is not only a huge burden from the patients’ perspective, as it is associated with a reduced quality of life and life expectancy, but also for healthcare systems. Thus, prevention of ESRD is not only important from a patient and medical viewpoint, but also from a societal viewpoint. Treating
patients with ACE inhibitors or ARBs will delay the state of ESRD, which has a poor prognosis. Our research was extensively discussed on national and international levels, at medical as well as health economic conferences.\textsuperscript{8-18} It has been made available to expert panels and guideline committees and will at least be considered in the next update of treatment guidelines of diabetic and non-diabetic renal disease. Thus, the results of our studies suggest a change of current treatment guidelines. Nevertheless, clinicians should already adapt those results and adjust their treatment regimens. Besides clinical aspects, the economic aspects of treatment, as investigated in Chapters 5 to 7 of the thesis, should be discussed more with clinicians, e.g. at clinical conferences. Here, health economic topics are hardly to be found. This is most likely due to the fact that HTA does not play a role in medical education and that physicians are not familiar with tools and methodologies of health economists. An alternative would be to offer post-doctoral courses for clinicians to get deeper insights into HTA. Optimally, HTA should be better integrated into medical education so that medical students are confronted early and more intensely with HTA topics. Hence, a solid background could be created which allows physicians in training, and eventually physicians in clinical practice, to better understand the drawbacks and opportunities of health economic study results.

All actions mentioned above would have at least two positive effects: first, results from HTA would be transferred to practice faster; and second, overall acceptance of HTA could be improved and anchored into clinical medicine and patient-related treatment decisions.

Further research in other countries besides Germany and the Netherlands is needed to investigate and confirm the transferability of our results from the cost-effectiveness analyses of ACE inhibitors and ARBs. Although the drug classes under investigation are no longer patent-protected, pharmaceutical companies could contribute to the support of research in this field, where significant benefits for patients in terms of quality of life and reduced morbidity can be achieved. As this also accounts for mortality (a life-prolonging effect from treatment was shown), long-term costs and outcomes of a very widespread disease would be influenced in a positive way.

In the light of this research and clinical evidence published on this topic, those drugs should not be kept back from patient groups that would – in the majority of cases – benefit from early treatment. This would lead to better
outcomes in these patient populations and consequently to a societal benefit due to cost-savings in this disease area.

Summing up, our research has some important implications for researchers, clinicians and policymakers. It might be just a small step, and there are still some barriers to overcome, but bringing study results from health economics to clinicians is something this dissertation might have contributed to.
References

13. Adarkwah CC, Gandjour A. Cost-effectiveness of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers in newly diagnosed type 2 diabetes in Germany. Annual Conference of the German Association of Primary Care and Family Medicine (DEGAM), 20th-22nd September 2011, Salzburg, Austria.


Summary
Summary

The aim of the dissertation is threefold: first, to assess whether costs for consumption and leisure activities need to be considered in the incremental cost-effectiveness ratio (ICER) in cost-effectiveness analyses (Part Ia – methodological aspects, utility measurement, chapter 2), second, to review evidence and highlight important bias types in trial-based as well as in model-based economic evaluations (Part Ib – methodological aspects, bias in economic evaluation, Chapters 3 and 4) and third, to assess the cost-effectiveness of angiotensin-converting enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs) in diabetic as well as non-diabetic renal disease in different countries (Part II – economic evaluation studies, Chapters 5 to 7).

Chapter 2 of the thesis presents the results of a randomized study on patients with inflammatory bowel disease in Germany (n=104). The purpose of the study was to conduct an empirical survey on whether the effects of ill health on consumptive activities are spontaneously considered in a health state valuation exercise and how much this matters. Patients were randomly provided or not provided with explicit instructions regarding the consideration of consumption and leisure effects in a time trade-off (TTO) exercise. The study showed that explicit instructions to consider non-health-related utility in TTO exercises did not influence TTO scores. However, spontaneous consideration of non-health-related utility in patients without explicit instruction (60% of respondents) led to significantly lower TTO scores. Based on these findings, we recommend that consumption costs be included in the numerator of the ICER, at least for those respondents who spontaneously consider non-health-related utility from treatment. However, double-counting always has to be avoided.

To the best of our knowledge, for the effects of ill health on consumptive activities, no such empirical evidence has been gathered so far, although this study shows that it plays an important role in the valuation of utilities. Results also showed that exercises eliciting health valuations from the general public may include a description of the impact of disease on consumptive activities. Overall, this research might contribute to a more accurate ICER by providing further insights into the TTO understanding of patients.
Chapters 3 and 4 complete the methodological part of the thesis, both dealing with bias types and addressing the second aim of the thesis. Chapter 3 aims to give an overview of the potential risks of bias in trial-based economic evaluations and to identify how key sources for bias can be revealed and overcome. In total, eleven biases were identified and assigned to a particular trial phase. A distinction was made between pre-trial biases, biases during the trial and biases that are relevant after the actual trial. We discussed these bias forms in detail and presented strategies to detect and overcome them in a systematic way. The biases identified also account for model-based studies. Chapter 4 is directly related to Chapter 3. Here, we focused on bias types in model-based economic evaluation. It is obvious that bias types relevant for trial-based economic evaluation are usually relevant for model-based studies as well. So in this chapter, we focused on selected model-specific issues, using the Philips checklist as a tool to structure our results.

Eleven specific biases for model-based economic evaluations were identified and classified related to structure, data and consistency of the model, including structural assumptions, model type, time horizon, data selection (such as treatment effects), assessment of uncertainty and internal validation. In the end, together with biases which account for trial-based as well as for model-based studies identified in Chapter 3, a checklist for assessing bias in economic evaluations was developed (ECOBIAS). The ECOBIAS checklist encompasses a common part for model- and trial-based economic evaluation and a model-specific part. Part A of the checklist incorporates eleven biases that account for both types of economic evaluation, whereas Part B covers eleven model-specific aspects of bias.

Both studies demonstrated that there are several ways that biases can influence the final results of trial-based as well as model-based economic evaluations, which could change a result from being cost-effective to not being cost-effective at all. Combining the biases that can occur in trial-based with those in model-based economic evaluations identified in a previous article by the author group, the ECOBIAS checklist for identifying and avoiding bias in economic evaluation is introduced. To our current knowledge, we provide the first checklist (ECOBIAS checklist) on biases in trial-based and model-based economic evaluations. It is important to understand that different bias types can be related to each other and that the elimination of one can introduce another. The checklist can be used by researchers and
can be completed with submission of an article, but it can also be used by policymakers to check for bias in a systematic way.

In Chapter 5, a cost-utility analysis on the cost effectiveness of ACE inhibitors in newly diagnosed type 2 diabetes was performed. The aim of this study was to assess the most cost-effective time to start an ACE inhibitor in the event of cough in patients with type 2 diabetes in Germany. Three strategies were compared: treating all patients at the time of diagnosing type 2 diabetes, screening for microalbuminuria and screening for macroalbuminuria. A lifetime Markov decision model with simulated 50-year-old patients with newly diagnosed diabetes mellitus was developed using published data on costs and health outcomes and simulating the progression of renal disease. We showed that the treat-all strategy is associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. Our results were robust, even when considering a variety of different assumptions of uncertainty. It was concluded that patients with type 2 diabetes should be treated with an ACE inhibitor immediately after diagnosis if they do not have contraindications. Although a significant number of newly diagnosed type 2 patients may receive blood pressure medications, it has not been shown that these patients are primarily prescribed an ACE inhibitor, which underlines the significance of this analysis. Treatment guidelines in Germany and the Netherlands do not strictly recommend prescribing an ACE inhibitor in this patient population, so our analysis should lead to a correction of current treatment guidelines in a way that newly diagnosed type 2 diabetic patients should receive a low dose ACE inhibitor, even in the absence of hypertension or heart failure, which are currently the standard indications for this drug class. It is therefore clear that an ACE inhibitor should be considered a first-line treatment for hypertensive diabetic patients as well. As the topic investigated in Chapter 5 is of great importance not only from patient’s perspective but also from an economic viewpoint since diabetes is a huge burden for healthcare systems, at least in the Western world, Chapter 6 investigated this research question in the Dutch setting. The majority of data taken into consideration for the model in Chapter 5 were country-specific or at least adapted to the German setting. So the generalizability of the results is not given. We therefore built a new model for examining the cost-effectiveness of ACE inhibitors (or angiotensin II receptor blockers (ARBs) if coughing occurs as a side effect) in patients with newly
diagnosed type 2 diabetes in the Netherlands. We showed a similar trend as in the German setting: In the base-case analysis, the treat-all strategy was associated with the lowest costs and highest benefit and therefore dominates screening both for macroalbuminuria and microalbuminuria. A multivariate sensitivity analysis shows that the probability of savings was 70%. It was concluded that also in the Netherlands, patients with type 2 diabetes should be prescribed an ACE inhibitor immediately after diagnosis if they do not have contraindications. An ARB should be considered for patients who develop a dry cough when treated with an ACE inhibitor. In addition, accounting for both models, the potential for cost savings would be even larger if the prevention of cardiovascular events due to the ACE inhibitor treatment were considered.

Although effectiveness was shown in different trials, cost-effectiveness models did not yet exist in this area for the Dutch and German settings. A systematic review on ACE inhibitors and ARBs in type 2 diabetic nephropathy was published recently. Here, 30 articles were evaluated, including Chapters 5 and 6 of this thesis. Our articles were the only two studies that discussed that dry cough, as a possible side effect of ACE inhibitors, could potentially have an impact on the cost-effectiveness of the latter.

With respect to diabetes, diabetic nephropathy, with its final state of end-stage renal disease (ESRD), is a sequela that causes an enormous economic burden for healthcare systems due to high costs for either dialysis or renal transplantation. Considering this burden, other causes for ESRD are worth investigating in more detail. Diabetes itself accounts for about 40% of patients in renal replacement therapy and therefore represents the main cause for ESRD. Contrariwise, about 60% of patients with ESRD suffer from non-diabetic renal disease, which is focused on in Chapter 7. There is evidence that ACE inhibitors need to be handled with care in this patient population. Therefore, national as well as international clinical guidelines do not consistently recommend treating non-diabetic proteinuric patients with advanced renal disease with an angiotensin-converting enzyme (ACE) inhibitor. This study aimed to determine the cost-effectiveness of ACE inhibitor therapy in this patient population in the Netherlands by comparing two strategies: treating patients with advanced renal disease with an ACE inhibitor versus no treatment. A lifetime Markov decision model was developed to simulate the progression of renal disease using country-specific published data on costs.
and health outcomes. We found that treatment with ACE inhibitors leads to higher benefits and lower costs and dominates the no-treatment strategy. So it could be concluded that ACE inhibitor treatment for non-diabetic patients in the Netherlands, even those with advanced renal disease, is highly cost-effective and should therefore be considered. This should have a strong impact on current treatment guidelines. Our study was the first cost-effectiveness study to investigate this patient population with respect to ACE inhibitor treatment. Assessing cost-effectiveness in diabetic as well as non-diabetic patients with renal disease represented the second part of the thesis, addressing the third objective as mentioned above. Even though most guidelines recommend using a social perspective, where all costs and outcomes are considered important regardless of who pays for the costs and who will experience the effects, our analyses were performed from a healthcare perspective, which was considered more suitable and was stated clearly as well as justified in detail. Here, only costs and outcomes relevant to the healthcare system were considered. Not stating the perspective correctly or stating a societal view but not including all costs and outcomes might induce bias, which was also systematically described and investigated in this thesis as mentioned above. The overall findings of this research lead to a number of recommendations for clinical practice, policymaking and future research, which are mentioned and discussed in detail in the final part of the general discussion (see Chapter 8).
Acknowledgement
Acknowledgement

I am very grateful to all the people who participated in my PhD project, and I still cannot believe that this project has come to an end. It was a wonderful experience to further dive into research and to extend my knowledge in the field of HTA.

First of all, I would like to thank Silvia Evers, my promoter. Thank you, Silvia, for your continuous support and encouragement during the last few years. You are my guide and role model in science, my “research-mother”. I could not have had a better supervisor for my PhD project, and I really appreciate all the things you’ve done for me. I also immensely thank Mickael Hiligsmann, my co-promoter. Thanks for your support, for all the fruitful discussions we’ve had, for being critical, for giving input and being responsive all the time, and for helping me with administrative issues. I hope to continue the collaboration with my supervisory team in the future. I also thank Prof. Afschin Gandjour, who played an important part in this research. He introduced me to the world of health economics and HTA in 2005 and guided me through the HTA world for some years.

Next, I would like to thank my colleague in private practice, Dr. Klaus Mertens. Without his enormous flexibility and willingness at times to let me participate in conferences instead of working in the practice, this project would not have been realized in this way. Thanks, Klaus, for our well-organized and friendly relationship, cooperation and friendship, which I hope to continue the coming years. I would like to thank my professors and colleagues at the Department of General Practice / Family Medicine at the University of Marburg. They released me from work to present my HTA research at various conferences.

I want to thank the three co-authors of the articles of this dissertation, Dr. Maren Akkermann, Dr. Paul van Gils and Amir Sadoghi. I really enjoyed working with you, and I again thank you for your individual contributions to the research projects. I also thank the members of the Lowland Health Economists' Study Group (lolaHESG) for discussing two of the included articles intensively at the annual conferences in 2014 and 2015 and for giving valuable feedback.
As an external PhD student, I was in Maastricht in person only from time to time. Whenever I visited the HSR department, I was cordially received. A special thanks to the secretaries, Brigitte Caenen, Suus Koene and Janet Verheilen. I was in touch with all three from time to time to make appointments, discuss organisational issues and for conference planning and payment issues. You significantly contributed to my enthusiasm for Maastricht University due to your pleasant attitudes and your helpfulness.

I would like to thank Hannerieke van der Boom, the CAPHRI PhD coordinator for touching base regularly and helping with some organisational issues. Furthermore, I would like to say thanks to the whole HSR department and to CAPHRI for admitting me as an external PhD candidate and for being very supportive. Thanks for an enormous amount of support, also financially, with respect to conference fees and travel expenses, which allowed me to present parts of my research at national and international conferences. Performing research at CAPHRI, even as an external PhD candidate, has been a pleasure and is an experience I will never forget.

I gratefully thank the evaluation committee for showing interest in my work and taking time to review my dissertation.

I very much thank Tiny Wouters, who did the professional layout of this thesis for me. You are so uncomplicated and very fast, and I can recommend working with you to every other PhD student in a similar situation that I was in. Thanks to Florian Afflerbach, architect, drawer, artist and friend, who did the brilliant and unique cover layout.

Finally, I would like to thank some of the people who play an important role in my life outside of research. Thanks to my parents. You made me what I am today, and you encouraged me to keep on working on the project, even in heavy seas. I am very grateful and thank God that you can celebrate this unique moment with me, and I dedicate this piece of research to you.

I also in particular thank my partner, Annette, for your continuous support during the last few years. Without you, I would not have finished the project in 2015, and I will try to give you the same support with your work and research projects in the coming years.
Curriculum Vitae
Curriculum Vitae

Charles Christian Adarkwah was born on October 27, 1982, in Siegen, Germany.

He completed his secondary education at Friedrich-Flick-Gymnasium Kreuztal in 2002 at the top of his class. In 2003, Charles began his studies in medicine at the Universities of Giessen, Cologne and Bonn with study visits abroad at Harvard University (Boston), Yale University (New Haven) and Cornell University (NYC). He obtained the State Medical Exam (Staatsexamen) in 2009. He started working on his doctoral thesis in medicine at the Institute for Health Economics and Clinical Epidemiology at the University of Cologne in 2005 (Head: Prof. Lauterbach, Supervisor: Prof. Afschin Gandjour) and received a medical degree (MD) in 2011 (magna cum laude, Dr.med.). In 2009, he started working fulltime as a physician and researcher at the Department of Medicine III at RWTH University Hospital Aachen, Germany (Heads: Prof. Wolfram Karges, Prof. Christian Trautwein) and at the time started the Master of Science in Health Economics, Policy and Management as a part-time student at Maastricht University. Charles graduated from the program with distinction in 2011. His master thesis, supervised by Prof. Silvia Evers, was awarded with the Catharina Pijls Incentive Prize 2011 (award for an excellent master thesis in the field of Health Sciences). Charles started working on his PhD in 2011 as an external PhD student. In 2012, he left the University Hospital Aachen to complete his clinical training in a large group practice for family medicine (Dr. Mertens, Dr. Adarkwah, Dr. Hoxha) in Kreuztal. Simultaneously, he started working as a researcher at the Department of General Practice/Family Practice (Heads: Prof. Erika Baum, Prof. Annette Becker, Prof. Norbert Donner-Banzhoff). Charles was board certified in Internal Medicine and Family Medicine in 2014. He was further board certified in Palliative Medicine in the same year. He is currently working as a general practitioner in a group practice, and as a researcher and lecturer at the University of Marburg. He is further lecturing at FOM University of Applied Sciences.

With his research, he tries to bridge the gap between clinical practitioners and health economists.
Peer reviewed scientific publications


Published abstracts


Presentations at national & international conferences


3. Adarkwah CC, Gandjour A. Cost-effectiveness of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers in newly diagnosed type 2 diabetes in Germany. German Diabetes Association Annual Conference (DDG), June 3rd, 2011, Conference Center Leipzig, Leipzig, Germany. (poster presentation)


7. Adarkwah CC, Gandjour A. Cost-effectiveness of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers in newly diagnosed type 2 diabetes in Germany. Annual Conference of the German Association of Primary Care and Family Medicine (DEGAM), 20th-22nd September 2011, Salzburg, Austria. (oral presentation)


24. **Adarkwah C**, Sagoghi A, Gandjour A. Should Cost-Effectiveness Analysis Include the Cost of Consumption Activities? An Empirical Investigation. 6. lolaHESG conference, 22.-23.05.2014, Oostvoorne (Rotterdam), NL. (discussion paper; oral presentation)


38. Adarkwah CC, van Gils PF, Hiligsmann M, Evers S. Risk of bias in model-based economic evaluations. 7. lolaHESG conference, 28.-29.05.2015, Maastricht, NL (discussion paper; oral presentation)

Grants / Awards

1. German Research Price for General Practice 2013 (Dr. Lothar Beyer-Preis, 3. Preis) for project „OptRisk: Optimierung der Risikoberatung durch Darstellung der Veränderbarkeit der individuellen ereignisfreien Lebenszeit, 3,000 Euro (price obtained by study group)

2. Catharina Pijls Incentive Prize 2011, Maastricht University; Award for excellent Master thesis in the field of Health Sciences, 1,000 Euro

3. Travel Grant of the German Renal Foundation (Dt. Nierenstiftung) 2011, 1,500 Euro

4. Grant of the Dr.-Friedrich-Flick-Stiftung Düsseldorf 2002, Award for excellent A-Level, 2,600 Euro