# Decision making in drug reimbursement

Margreet Franken

# Decision making in drug reimbursement

Margreet Franken

**Funding**
The studies performed in this thesis were performed in the context of the Escher project (T6-202), a project of Top Institute Pharma (chapter 3–6), and financially supported by the Dutch Health Care Insurance Board (chapter 9 and 10).



ISBN: 978-94-6169-587-1

Cover design: Optima Grafische Communicatie, Rotterdam, The Netherlands
Layout and printed by: Optima Grafische Communicatie, Rotterdam, The Netherlands

# Decision making in drug reimbursement

## Besluitvorming inzake geneesmiddelenvergoeding

**Proefschrift**

ter verkrijging van de graad van doctor aan de

Erasmus Universiteit Rotterdam

op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 5 december 2014 om 11.30 uur

door

**Melinde Gretha Franken**

geboren te Barneveld

## Promotiecommissie

**Promotoren**
Prof.dr. C.A. Uyl-de Groot
Prof.dr. A. de Boer

**Overige leden**
Prof.dr. A. Boer
Prof.dr. I. Cleemput
Prof.dr. F.F.H. Rutten

**Copromotor**
Dr. M.A. Koopmanschap

## Contents

# 1

# General introduction

## Background

Life expectancy is increasing in the Organisation for Economic Co-operation and Development (OECD) countries. Simultaneously, however, many countries have experienced rising health care expenditure, albeit at a reduced rate during the last years due to the global economic recession.[1-3] For instance, health care expenditure in The Netherlands has increased from 6.5 billion to 46.9 billion and 94.2 billion euro in 1972, 2000 and 2013, respectively.[4] More apprehensively, not only absolute expenditure but also the relative share of Gross Domestic Product (GDP) spent on health care has increased across OECD countries. For example, the Dutch share of GDP spent on health care increased from 8.7%, to 11.2%, and 15.6% in 1972, 2000, and 2013, respectively.[4]

It is expected that health care expenditure will continue to rise globally due to ageing populations and the continual development of new medical technologies.[5] Since the organisation of health care is mostly implemented at a national level, governments face major challenges to keep their health care system financially viable in the long term. The major challenge is to appropriately balance everyone's legal right to health and access of health care (*Article 25.1 Universal Declaration of Human Rights*) given the increasing demand for and supply of health care with the financial sustainability of the system. Consequently, it is inevitable that priorities must be set to ensure efficient and equitable use of limited health care resources.

Scarcity of resources implies that choices must be made regarding access to health care. The health economic science is concerned with the question of how to allocate scarce resources given individual and social objectives. In the past decades, the health economic field of Health Technology Assessment (HTA) has received particular attention within health care policy making. HTA is a multidisciplinary field of policy analysis that entails the medical, social, ethical, legal, organisational, and economic implications of the development, diffusion, and use of a health technology,[6] with the aim to facilitate informed social decision making regarding the application of the technology.[7] Concerns about increasing costs of medical technologies as well as concerns about the non-evaluated benefits of many technologies stimulated the development and use of HTA.[8] Since the 1990s, many European countries established formal HTA agencies to inform health care decision making.[9,10]

## Priority setting for pharmaceutical products

The pharmaceutical market is one of the most regulated sectors. Pharmaceutical expenditure accounted for a total of 800 billion USD across OECD countries in 2011, comprising on average 17% of total health care expenditure.[2] Until 2009, the share of pharmaceutical expenditure increased faster than total health care expenditure, but due to the economic recession, annual growth rates decreased significantly. Interestingly, there is a wide variation across OECD countries in pharmaceutical spending per capita (OECD average USD483 PPP; ranging from USD178 PPP in Chile to USD985 PPP in United States of America) as well as share of GDP (OECD average 1.50%; ranging from 0.56% in Luxembourg to 2.63% in Hungary).[2]

A pharmaceutical product is only allowed to enter the market after the demonstration of its efficacy, quality and safety to the market licensing authority (e.g., the European Medicines Agency, and the United States Food and Drug Administration). However, in order to contain health care expenditure, many countries developed policies and decision structures to control the basic benefit package. Most of these policies have been directed at controlling pharmaceutical expenditure.[11] This implies that, in order to justify public funding, many countries additionally require (HTA) evidence regarding the cost efficiency of a pharmaceutical product. This requirement is often perceived as an additional barrier to market access and has therefore been labelled as 'the fourth hurdle'.[12] Due to differences in health care organisation across countries there are differences in drug reimbursement systems and reimbursement policies. There is, however, little evidence on the efficiency and sustainability of these different reimbursement policies. A detailed comparison of various systems could help identifying systems' strengths and weaknesses and could thus facilitate policy learning and provide lessons to improve system efficiency and sustainability. However, so far, most studies have been descriptive in nature; more analytically oriented studies could enhance understanding of decision-making processes.[13,14]

This thesis focuses on decision making in drug reimbursement. It consists of a descriptive part of European drug reimbursement systems' procedures, processes, and their applied criteria, but also provides a detailed analysis of the actual use of reimbursement criteria in everyday decision making, evaluations of policy tools to handle uncertainty of the evidence, an evaluation of strengths of and challenges for the systems, as well as suggestions on how the decision-making process could be improved.

## Drug reimbursement decision making

Many countries institutionalised priority setting for pharmaceutical products at the national level; for instance, based on a 'positive' or a 'negative' list. In this thesis, the term *drug reimbursement system* is used for the established policy system that determines whether or not a drug is entitled for reimbursement (i.e., public funding). We distinguish four different phases in the reimbursement decision-making process (see Figure 1.1). Firstly, the *assessment phase* entails the quantification of the clinical, pharmacotherapeutic and pharmacoeconomic value of a drug in comparison to other available drugs. The assessment can also include a description of other health care sector related arguments (e.g., ethical and organisational issues). Assessment is descriptive in terms of quality and uncertainty of evidence, thus, it does not include a value judgement. It provides a description of the available evidence, the level of quality of the evidence and the level of uncertainty related to the quantifications. Secondly, the *appraisal phase* entails the evaluation of the social value of a drug by weighing the assessment outcomes against other criteria related to the objectives of the health care system (e.g., necessity to pay out of public funding). Appraisal seeks to gauge societal willingness to pay for a specific drug out of public funding. A consequence of weighing criteria is that the appraisal includes a value judgement of the drug. Thirdly, the decision whether or

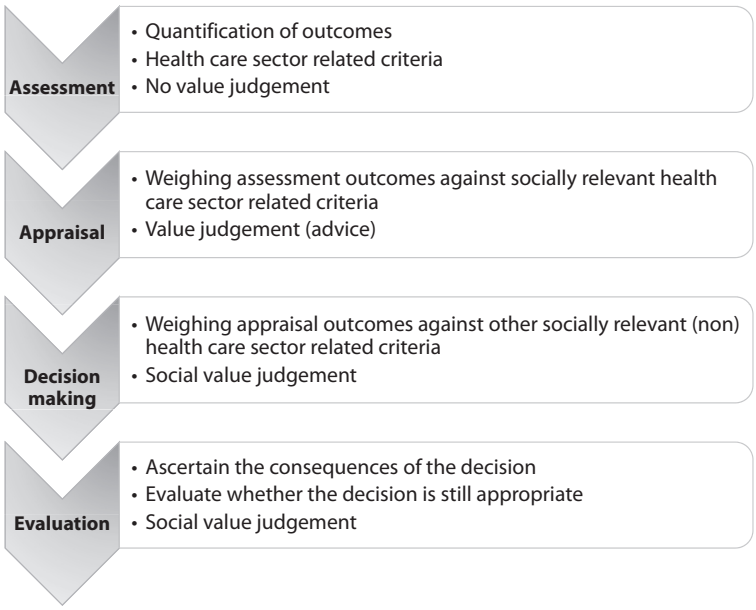


**Figure 1.1** Four phases in the drug reimbursement process*
* Adapted from Le Polain, Franken et al.[15]

not to reimburse the drug is made in the *decision-making phase*. The decision is based on a value judgement from a broader social perspective. As such, the value judgement includes considerations of both health system objectives as well as non-health system objectives (e.g., employment in pharmaceutical industry). Finally, the last phase in the process is the *evaluation phase*. This phase entails ascertaining the consequences of the decision and evaluating whether the initial decision is still appropriate in case new evidence becomes available (in which case the cycle of assessment, appraisal, and decision making will be reiterated).

## Uncertainty of the evidence in the decision making

Reimbursement decisions are inherently made under uncertain conditions due to the absence of complete information.[16,17] Evidence required for market access is mainly based on randomised clinical trials (RCT). These trials are conducted under 'ideal circumstances' and assure internal validity by randomly assigning patients to a treatment strategy. However, external validity may be hampered because patients treated in everyday practice may not be comparable to 'ideal patients' in an RCT setting. Moreover, RCTs often use intermediate efficacy estimates which may not be fully predictive of effectiveness endpoints such as improving quality of life and prolonging life in a real-world setting. Consequently, uncertainty may arise on the actual clinical benefit, the adoption and diffusion, value for money, and the economic impact of the new drug.[18]

However, evidence from an RCT may be very promising and can thus put a high pressure on policymakers to ensure early access to promising drugs for patients in need. For example, in the late 1990s the Dutch minister ensured early access to promising drugs for AIDS.[19] More recently, various policies are implemented to guarantee access to new promising treatments for cancer[20] (e.g., the Cancer Drugs Fund in the United Kingdom). Van Luijn et al.[21] reported that between 1999 and 2005 only 48% of newly approved drugs had comparative efficacy data (i.e., compared with an active treatment) available at the moment of European market authorisation. Similarly, Goldberg et al.[22] reported that between 2000 and 2010 only 51% of newly approved drugs had comparative efficacy data available for market authorisation in the United States. Consequently, policymakers face the challenge to strike an optimal balance between ensuring timely access and having sufficient evidence of drug's comparative value. Policymakers could consider postponing the reimbursement decision until sufficient evidence becomes available.[16] This is, however, at the cost of delaying access. Another option would be to link the price paid for the drug not only to its actual value but also to the quality of the available evidence; this would incentivise manufactures to invest in good quality evidence.[23] In

the last decade, many systems introduced policies to reduce the clinical and economic uncertainty of the performance of novel drugs by making reimbursement conditional on additional evidence collection.[18,24-26] The gathered evidence should fill the gap between the evidence from RCTs and treatment in everyday practice. The policies address clinical and/or financial uncertainty; for example, by means of implementing patient access schemes,[27] managed entry agreements,[28] coverage with evidence development schemes,[29-31] outcomes research requirements,[32,33] and finance-based[34,35] or outcomes-based [35,36] risk sharing agreements. There is, however, little empirical evidence regarding the effectiveness of these different schemes.

## Objectives

The overall aim is to describe, analyse and evaluate systems that make decisions on reimbursement of drugs. The thesis begins with exploring preferences for the objectives of a health system. To obtain insights into drug reimbursement decision-making procedures and processes, we describe and compare five European drug reimbursement systems. Based on this, a framework is presented to improve upon the legitimacy of the decision-making process. Subsequently, the use of two reimbursement criteria (cost-effectiveness and disease severity) is assessed in actual decision making. Finally, one of the policy tools to handle uncertainty of evidence (coverage with evidence development) is evaluated regarding its effectiveness, feasibility, and appropriateness.

The following research questions are addressed:

1. What are the objectives of a health care system?
2. What criteria are important in drug reimbursement decision making?
3. How do drug reimbursement decision-making systems handle uncertainty of evidence?
4. To what extent do drug reimbursement systems satisfy the conditions of legitimate decision making?
5. What are the strengths of and challenges for drug reimbursement decision-making systems?
6. Based on the research findings, what suggestions can be put forward to improve decision making in drug reimbursement?

## Outline

This thesis is structured in three sections. Part A is a general explorative introduction into social preferences for the goals of health systems. Chapter two describes five main goals of a health system including their theoretical foundation, and presents a framework, based on a multi-attribute choice technique, for measuring relative preferences for these goals. This chapter also presents goal valuations which were obtained in The Netherlands.

Part B focuses on drug reimbursement procedures, processes and criteria. Chapter three describes a comparative analysis of five European drug reimbursement systems (between Austria, Belgium, France, The Netherlands, and Sweden). The chapter provides insights into reimbursement procedures and processes, and identifies strengths and weaknesses of the five systems. Based on this comparative analysis, chapter four presents a framework for policymakers to improve upon the legitimacy of the decision-making process. Chapter five to seven explore the actual importance of two reimbursement criteria (the disease severity and cost-effectiveness criterion) in everyday decision making. Firstly, in chapter five, the role of the criterion disease severity is assessed in four European decision-making settings (in Belgium, France, The Netherlands, and Sweden). Secondly, chapter six provides a comparative analysis of the actual importance of the criterion cost-effectiveness in Dutch and Swedish drug reimbursement decision making. Chapter seven more specifically focuses on the actual impact of cost-effectiveness in The Netherlands.

Part C provides empirical evidence of the coverage with evidence development policy tool to handle uncertainty of evidence in drug reimbursement decision making. Chapter eight assesses whether the Dutch policy is effective towards reaching its objective. The next chapter provides insights into the practical feasibility of developing evidence on drug use and cost-effectiveness in everyday clinical practice. The last chapter shows whether this particular policy reduces policymaker uncertainty and debates whether other policy options could be appropriate to ensure sufficient value for money of expensive drugs.

Finally, chapter eleven reports the main findings and discusses the merits and limitations of the research presented in thesis, and presents suggestions on how to improve decision making in drug reimbursement and explores challenges for future research.

It should be noted that chapters two to ten are based on publications in, or intended for, scientific peer reviewed journals. These chapters can, therefore, be read independently, and some overlap may exist between these chapters.

# A

## Social preferences for health system objectives

# 2

# Health system goals: A discrete choice experiment to obtain societal valuations

With Xander Koolman

Health Policy 2013; 112(1-2):28-34

## Abstract

The aim of this study is to improve previous approaches to health system goals valuation.

We reviewed literature on health system performance and previous comparative performance assessments, and combined this with literature on process utility to create a theoretical foundation for health system goals. We used a discrete choice experiment to elicit goal weights. To obtain social justice weights respondents were placed behind a 'veil of ignorance'. To ensure that respondents understood their task, we instructed them in a classroom setting.

We identified five health system goals. All five goals significantly affected choice behaviour. An equitable distribution of health obtained the highest weight (0.34), followed by average level of health (0.29) and financial fairness (0.24). Both process outcomes (utility derived from the process and its distribution) received much lower weights (0.07 and 0.06, respectively).

Our framework adds to that of the World Health Organization. We demonstrated the feasibility of measuring societal valuation of health system goals with a multi-attribute technique based on trade-offs. Our weights placed much greater emphasis on health and health inequality than on process outcomes. Our study improves the methodology of international health system performance comparison and thereby enhances global evidence-based health policy information.

## Introduction

Health systems around the world have contributed to better health and life expectancy with varying degrees of success. Even in countries with seemingly similar resources outcomes vary markedly.[37] To date, policy effects on the performance of health systems remain largely unclear. Monitoring and evaluating performance can generate this vital policy information. Moreover, cross-country comparisons enable countries to learning from others.

The challenge, however, is how to assess health systems that are extremely complex and have multi-dimensional goals. This complex task has been explored by the World Health Organization (WHO) and the Organisation for Economic Co-operation and Development (OECD).

Our aim is to improve the valuation of health system goals. To do so, we first unify literature on health system goals, equity, and process utility to create the underpinnings of a theoretical framework for health system evaluation. Second, we review previous approaches to deriving relative weights for health system goals. Third, we suggest an enhanced methodology based on a multi-attribute choice technique to elicit goal valuations using a 'veil of ignorance' perspective. Last, we present goal valuations for The Netherlands based on our proposed method.

## Theoretical framework for health system goals

Until the 1990s health economics was dominated by the assumption that 'health' was the dominant outcome of health systems. This links with the consequentialism moral theory, which focuses solely on outcomes irrespective of the process that led to them, and Jeremy Bentham's 'act utilitarianism', which states that the greatest amount of happiness for the greatest number of people determines choice behaviour. Although utility derived from health is an obvious outcome of a health system, research showed that people also care about the processes that precede health outcomes, irrespective whether they affect health.[38-40] Therefore, processes are not just means to an end, not just instrumental to an intrinsic goal, but are an intrinsic goal of the health system.

This utility derived from processes, procedural utility, has a base in social sciences. Parsons' social action theory (1937) already described the necessity of the subjective dimension of human action.[41] Psychologists have developed a comprehensive notion of basic psychological needs for the human self, evident in the 'self-determination theory of

intrinsic motivation' by Deci and Ryan.[42,43] The theory maintains that human motivation originates from three innate needs: autonomy, competence, and relatedness; individual well-being therefore depends on procedures that address them.[42,43]

The theory of procedural utility can be directly applied to health care. Consequently, both health outcomes and the process attributes of non-health outcomes are health system goals. Furthermore, it is widely recognised that health systems' costs should be related to capacity to pay rather than the risk of illness.[44] Therefore, health systems have three independent outcome-oriented objectives: health utility, process utility, and financial fairness.

### Health utility

Health systems aim to improve health and strive for the highest possible health status of the entire population, taking both morbidity and mortality into account. Behind a 'veil of ignorance' the distribution of health also matters; empirical evidence indicates that the public is willing to trade efficiency for social objectives such as equity.[45-47] Therefore, health utility consists of two goals: average level of health and the equitable distribution of health.

### Process utility

Procedural utility can arise from two sources.[48] First, interaction between people can generate utility since people evaluate actions by how they are treated by others. Second, people have preferences for good institutions in addition to health outcomes (e.g., preferences on allocative and redistributive decisions) that address the innate needs of human motivation (autonomy, competence, and relatedness). Institutions also establish the fundamental rules for societal decision making. As a result, process attributes of health systems are twofold: utility derived from interaction between people and the health system (how people are treated by the health system), and utility obtained from living under institutions (how allocative and redistributive decisions are taken). Although distributional fairness of process utility is not well founded in moral theory, we followed the WHO framework and therefore included both process utility and its distribution in our framework.

### Financial fairness

Murray et al.[49] claim that a health system is fairly financed *"if the ratio of total health system contribution of each household through all payment mechanisms to that household's capacity to pay is identical for all households, independent of the household's health status or use of health system."* This signifies two key challenges. First, households should not pay an excessive share of their income for health care or become impoverished.[50]

Second, wealthy households should contribute more than poor households reflecting vertical equity and an element of progressivity.



## Existing international frameworks

Several countries, such as the USA, United Kingdom, The Netherlands, Australia, and Canada, have designed and implemented national schemes and indicators to measure health system performance.[51] Cross-country comparison, however, requires a comprehensive international framework such as those of the OECD and WHO.

*OECD framework*. The three main goals of the OECD framework are (i) health improvement and outcomes, (ii) responsiveness and access, and (iii) financial contribution and health expenditure.[52] Without suggesting any relative importance of the system goals it provides a framework to measure performance in several dimensions that seem to be based on the historical development of health systems. A composite score requires, however, each goal to be independent. The OECD framework consists of input and output variables, and intermediate as well as end goals. Consequently, using the OECD framework gives rise to methodological problems when weighing goals.

*WHO framework*. The WHO framework for performance measurement consists of three intrinsic goals of health systems: health, responsiveness, and fairness in financing.[37] The first two are assessed on both level and fairness of the distribution. The framework satisfies the required conditions (i.e., a complete set of intrinsic goals) to facilitate global performance assessment.

WHO's health system goals closely resemble those identified for our own theoretical framework. Health utility and its distribution are reasonably comparable to WHO's level and the distribution of health. Our two sources of process utility can be described by WHO's assessment of quality and equity of responsiveness. Last, one could suggest that financial fairness reflects WHO's fairness in financing.

## Valuing health system goals

### Previous approach

WHO's goal weights were acquired by measuring preferences of individuals with health system knowledge via an internet-based questionnaire, which included interactive, weight-assigning pie charts, descriptive multiple choice questions, and ranking tasks.

The final weights were rounded to the nearest one-eighth for the World Health Report 2000 (WHR) to make the composite goal easier to understand.[53] In 2000 and 2001, the WHO performed a follow up multi-country study to measure preferences from the public.[54] These questionnaires consisted of ranking tasks and pie charts. The Dutch sample included 1566 respondents: 1068 face-to-face and 498 postal interviews.[54]

**Critiques on the WHO approach**

Much critical attention and debate followed the publication of the WHO results. One of the major concerns was that WHO's relative weights were highly subjective since they were derived from respondents who were far from representative.[55-57] The WHO justified its method by stating: *"the purpose of the first survey was not to describe preferences in a population, but rather empirically derive a set of weights reflecting normative choices."*[58]

Richardson et al.[59] maintained that effective weights depend on variation in scores across countries as well as the nominal weights. Consequently, if there is no difference in, for example, health inequality then, regardless of the weight of 0.25, it would contribute nothing to the ranking scores. However, applying a standard set of weights appears to deny differences between countries' ideologies and theories of social justice.

Moreover, the Oswaldo Cruz Foundation[60] concluded that the composite index is very sensitive to modification in the relative weights and claimed that some countries can shift the scale by more than thirty points by small weight adjustments. On the contrary, Lauer et al.[61] concluded that even large changes would have a small impact and that all rankings remained within the uncertainty intervals. Nevertheless, Lauer et al.[62], in yet another publication, argued that the differences could be large for individual countries because of the impact of publication of the rankings.

The most essential critiques concern the valuation methodology. Smith[57] argued that WHO's methodology was highly questionable and that it is unlikely that it would elicit the required relative marginal valuation of an extra unit of performance. Moreover, Williams[56] claimed that the main issue was the use of rankings, scores, and rating scales rather than facing respondents directly with trade-offs.

Making trade-offs is at the heart of economics. In a multi-attribute environment, such as a health system, individuals choose between attributes based on their relative importance. The WHO limited their survey by using pie-charts, rankings, and descriptive multiple choice questions; their instrument did not allow for deliberation.

### Enhanced methodology

The Multi Attribute Utility Theory (MAUT) provides a foundation for valuing complex and multi-dimensional objectives. It has its roots in classical measurement theory and theories of economic choice behaviour. In the last decade, multi-attribute valuation methods and especially discrete choice experiments (DCEs) have become increasingly popular. DCE draws upon Lancaster's economic theory of value[63] and is firmly rooted in the random utility theory.[64]

Furthermore, Dolan et al.[65] distinguish respondents' perspectives in two dimensions: who the respondent should have in mind and at what point in time. The former dimension concerns oneself, other people, or all people; that is, preferences are personal, social, or socially inclusive, respectively. The time dimension relates to the context in which the valuation is obtained, that is, whether it is *ex ante* or *ex post*. The structure of a health system depends in large part on society's choices concerning resource allocation; relatedly, a person's societal position often reflects a feeling of social (un)fairness. Therefore, it is essential to elicit *ex ante* socially inclusive personal preferences for social justice valuations. Valuations should therefore be obtained in a procedurally fair way reflecting a social justice perspective; i.e., respondents assume a Rawlsian 'veil of ignorance'. The main characteristic of Rawls's approach is that people make choices without knowing where their own position might be in a society.[66] The 'veil of ignorance' ensures that the principles of justice are blind to age, health status or societal position.

An enhanced methodology should therefore be based on a multi-attribute choice technique such as a DCE. This technique applies direct trade-off questions, improves conscientious deliberation, and elicits marginal valuations. To ensure social justice valuations, goal valuations should be obtained using an *ex ante* perspective.

## Methods

We conducted a DCE to obtain goal valuations in The Netherlands. Our five-step procedure, typical of a DCE, was to: (1) identify and describe the attributes for health systems; (2) assign attribute levels based on goal variation; (3) combine attribute levels and create hypothetical scenarios; (4) establish goal valuations; and (5) analyse and interpret the data.

### Attributes and levels

Our health system goals satisfy the attribute criteria by being complete, operational, decomposable, non-redundant and minimum-sized.[67] For use in a questionnaire, at-

tributes must be meaningful, relevant and easy to understand, and their levels should be plausible, actionable and tradable.[68] Variation in their levels should mimic existing variation.

We pilot-tested attribute and level descriptions in 54 participants [data not shown]. They were interviewed face-to-face or by telephone and were asked to fill out a questionnaire while thinking aloud. Pilot questionnaires offered 30 questions comprising paired comparison of two attributes in which vignettes varied by one level. The interviewer could ask about the basis for respondents' trade-offs and the extent to which they understood the descriptions. Answers provided insight into respondents' ways of thinking and revealed any misinterpretation or misunderstanding of the notions of the attributes. Respondents initially found it difficult to make a clear distinction between the intrinsic and instrumental contributions of process factors. The pilots also showed that respondents changed their behaviour when choosing from behind a veil of ignorance, some explicitly mentioning that equity was more of a concern. Preferences concerning solidarity versus individualism were partly balanced out. Table 2.1 shows the attributes and the final descriptions of the attributes and levels which were used in the DCE exercise.

**Table 2.1** Attributes and descriptions of attributes and levels used in the DCE exercise

| Attribute | Attribute description | Level description |
|---|---|---|
| Average level of health | Average health adjusted life expectancy | 74, 72, 70 and 68 years |
| Distribution of health | Differences in health adjusted life expectancy across social groups | 3, 5, 7 and 9 years difference |
| Average level of process outcome | Patient experiences | very good experiences, good experiences, rather good experiences, reasonable experiences |
| Distribution of process outcome | Differences in experiences between patients across social groups | no difference, small difference, some difference and fair difference |
| Financial fairness | Persons in poverty because of health system payments | 0, 1, 2 and 3 person(s) per 200 persons |

*Health outcome*

Utility obtained from health outcome reflects the average level of health in a population expressed by life expectancy that takes both morbidity and mortality into account. The levels are actual Health Adjusted Life Expectancy (HALE) years in The Netherlands: 74, 72, 70 and 68 years.

*Distribution of health outcome*

We assessed existing inequalities in HALE in The Netherlands. Inequality is present across social classes, educations groups, income levels, ethnic groups, gender and geographic areas. The differences across social groups in HALE were 3, 5, 7 and 9 years.

*Process outcome from interactions*

Because the pilots revealed that respondents experienced difficulties in understanding the process-outcome concept, the concept and its implications were clarified and illustrated with recognisable examples from the Dutch health system during the DCE exercise. The levels in experience were very good, good, rather good, and reasonable. This variation mimics the actual variation in patient experiences in The Netherlands as measured by a population based study of the WHO (average responsiveness score of 85.7 on a scale of 0 to 100).[69]

*Distribution of process outcome*

Although distributional fairness of process utility is not well founded in moral theory, we followed the WHO framework and allowed our empirical test to indicate the value respondents attached to an unequal distribution of process utility. This attribute thus indicates the inequality in the distribution of process outcome across groups. In The Netherlands options exist to bypass waiting lists and 'buy' quality but the majority of the population sees this as inequitable. The Health Insurance Act has since 2006 been moving towards managed competition in health care and it is thus likely that more options will become available to buy additional insurance or pay out-of-pocket for extra quality and other process attributes in the future. Therefore, our levels are actual differences in distribution and partly reflect potential future developments. The levels of difference offered were none, small, some, and fair. The differences in the levels were clarified and illustrated with recognisable examples from the Dutch health system.

*Financial fairness.* The last attribute describes financial protection against costs of illness. We explored national health insurance premiums and potential co-payments, and connected these financial flows with catastrophic payment and impoverishment. We selected a quasi-relative poverty measure found to be recognisable and accurate for the core perception of poverty within the Dutch population.[70] We expressed the levels as consequences of variations in (co-) payments. The levels offered were 0, 1, 2 and 3 extra person(s) per 200 persons unable to satisfy basic needs because of health system payments.

## Experimental design

A full factorial design would generate 1024 ($4^5$) scenarios. To reduce the DCE exercise to a manageable level, we applied a fractional factorial design. We obtained an orthogonal

array from the online Sloan library in order to assign 16 choice sets (http://www.research.att.com/~njas/oadir). We applied an optimal design generator (12332, 21123 and 33211) based on strategies described by Street et al.[71] Such a fold-over strategy provides an optimal design with a high efficiency for estimating main effects while satisfying the statistical properties of level balance, orthogonality, minimal overlap, and utility balance.[68] Main effects usually account for 70–90% of explained variance.[64] The order of vignettes was randomly varied for all three choice sets.

We did not include an opt-out option because (i) it can generate problems such as applying heuristics to prevent making difficult choices or preferring a status quo and (ii) it is impossible to opt-out by choosing not to have any health system. Our respondents were choosing from behind a veil of ignorance and thus we assumed that there was no default system, i.e., an opt-out option was non-existent.

### Data collection

We selected 63 persons familiar with health systems to ensure that respondents understood their task. Moreover, respondents were instructed by the principal investigator in a classroom setting via a PowerPoint presentation and a 20-minute interactive discussion about the meaning of the attributes and their levels. In particular, the distinction between the intrinsic and instrumental contribution of both process attributes was explained and the levels were made operational by illustrative examples from the Dutch health system. Each respondent was provided with the information on paper to refer to if needed. Respondents had to choose in which country they would prefer to be born.

### Data analysis

The software program STATA was used to perform data analysis. Given the exploratory character and aim of our study to investigate the feasibility to estimate health system goal valuations using a multi-attribute technique based on trade-off questions, we used a conditional logit model to estimate the coefficients. The estimated function for the valuations of health system goals was: $Ylatent = \beta_0 + \beta_1$* average level of health + $\beta_2$* distribution of health + $\beta_3$* process outcome + $\beta_4$* distribution of process outcome + $\beta_5$* financial fairness + $\varepsilon$.

## Results

### Econometric model

All responses were included in our data analysis since all respondents correctly answered the dominant question. Table 2.2 presents the results of the conditional logit model, which provides good insight into respondents' trade-off behaviour because the

**Table 2.2** Results conditional logit model

| Attribute | Coefficient | Std. error | z value | P > \|z\| | 95% Conf. interval |
|---|---|---|---|---|---|
| Average level of health | -1.097 | 0.067 | -16.31 | 0.000 | -1.229; -0.966 |
| Distribution of health | -1.294 | 0.073 | -17.84 | 0.000 | -1.436; -1.152 |
| Average level of process outcome | -0.277 | 0.058 | -4.73 | 0.000 | -0.391; -0.162 |
| Distribution of process outcome | -0.236 | 0.061 | -3.84 | 0.000 | -0.356; -0.116 |
| Financial fairness | -0.927 | 0.063 | -14.64 | 0.000 | -1.051; -0.803 |
| Log likelihood | -880.674 | | | | |
| Chi-Square test | 948.75 | | | | |
| Pseudo $R^2$ | 0.350 | | | | |

outcome probabilities are based only on the attributes and their levels. The summary statistics show that the model had a decent fit with a pseudo $R^2$ of 0.35 and a statistically-significant Chi-square test of 948.75.

The coefficients reflect the relative importance of the system goals. As expected, they revealed that respondents favour the best attainment in all five goals. The associated p-values indicated that all attributes have a statistically significant effect on choice behaviour. The table shows that health distribution received the highest importance followed by average level of health, financial fairness, average level of process outcome, and distribution of process outcome, respectively.

### Marginal rate of substitution

The marginal rate of substitution (MRS) is computed by dividing the coefficients and demonstrates the trade-off between attributes. Table 2.3 shows that individuals are willing to give up 0.848 level in health distribution to gain one level in average health ($\beta_1/\beta_2$), meaning that individuals will trade 2 years of average health for 1.70 years of health inequality.

### Valuation as percentage

To enable direct comparison between WHO's relative weights and our valuations, our coefficients needed to be converted into percentages by dividing the coefficient of one attribute by the sum of all coefficients. This method is based on the same assumption as the MRS calculation, that is, linearity of the coefficients and a comparable realistic amount of variation between the levels for each of the attributes. Table 2.4 shows the results.

Although the sum of the quality (0.36 vs. 0.37) and equity (0.64 vs. 0.63) objectives are similar, their weights are derived differently. Specifically, our valuations place much greater emphasis on health (0.63 vs. 0.49) and far less on process outcomes (0.13 vs. 0.29). Furthermore, our results show a greater weight for health distribution (0.34) compared to the average level of health (0.29), whereas these weights were almost similar in WHO's survey (0.25 and 0.24, respectively). Fairness in financing is roughly equally weighted in both methodologies.

**Table 2.3** Marginal rate of substitution

| Coefficients | Attribute/ Attribute | Observed MRS |
|---|---|---|
| $\beta_1/\beta_2$ | average level of health/ distribution of health | 0.848 |
| $\beta_1/\beta_3$ | average level of health/ average level of process outcome | 3.968 |
| $\beta_1/\beta_4$ | average level of health/ distribution of process outcome | 4.651 |
| $\beta_1/\beta_5$ | average level of health/ financial fairness | 1.184 |
| $\beta_2/\beta_3$ | distribution of health / average level of process outcome | 4.677 |
| $\beta_2/\beta_4$ | distribution of health / distribution of process outcome | 5.482 |
| $\beta_2/\beta_5$ | distribution of health / financial fairness | 1.396 |
| $\beta_3/\beta_4$ | average level of process outcome/ distribution of process outcome | 1.172 |
| $\beta_3/\beta_5$ | average level of process outcome/ financial fairness | 0.298 |
| $\beta_4/\beta_5$ | distribution of process outcome/ financial fairness | 0.255 |

**Table 2.4** Comparison of goal valuations

| | WHO 1st survey[a] | WHR applied weights[a] | WHO 2nd survey (Netherlands) (face-to-face/ postal)[b] | Discrete choice experiment valuations |
|---|---|---|---|---|
| Average level of health (1) | 0.24 | 0.25 | | 0.29 |
| Distribution of health (2) | 0.25 | 0.25 | | 0.34 |
| Average level of process outcome (3) | 0.13 | 0.125 | | 0.07 |
| Distribution of process outcome (4) | 0.16 | 0.125 | | 0.06 |
| Financial fairness (5) | 0.22 | 0.25 | | 0.24 |
| Health outcomes (1+2) | 0.49 | 0.5 | 0.415 | 0.63 |
| Process outcomes (3+4) | 0.29 | 0.25 | 0.309/ 0.306 | 0.13 |
| Financial fairness (5) | 0.22 | 0.25 | 0.277/ 0.279 | 0.24 |
| Quality (1+3) | 0.37 | 0.375 | 0.376/ 0.432 | 0.36 |
| Equity (2+4+5) | 0.63 | 0.625 | 0.624/ 0.569 | 0.64 |

[a] *Gakidou et al. 2000*[58]
[b] *Gakidou et al. 2003*[54]

## Discussion

The focus of our research was twofold. First, we identified five health system goals and explored their theoretical foundation. We included the process of health care delivery in a utility framework supporting that the process of care giving is an end goal of health systems, irrespectively whether it affects health, and thus can be traded-off against health outcomes. Second, we obtained valuations for the goals using a multi-attribute technique based on trade-offs. We used actual variation in goal attainment, elicited marginal weights from behind a 'veil of ignorance' reflecting the original position, and created a setting to help respondents understand the concepts and their task. Consequently, we tested all five goals statistically significantly affect choice behaviour independently. An equitable health distribution has the highest valuation, followed by average level of health and financial fairness. Both attributes measuring process utility receive much lower weights.

By assuming a linear additive model for a latent preference variable, it was possible to compare our valuations with WHO's weights. Some might argue that our DCE valuations only provide a rough estimate of actual percentages. Nevertheless, our results are quite different from WHO's weights. Specifically, much weight of both process attributes shifts toward both health attributes. The shift may be due to the fact that we specifically made respondents aware that the intrinsic contribution of process attributes do not directly influence health. The extent to which WHO's respondents were conscious of this distinction is unknown.

Several researchers have claimed that people are willing to sacrifice overall health to achieve a more equitable distribution of health[45,47,62,72] Our results show this greater weight for health inequality whereas WHO's weights are equal for both health attributes. First, an equitable health distribution attains a higher valuation compared to average level of health. Second, the computed MRS suggests that individuals are willing to trade 2 years of average health for 1.7 years in health inequality. We believe that this might be attributable to our enhanced methodology to derive marginal valuations from behind a veil of ignorance.

The complexity of the attributes and levels forced us to make a trade-off between potential interviewer bias caused by an interactive classroom setting and task simplicity. Furthermore, our respondents were well-educated and familiar with the Dutch health system. It is possible that we introduced selection bias and the goal valuations are not representative of the preferences of the Dutch population. Gakidou et al.,[54] however, conclude that WHO's valuations varied only slightly between informed respondents and

the population at large. We also obtained goal valuations from behind a veil of ignorance, although whether such a scheme guarantees societal preferences that are genuinely impartial is debatable. Two findings in favour of impartiality was that respondents' equity concerns were altered by the 'original position' due to 'the veil of ignorance' and choice behaviour in the pilot study changed when respondents assumed an unknown position in society.

At the methodological level, economists are inclined to consider DCEs superior to ranking and rating since they are based on the random utility theory. Therefore, we argue that the research resulting from our enhanced methodology can be seen as a follow-up to the WHO surveys. We do, however, acknowledge methodological issues of DCEs, most of which are related to human cognitive processes. For example, choice experiments assume that people have stable preferences and are willing to trade between all attributes and a 'veil of experience' can influence decision making through status quo bias and the endowment effect.[73] However, these issues have received much attention within health economics and should not be seen as 'threats' to economic methods of valuation.[74]

## Conclusions

Our study demonstrates the feasibility of measuring health system goal valuations using a multi-attribute technique based on direct trade-off questions. We believe that our valuations improved on WHO's derived weights because we applied a comprehensive enhanced methodology, used actual variation in goal attainment, elicited marginal weights from behind a veil of ignorance, and created a setting to help respondents understand the concepts and their task. And because new weights could affect countries' rankings and comparisons over time, we advocate that appropriate weights be applied in future (international) comparisons of health system performance. Our study provides a promising and challenging basis on which to improve the methodology of global health system performance measurement. Advancing performance measurement is essential to cross-country comparison, which, in turn enhances global evidence-based health policy information.

## Acknowledgments

This chapter has benefitted from the questions and suggestions of participants of a WHO-headquarters seminar. We also thank Marc Koopmanschap for providing valuable feedback on the manuscript.

# B

## Drug reimbursement procedures, processes and criteria

# 3

# Similarities and differences between five European drug reimbursement systems

With Maïté le Polain, Irina Cleemput, and Marc Koopmanschap

International Journal of Technology Assessment in Health Care 2012; 28(4):349-357

## Abstract

The aim of our study is to compare five European drug reimbursement systems, describe similarities and differences, and obtain insight into their strengths and weaknesses and formulate policy recommendations.

We used the analytical Hutton Framework to assess in detail drug reimbursement systems in Austria, Belgium, France, The Netherlands and Sweden. We investigated policy documents, explored literature and conducted 57 interviews with relevant stakeholders.

All systems aim to balance three main objectives: system sustainability, equity and quality of care. System impact, however, is mainly assessed by drug expenditure. A national reimbursement agency evaluates reimbursement requests on a case-by-case basis. The minister has discretionary power to alter the reimbursement advice in Belgium, France and The Netherlands. All systems make efforts to increase transparency in the decision-making process but none uses formal hierarchical reimbursement criteria nor applies a cost-effectiveness threshold value. Policies to deal with uncertainty vary: financial risk-sharing by price/volume contracts (France, Belgium) versus coverage with evidence development (Sweden, The Netherlands). Although case-by-case revisions are embedded in some systems for specific groups of drugs, systematic (group) revisions are limited.

As shared strengths, all systems have clear objectives reflected in reimbursement criteria and all are prepared to pay for drugs with sufficient added value. However, all systems could improve the transparency of the decision-making process; especially appraisal lacks transparency. Systems could increase the use of (systematic) revisions and could make better use of HTA (amongst others cost-effectiveness ) to obtain value for money and ensure system sustainability.

## Introduction

The sustainability of drug reimbursement systems is increasingly under pressure by continuously rising health care expenditures. A detailed comparison of various European drug reimbursement systems provides an overview of systems' similarities and differences, and could help identify systems' strengths and weaknesses and thus provide opportunities to improve their efficiency and sustainability.



Previous studies have investigated the use of health technology assessment (HTA) in coverage decision making,[75-78] specific drug policies,[79] or parts of drug reimbursement systems[80-83] such as pricing and reimbursement,[76,84] stakeholder involvement,[85] and the role of reimbursement criteria.[81,86-88] However, based on a literature review, Vuorenkoski et al.[13] concluded that most studies are descriptive in nature. They suggest that more analytically oriented studies would enhance our understanding of how reimbursement decision-making processes perform against system objectives. Therefore, we compare five European drug reimbursement systems, providing a detailed and comprehensive comparative analysis between the systems' objectives, institutions, processes, formal reimbursement criteria, and output and implementation in real life. In specific, the degree of detail in our analysis, the link with policy goals, and the breadth of our investigations improves upon previous research. We draw general conclusions with respect to systems' similarities and differences, strengths and weaknesses and formulate policy recommendations.

## Methods

We used the analytical Hutton Framework[89] to describe, analyse and compare the Belgian, Austrian, Dutch, French and Swedish drug reimbursement systems. Although the country selection was partly arbitrary aiming to include our own countries, the selection of the other countries was based on observed important differences in systems' structure, organisation and procedures. Our sample includes systems with (i) various historical contextual backgrounds such as having a Beveridge-type (Sweden), Bismarck-type (Austria, Belgium, France, and The Netherlands), and managed competitive (The Netherlands) system; (ii) various types of final decision makers, i.e., the reimbursement agency (Austria and Sweden) or minister of health (Belgium, France and The Netherlands); and (iii) various implementation levels (national in Austria, Belgium, France and The Netherlands and regional in Sweden).

We investigated policy documents, explored literature and conducted interviews. Experts in each country validated individual country reports. The aim of the interviews was to retrieve (up-to-date) information unavailable in policy documents and literature, and to obtain further insight into how the systems work in practice. The selection of interviewees was based on their specific involvement in drug reimbursement. Interviewees were policymakers, representatives of the reimbursement agency/social insurance institution, expert committee members, patients, or representatives of the pharmaceutical industry. Interviews were performed by mail questionnaire (1), phone (2), or face-to-face (34), totalling 57 persons (3, 24, 5, 14 and 11 in Austria, Belgium, France, The Netherlands and Sweden, respectively). The number of interviewees was deliberately higher in our own countries in which we started; because of time restrictions, but mainly due to learning effects we could reduce the number of interviewees in the subsequent countries.

The descriptive Hutton framework provides a structure that comprehensively details reimbursement systems (including drug reimbursement systems), distinguishing between policy implementation and technology decision levels.[89] The policy implementation level describes how the system is embedded in the broader political system. It encompasses the (legal) establishment, objectives, implementation, and accountability of the system. The technology decision level describes the process of an individual reimbursement request and its phases: assessment, decision making, and outputs and implementation. Based on the framework, information on the characteristics of reimbursement systems can be grouped into a four-area research matrix: constitution and governance, methods and processes, use of evidence, and accountability and transparency (see Table 3.1).

**Table 3.1** Elements of the Hutton framework

| | **Elements of the system** | | | |
|---|---|---|---|---|
| **Policy implementation level** | Establishment | Objectives | Implementation | Accountability |
| **Technology decision level** | *Constitution and governance* | *Methods and processes* | *Use of evidence* | *Transparency, accountability* |
| a) Assessment | Consultation and involvement of stakeholders | Methodology | Evidence-base for assessment | Presentation and communication of assessment results |
| b) Decision | Who makes the decision | Decision-making process | Evidence-base and additional influences | Content and documentation of the decision |
| c) Outputs and implementation | Appeal and dissent | Implementation and communication | Monitoring and reappraisal | Evidence of the impact of the decision |

*Source: Hutton et al.*[12]

We added the concept of appraisal at the technology decision level. ***Assessment*** is the quantification of the clinical, pharmacotherapeutic and pharmacoeconomic value of a drug. It is descriptive in terms of quality and uncertainty of evidence. ***Appraisal*** seeks to gauge society's willingness to pay for a drug by weighing assessment outcomes against other (societal) criteria which reflect health system objectives. ***Decision making*** is a value judgement from a broader societal perspective, considering health system objectives as well as non-health care related objectives.



## Results

### Contextual background

All five countries have health care systems that cover more than 99 percent of their populations. The Swedish system originates from a Beveridge-type national health system; the other four originate from a Bismarck-type social insurance system. The Dutch system uses managed competition between providers and insurers. Health policy is mainly developed and regulated at the national level, but implementation and financial responsibility can be regional or rely on external actors (e.g., insurers).

Based on OECD 2008 figures, health care expenditure varies from 9.4 to 9.9, 10.2, 10.5 and 11.2 percent of GDP in Sweden, The Netherlands, Belgium, Austria and France, respectively.[90] A larger variation is observed in pharmaceutical expenditure as a share of total health care expenditure: 11.0, 13.2, 13.3, 16.4, and 16.4 percent in The Netherlands, Sweden, Austria, Belgium, and France respectively.[90]

The countries share similar system objectives: system sustainability, equity, and quality of care. Countries can make different trade-offs to balance the objectives and obtain a socially acceptable equilibrium. All have an open-ended pharmaceutical budget moderated by annual goals. Although pricing policies are not within the scope of this study, pricing and reimbursement are often strongly linked. All five countries use budget control mechanisms and supply/demand-side tools such as price regulations, international price referencing, internal reference pricing, financial risk-sharing agreements, (incentivised) prescription guidelines, and co-payments. The final price or reimbursement basis at least partially depends on the drug reimbursement evaluation.

### Policy implementation level

All five drug reimbursement systems explicitly seek equitable and affordable access to high quality health care in a sustainable manner. Other shared objectives are transparency towards pharmaceutical companies and rewarding innovation and investments in

research and development (R&D). None of the systems is clear about the actual place of such 'non-health' objectives. In the past decade, all countries have reformed their reimbursement systems' legal basis. The reforms aimed to improve efficient decision making in the context of increasing health care expenditure and were partly triggered by the EU Transparency Directive 89/105/EEC, requiring transparency of the decision-making process.

Except for expensive inpatient drugs in The Netherlands and ad-hoc procedures initiated by the reimbursement agency in Austria, the reimbursement process for a new product is initiated by the manufacturer. In all countries, outpatient drugs need to be assessed and enlisted to be eligible for reimbursement. Systems for inpatient drugs vary: they are part of the drug reimbursement system in Belgium, France and The Netherlands (expensive drugs only), the responsibility of county councils in Sweden, and hospitals, Länder, communities, and other hospital owners in Austria.

A shared characteristic is the existence of a national reimbursement agency: HVB in Austria, INAMI/RIZIV in Belgium, HAS in France, CVZ in The Netherlands and TLV in Sweden. In all countries, a technical department is responsible for compiling scientific evidence. The department prepares the assessment and drafts the preliminary summary report. An independent expert committee assesses and appraises the evidence and is responsible for advising the final decision maker (i.e., the minister of health in Belgium, France and The Netherlands, HVB in Austria). In Sweden, the expert committee also makes the final decision. Expert committees are considered independent because members, who must disclose conflicts of interest, are appointed for their scientific skills and expertise as representatives of society's prevailing interest. Only The Netherlands has, besides the expert committee (CFH), a separate appraisal (ACP) committee that also advises, based on societal considerations, the final decision maker. A closer look at the composition of the expert committees reveals divergences (see Table 3.2). Belgium has the largest expert committee (31 members, 23 of which have voting rights); Sweden has the smallest.

We distinguished two main differences in the composition of the committees. The Belgian and Austrian committees represent all relevant stakeholders. Sweden, The Netherlands, and France rely heavily on academic and other scientific experts. Stakeholders can be consulted but are not entitled to deliberate or vote. In 2010, Sweden reduced the number of committee members, replacing scientific experts with health care planning experts. The Belgian and French committees include consultants from the pharmaceutical industry without voting rights. The Austrian committee has two representatives of employees and consumers. The Dutch appraisal committee and the Swedish expert committee have a patient representative. The reimbursement advice (or decision) is

**Table 3.2** Composition of the expert committees

| | Austria | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|---|
| **Expert committee** | HEK | CRM/CTG | CT | CFH (ACP) | TLV Expert Board |
| **Voting members** | 20<br>- 3 academics<br>- 10 sickness funds<br>- 2 physicians<br>- 1 pharmacist<br>- 2 employees/ consumers<br>- 2 pharmaceutical industry | 23<br>- 1 chairperson<br>- 7 academics<br>- 8 sickness funds<br>- 4 physicians<br>- 3 pharmacists | 20<br>- 1 chairperson (from HAS)<br>- 19 members with medical or pharmacological expertise | max 24 (9)<br>CFH:<br>- 1 chairperson (from CVZ)<br>- members have expertise in pharmacological, medical, health sciences and economics<br><br>ACP:<br>- 3 CVZ (board of directors)<br>- 6 members with societal expertise (e.g., patient, ethicist, economist) | 7<br>- 1 chairperson (from TLV)<br>- 1 pharmacologist<br>- 1 (health) economist<br>- 1 patient<br>- 3 health care planners |
| **Permanent consultative members** | 1<br>- federal government | 8<br>- 4 ministries<br>- 3 pharmaceutical industry<br>- 2 INAMI/ RIZIV | 8<br>- 4 public institutions<br>- 1 pharmaceutical industry<br>- 3 sickness funds | 2<br>- ministerial observers | n/a |

*ACP = Appraisal committee (Advies Commissie Pakket);*
*CFH = Expert Pharmaceutical Advisory Committee (Commissie Farmaceutische Hulp);*
*CRM/ CTG = Drug Reimbursement Committee (Commission de Remboursement des Médicaments/ Commissie voor Tegemoetkoming Geneesmiddelen);*
*CT = Transparency Committee (Commission de la Transparence);*
*CVZ = Health Care Insurance Board (College voor Zorgverzekeringen);*
*HAS = Natianal Authority for Health (Haute Autoritée Santé);*
*HEK = Pharmaceutical Evaluation Board (Heilmittel-Evaluierungskommission);*
*INAMI/ RIZIV =National Institute for Health and Disability Insurance (Institut National d'Assurance Maladie-Invalidité);*
*TLV = Dental and Pharmaceutical Benefits Agency (Tandvårds- och Läkemedelsförmånsverket)*

based on majority voting or, in The Netherlands, consensus. Belgium is unique in that a two-thirds majority is required; without it no advice is formulated.

In all countries, the minister of health is responsible for defining the overall drug reimbursement policy and steering the system; the ministry is accountable to Parliament. All systems have a trend towards increasing the transparency of decision making, but all systematically assess the impact of the system by monitoring drug expenditure rather than other system objectives.

## Technology decision level

Table 3.3 provides a summary of our findings regarding the technology decision level for individual drug reimbursement requests.

*Assessment and appraisal*

The authorities responsible for the final reimbursement decision rely on advice from the expert committees. Reimbursement advice results from the often-intertwined processes of assessment and appraisal. The technical department starts the assessment and informs the expert committee, which appraises the reimbursement request and advises the final decision body. Even though The Netherlands has separate assessment and appraisal committees, the processes are still intertwined.

*Therapeutic value*

A common key characteristic is the evaluation of the therapeutic value. All interviewees acknowledged that efficacy, effectiveness, safety, and adverse effects were the most important formal criteria. Although the criteria are related, they are subject to different interpretations and have various outcomes. In Austria and France, therapeutic value is rated in categories. Austria applies six categories ranging from 'no added benefit' to 'important benefit for the majority of patients'. France distinguishes five levels of improvement in the medical service rendered (ASMR) ranging from 'no improvement' to 'major

**Table 3.3** Technology decision level

| | Austria | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|---|
| ***Assessment*** | | | | | |
| Main actor(s) | | | | | |
| Preparation, processing & reporting | HVB | INAMI/RIZIV | HAS | CVZ | TLV |
| Expert committee | HEK | CRM/CTG | CT | CFH | TLV Expert Board |
| Assessment criteria | | | | | |
| - Efficacy | Yes | Yes | Yes | Yes | Yes |
| - Effectiveness | Yes | Yes | Yes | Yes | Yes |
| - Safety & adverse effects | Yes | Yes | Yes | Yes | Yes |
| - Ease of use/comfort | Yes | Yes | Yes | Yes | Yes |
| - Added therapeutic value | Yes | Yes | Yes‡ | Yes | Yes |
| - Cost-effectiveness | Yes* | Yes | No (new drugs†) | Yes** | Yes |
| - Other(s): | Extensive list of criteria | Therapeutic and social needs | Public health, treatment properties, compliance | Applicability, feasibility, experience | All effects on a person's health and quality of life |

**Table 3.3** Technology decision level (continued)

| | Austria | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|---|
| ***Appraisal*** | | | | | |
| Main actor | HEK | CRM/CTG | CT | ACP (CVZ + CFH) | TLV Expert Board |
| Explicit appraisal criteria | Yes | Yes | Yes | Yes | Yes |
| Appraisal criteria | All assessment criteria judged in the light of system's objectives. | Added therapeutic value, clinical effectiveness, budget impact, cost-effectiveness and price/ reimbursement basis | SMR[x] criteria: efficacy, adverse effects, place of the drug with regard to alternatives, disease severity, treatment properties, public health benefit | Added therapeutic value, cost-effectiveness, medical need, disease severity, rarity, public health, accessibility, own responsibility, societal affordability | Human value, need and solidarity, and cost-effectiveness |
| Threshold (range) for cost/QALY | No | No | No | No | No |
| Expert committee report publicly available | No | Yes | Yes | Yes | Yes<br>No if applicant withdraws request |
| Expert committee advice binding | No | No | No | No | Yes |
| ***Decision*** | | | | | |
| Decision-making body | HVB | Minister | Minister | Minister | TLV |
| Discretionary power final decision maker | Yes, deviation rarely occurs | Yes, deviation sometimes occurs | Yes, deviation rarely occurs | Yes, deviation rarely occurs | n/a |
| Stakeholders involvement | No | Yes | Yes | Yes | n/a |
| Motivation publicly available | Yes | Yes | Yes | Yes | Yes |
| Reimbursement restrictions (e.g., specific indications) | Yes (Yellow box) | Yes (Chapter IV) | Yes | Yes (Annex 2) | Yes |
| Temporary decision | No | Yes (Class 1) | Yes (all drugs) | Outpatient: No<br>Expensive inpatient: Yes | Yes (case-by-case) |
| Risk sharing agreements | No | Yes, financial based (Class 1 with negative/ no proposal) | Yes, financial based (price-volume agreements) | No | No |

**Table 3.3** Technology decision level (continued)

| | Austria | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|---|
| ***Outputs and implementation*** | | | | | |
| Appeal and dissent | Yes | Yes | Yes | Yes | Yes |
| - Grounds for appeal | Procedural and substantive grounds | Procedural grounds | Procedural grounds | Procedural grounds | Procedural grounds |
| - Initiator | Applicant | Any stakeholder | Any stakeholder | Any stakeholder | Applicant |
| - Appeal options | UHK | State Council | State Council | Expert Review + Administrative Court | Administrative Court |
| Implementation | | | | | |
| - Mechanisms | National drug formulary | National drug formulary | National drug formulary | National drug formulary; Pharmaco-therapeutic groups | County councils & Drug Therapeutic Committees |
| - Local variations | No | No | No | No | Yes |
| Revisions | | | | | |
| - Ad hoc | Yes | Yes | Yes | Yes | Yes |
| - Systematic | No | Yes (Class 1) | Yes (all drugs every 5 years) | Outpatient: No Expensive inpatient: Yes | Yes (drugs enlisted < 2002) |
| - Consequences revisions | Changes in conditions, delisting | Changes reimbursement modality; delisting (rarely) | Delisting | Outpatient: delisting (rarely) Inpatient: awaiting | Delisting |
| Impact assessment | Drug expenditure | Drug expenditure | Drug expenditure | Drug expenditure | Drug expenditure |

** quality and uncertainty of evidence ** robustness of evidence*

*[†] New law (Article 14; Law No 2011-2012 December 29th, 2011): drug reimbursement applications are assessed relative to therapeutic strategies, where available, under conditions defined by decree in Conseil d'Etat –Council of State– (conditions not yet published at time of publication)*

*[¥] HAS is currently drafting a proposal to replace the SMR and ASMR to one single criterion (Relative Therapeutic Benefit)*

*[†] HAS recently received an extended remit to assess methodological quality of economic assessments of new technologies (decree under review Conseil d'Etat –Council of State–)*

*ACP = Appraisal committee (Advies Commissie Pakket); CFH = Expert Pharmaceutical Advisory Committee (Commissie Farmaceutische Hulp);CRM/ CTG = Drug Reimbursement Committee (Commission de Remboursement des Médicaments/ Commissie voor Tegemoetkoming Geneesmiddelen); CT = Transparency Committee (Commission de la Transparence); CVZ = Health Care Insurance Board (College voor Zorgverzekeringen); HAS = Natianal Authority for Health (Haute Autoritée Santé); HEK = Pharmaceutical Evaluation Board (Heilmittel-Evaluierungskommission); HVB = Main Association of Austrian Social Security Institutions (Hauptverband der Österreichischen Sozialverzicherungsträger); INAMI/ RIZIV =National Institute for Health and Disability Insurance (Institut National d'Assurance Maladie-Invalidité); TLV = Dental and Pharmaceutical Benefits Agency (Tandvårds- och Läkemedelsförmånsverket); UHK = Independent Pharmaceutical Commission (Unabhängige Heilmittelkommission)*

improvement'. It should be noted that the French agency is currently drafting a proposal to replace the SMR and ASMR to one single criterion (Relative Therapeutic Benefit). Sweden uses a sliding scale such that price depends on the drug's cost-effectiveness. In contrast, the added therapeutic value is a binary yes/no decision in Belgium (class 1 or 2) and The Netherlands (list 1A or 1B).

In all countries, only drugs with added therapeutic value can obtain a higher reimbursement basis; in France, the added value also determines the level of patient cost share. For drugs with similar therapeutic value, the implications vary. In France, such drugs are reimbursed only if they realise savings. Dutch therapeutically-equivalent drugs are grouped and reimbursed equally. In Belgium, the reimbursement basis equals that of the comparator. Austria assigns such drugs a lower consumer price than the best therapeutic and reimbursable alternative.



*Cost-effectiveness*

All countries but France use cost-effectiveness as formal criterion. Although the French agency is explicitly encouraged to use cost-effectiveness, the expert committee has until now been reluctant to take it into account for assessing new drugs. France does consider cost-effectiveness in revision processes. Recently, the French agency received an extended remit to assess methodological quality of economic assessments of new technologies (decree under review Council of State –Conseil d'Etat–). In Belgium and The Netherlands, cost-effectiveness is taken into account only for drugs with recognised added therapeutic value. In Sweden and Austria, cost-effectiveness evidence requirements are most extensive for drugs claiming added therapeutic value.

Further exploration reveals divergence in countries' assessment of cost-effectiveness. In Austria and The Netherlands, only the quality of evidence and its level of uncertainty are assessed. The Swedish and Belgian committees, in contrast, also consider the actual cost-effectiveness ratio.

Even though four countries have cost-effectiveness as a formal criterion, none applies a strictly defined or transparent cost-effectiveness threshold (range). Most interviewees indicated that if one existed, it would be an increasing threshold depending on factors such as disease severity and medical need. They also acknowledged being more lenient towards orphan drugs and drugs for severe and life-threatening diseases.

*Appraisal*

Appraisal criteria and the weighing process are far less transparent than assessment. Belgium uses five appraisal criteria: added therapeutic value, price, budget impact,

cost-effectiveness, and therapeutic importance in light of unmet medical and societal needs. Austria has an exhaustive list of assessment elements and uses system objectives as appraisal criteria. In France, the medical service rendered (SMR) evaluation includes the following criteria: level of efficacy relative to adverse effects, disease severity, treatment properties (preventive, curative, symptomatic), the drug's position in therapeutic strategy, and public health benefit. In The Netherlands, the appraisal committee has developed formal appraisal criteria such as medical need, disease severity and rarity, public health, accessibility, societal and patient affordability, and lifestyle. In Sweden, the three priority principles –human value, need and solidarity, and cost-effectiveness– set formal appraisal criteria. Sweden promotes a value for money system; budget impact is thus not a formal national level criterion.

All systems apply various reimbursement criteria without an explicit hierarchy. Although the appraisal criteria are often derived from system objectives, they remain somewhat implicit and are often not transparent.

*Decision*

All European countries are required to make a final reimbursement decision within 180 days (excluding clock stops). Austria, France and Belgium apply strict timelines for advice (90, 90, and 150 days, respectively) and the reimbursement decision (180 days). For The Netherlands and Belgium, expert committee members' limited time and limited technical staff were frequently mentioned as bottlenecks.

In Austria, Belgium, France, and The Netherlands, decision making occurs in two phases. First, the expert committee comes to reimbursement advice. Second, the minister of health (or in Austria the association of Austrian Social Security Institutions [HVB]) makes the final reimbursement decision based on the advice. The Swedish expert committee makes the final decision without an advice phase. Although the minister in Belgium, France and The Netherlands rarely deviates from the advice, in Austria and Sweden the minister has neither final decision right nor discretionary power with respect to individual reimbursement decisions.

All countries but Austria publish their reimbursement advice (decision) reports although their extensiveness varies by country. (Additional) appraisal criteria especially are often not transparent and the weighing process is often not documented.

Outcomes of the decision-making process are similar: reimbursement, no reimbursement, or conditional reimbursement. All countries can apply restrictions for specific indications, patient groups, access restrictions and the like. In Austria, drugs in the so-called

red box (i.e., newly launched drugs and drugs that have applied for reimbursement) can already be reimbursed on an individual basis conditional on an ex-ante approval of a sickness fund 'head physician' before the reimbursement decision has been made.

Use of temporary decisions varies by country. All positive decisions in France are re-assessed after five years. No decisions in Austria are temporary. Only decisions on drugs with recognised added therapeutic value in Belgium and expensive inpatient drugs in The Netherlands are temporary. The Swedish reimbursement agency decides temporary reimbursement on a case-by-case basis which is based on uncertainty of the evidence.



*Outputs and implementation*
Applicants have formal opportunities in all countries to express their point of view or disagreement during the reimbursement process. They are also entitled on procedural grounds to appeal to the final decision at an administrative court.

All countries have mechanisms to support implementation by disseminating scientific evidence and improving appropriate drug use by means of national drug formularies and prescription guidelines. Only in Sweden every county council has its own guidelines. Impact assessment is often restricted to monitoring prescription volumes or drug expenditure.

We found substantially diverging policies regarding revision of enlisted drugs. The Austrian system has no systematic policy-enforcing revision process. In contrast, France systematically revises all decisions every five years, potentially changing reimbursement level or drug price. Sweden currently evaluates all drugs from the old reimbursement scheme (listed before 2002) according to therapeutic class. So far, this has resulted in guideline changes as well as delistings. In Belgium, all innovative drugs are systematically revised after 18 to 36 months. Changes in the reimbursement conditions occur but drugs are rarely delisted. Since 2006, expensive inpatient drugs in The Netherlands are revised after four years. No revision has been finalised thus far, hence its consequences are not clear. Recently, the Dutch minister announced that temporary decision making will be extended to outpatient drugs.

## Discussion

We compared five European drug reimbursement systems and provided a detailed analysis of systems' similarities and differences to obtain insight into their strengths and weaknesses and to formulate policy recommendations.

### Systems' strengths

At the policy implementation level, all five countries have enforced a national system that evaluates the societal value of a drug and determines whether the drug is worth reimbursement. All systems share clear objectives: system sustainability, quality of care and equity. System performance is monitored in terms of pharmaceutical expenditure, addressing system sustainability.

At the technology decision level, all put forward formal criteria which reflect systems' objectives. HTA is used at some phase in the decision-making process to trade-off between the objectives. All systems are prepared to pay for drugs with sufficient added therapeutic value. Stakeholder involvement is ensured either through consultation or direct representation in the expert committee. Mechanisms to support implementation are used by means of guidelines and drug formularies.

### Systems' weaknesses

At the policy implementation level, none of the systems systematically evaluates its performance regarding the quality of care and equity objective. All countries have a so-called supply driven system: the process starts with a manufacturer's reimbursement request and proceeds on a case-by-case basis. In principle this might lead to 'pragmatic incrementalism',[91] risking a low degree of consistency across decisions. Furthermore, most systems make limited use of tools to systematically (re-) evaluate drugs' relative value for money throughout their life cycle.

At the technology decision level, assessment and appraisal are in practice often strongly intertwined. All systems seem to use similar reimbursement criteria. However, none of the systems applies a formal hierarchy and the actual role of each criterion in the decision-making process is often not transparent; especially appraisal criteria lack transparency. Although all countries recognise the importance of HTA, all experience difficulties in defining its role and weight in the decision-making process.

### Study limitations

Our study only includes five countries. Nevertheless, we observed important differences in structure, organisation and procedures. The degree of detail in our analysis, the link with policy goals, and the breadth of our investigation contribute to previous studies and show opportunities to improve system efficiency and sustainability. Our analysis did not study individual reimbursement cases; such a case series analysis is part of our current work. This could produce additional insights.

### Implications for policy and recommendations

To increase legitimacy of societal decision making, all systems could improve transparency, especially the use of appraisal criteria and their role in the decision-making process. Assessment and appraisal could be better disentangled. We believe it would be possible to develop standard European guidelines for the assessment of clinical, pharmacotherapeutic and pharmacoeconomic evidence; especially because countries already keep track of the evaluation in other countries which most likely influences their own evaluation of especially the clinical evidence. EUnetHTA has been exploring such activities for relative effectiveness assessment, though not for pharmacoeconomic evaluations.[92] On the other hand, appraisal should remain country-specific because social values might vary across countries. Having the final reimbursement decision in the hands of the Ministry of Health (Belgium, France, and The Netherlands), might reflect central governments' wish to keep discretionary power.

Drug reimbursement decisions are inevitably made under uncertainty. Although tools to reduce consequences of uncertainty, such as (financially-based) risk-sharing schemes and temporary reimbursements have been introduced and seem to gain more attention, not all systems are currently sufficiently equipped to systematically deal with uncertainty. Results from risk-sharing agreements in France are promising,[93,94] as well as results from systematic revisions in France[93,95] and Sweden.[96]. After reimbursement, evidence development using outcomes research and patient registries could improve monitoring real-world outcomes.[97,98] Full package revisions might improve consistency of decision making over time and enhance overall value for money and thus ensure sustainability. Furthermore, countries could make better use of HTA to obtain value for money. HTA could play a more prominent role to systematically assess and determine the level of added societal value and set the price or reimbursement level accordingly.

Currently, the countries only evaluate performance of the system regarding sustainability. We recommend developing tools to assess the impact of drug reimbursement on the other two objectives: quality of care and equity. Finally, policymakers could reconsider the current supply-driven system; they could also consider shifting towards a more demand-oriented system in which they state for which new drugs addressing unmet medical needs they are willing to pay.

## Acknowledgements

We would like to thank Gérard De Pouvourville, John Hutton and Willy Palm for their helpful comments on a preliminary version of our project report. We are also grateful to

Frans Rutten, Ton de Boer, Adri Steenhoek, Elly Stolk, Erik Schokkaert and Carine Van de Voorde for their advice during our research project.

This study was partly performed in the context of the Escher project (T6-202), a project of Top Institute Pharma, Leiden, The Netherlands.

# 4

# European drug reimbursement systems' legitimacy: Five country comparison and policy tool

With Irina Cleemput, Maïté le Polain, and Marc Koopmanschap

International Journal of Technology Assessment in Health Care 2012; 28(4):358-366

## Abstract

In a democratic system, decision makers are accountable for the reasonableness of their decisions. This presumes (i) transparency, (ii) relevance of the decision criteria, (iii) revisability of decisions, and (iv) enforcement/regulation. We aim to (i) evaluate the extent to which drug reimbursement decision-making processes in different contexts meet these conditions and (ii) develop, starting from these findings, a framework for improving the transparency and the relevance of used decision criteria.

We evaluated the Austrian, Belgian, French, Dutch and Swedish drug reimbursement systems. Based on this evaluation, we developed a framework for improving the transparency of drug reimbursement decision-making processes. It makes explicit the questions often addressed implicitly during decision-making processes as well as criteria for answering each question.

Transparency of appraisal processes varies across systems. Justification with explicit criteria is generally limited. Although relevant criteria are similar across systems, their operationalization varies and their role in the appraisal process is not always clear. All systems seem to implicitly address five key questions, relating to (i) the medical, therapeutic and societal need for treatment; (ii) preparedness to pay for treating the condition as a principle; (iii) for using the treatment under consideration, (iv) preparedness to pay more compared with alternatives; and (v) actual willingness to pay from public resources.

Transparency of the appraisal process can be improved by using an explicit decision framework. Systematic use of such a framework enhances consistency across decisions, allows justification of value judgments, and thus enhances legitimacy of societal decision making.

## Introduction

In the context of continuously increasing public expenditure on pharmaceuticals, the efficiency and sustainability of drug reimbursement policies become increasingly important. While the health systems that shape such policies have the same primary objective (to increase and maintain health), constraints force policymakers to make choices toward system sustainability.[99] Despite variations in the organisation and financing of health care between member states, the European Commission defined three common health system objectives: equity and accessibility, quality of care, and sustainability.[100] The US Department of Health and Human Services highlighted similar objectives in its strategic plan 2010-2015.[101] Competition between these objectives often forces policymakers to make trade-offs. These trade-offs are primarily a matter of normative choice: countries will aim for a socially acceptable equilibrium between the different objectives. Judging drug reimbursement systems on this outcome is difficult but we can argue that a legitimate policy-making process that facilitates decisions in line with public values would optimally serve the stated objectives.



Key criteria for legitimacy or accountability for reasonableness according to Daniels and Sabin, are (i) transparency of the decision-making process, (ii) relevance of the decision criteria, (iii) revisability of decisions in light of new evidence and arguments, and (iv) enforcement/regulation of the previous criteria.[102] This study evaluates to what extent these criteria are fulfilled in five European drug reimbursement systems with a different organisational and procedural context. The findings of this evaluation prompted for the development of a framework for improving the transparency of drug reimbursement decision-making processes and the relevance of drug reimbursement criteria. This study reports on both the evaluation and the decision framework.

## Methods

We assessed legitimacy of drug reimbursement decision making in five European countries as follows. First, we performed an in-depth analysis of five different European drug reimbursement systems using the analytical Hutton Framework.[12] Table 4.1 characteristics of the five systems relevant for this study. Detailed methods and results of the analysis of the performance of these five systems are presented elsewhere.[103] Using data triangulation, we investigated policy documents publicly available in English, French, German and Dutch at the Websites of the reimbursement agencies, explored (gray) literature and other relevant publications obtained by means of Medline and Cochrane Library searches and provided by our interviewees, and conducted interviews.

**Table 4.1** Summary characteristics of five national reimbursement systems

| | Austria | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|---|
| ***Regulation/ Enforcement*** | | | | | |
| National reimbursement agency | HVB | INAMI/ RIZIV | HAS | CVZ | TLV |
| Expert advisory committee | HEK | CRM/CTG | CT | CFH (and ACP) | TLV Expert Board |
| Scope of national agency | Outpatient drugs | Inpatient and outpatient drugs | Inpatient and outpatient drugs | Expensive inpatient and outpatient drugs | Outpatient drugs |
| Final decision maker | HVB | Minister of Health | Minister of Health | Minister of Health | TLV |
| Implementation of the outcome | Positive list | Positive list | Positive list | Positive list | Positive list |
| Pharmaceutical budget | Open-ended | Open-ended | Open-ended | Open-ended | Open-ended |
| Monitoring outcomes | Drug expenditure | Drug expenditure | Drug expenditure | Drug expenditure | Drug expenditure |
| Appeal options | Content and procedural grounds | Procedural grounds | Procedural grounds | Procedural grounds | Procedural grounds |
| ***Transparency*** | | | | | |
| Reimbursement reports publicly available | No | Yes | Yes | Yes | Yes (not if case is withdrawn) |
| ***Relevance*** | | | | | |
| Appraisal criteria (national level) | | | | | |
| Medical, therapeutic and societal need | Yes | Yes | Yes | Yes | Yes |
| Added therapeutic value | Yes | Yes | Yes | Yes | Yes |
| Cost-effectiveness | Yes | Yes | No for first decision<br>Yes for revision | Yes | Yes |
| Budget impact | Yes | Yes | Yes | Yes | No |
| Threshold (range) for cost/ QALY | No | No | No | No | No |
| ***Revision*** | | | | | |
| Ad hoc revision | Yes | Yes | Yes | Yes | Yes |
| Systematic revision | No | Yes, all drugs with recognised added value | Yes, all drugs every 5 years | Yes, only expensive inpatient drugs | Case by case & all drugs enlisted before 2002 |

*HVB= Main Association of Austrian Social Security Institutions [in German: Hauptverband der Österreichischen Sozialversicherungsträger]; HEK= Pharmaceutical Evaluation Board [in German: HeilmittelEvaluierungsKommission]; INAMI/ RIZIV = National Institute for Health and Disability Insurance [in French: Institut National d'Assurance Maladie-Invalidité; in Dutch: Rijksinstituut voor Ziekte- en Invaliditeitsverzekering]; CRM/ CTG = Drug Reimbursement Committee [in French: Commission de Remboursement des Médicaments; in Dutch: Commissie voor Tegemoetkoming Geneesmiddelen]; HAS = National Authority for Health [in French: Haute Autorité de Santé]; CT = Transparency Committee [in French: Commission de la Transparence]; CVZ = Health Care Insurance Board [in Dutch: College voor Zorgverzekeringen]; CFH = Expert Reimbursement Advisory Committee [in Dutch: Commissie Farmaceutische Hulp]; ACP = Appraisal committee [in Dutch: Advies Commissie Pakket]; TLV = Dental and Pharmaceutical Benefits Agency [in Swedish: Tandvårds- och läkemedelsförmånsverket]*

Interviewees were selected based on their involvement in the drug reimbursement procedure; they were policymakers from different organisations (n = 48), a patient representative (n = 1), or representatives of the pharmaceutical industry (n=8). Interviews were performed by mail questionnaire (1), phone (2), or face-to-face (34), totalling fifty-seven 57 persons (3, 24, 5, 14, 11 in Austria, Belgium, France, The Netherlands and Sweden, respectively). The number of interviewees was deliberately higher in our own countries in which we started and in Sweden. Although the Swedish Website provides a great amount of information in English, we needed to ensure complete data on the Swedish system not limited by language restrictions, because of time restrictions, but mainly due to learning effects we could reduce the numbers of interviewees in the subsequent countries. The aim of each interview was to retrieve (up-to-date) information unavailable in policy documents and literature and to obtain further insight into how the systems work in practice. Experts in each country validated all our individual country reports. For this analysis, we selected five European countries: Austria, Belgium, France, The Netherlands and Sweden. Although this sample size is relatively small, we performed a detailed analysis requiring an intensive search for formal as well as informal information. In these five countries we observed important differences in structure, organisation and procedures of the systems. Our sample includes (i) health care systems with various historical contextual backgrounds: Beveridge-type (Sweden), Bismarck-type (Austria, Belgium, France, and The Netherlands), and managed competitive (The Netherlands) systems; (ii) various types of final decision makers: the reimbursement agency (Austria and Sweden) and minister of health (Belgium, France, and The Netherlands); and (iii) various implementation levels: national (Austria, Belgium, France, and The Netherlands) and regional (Sweden).

Second, based on the findings of the in-depth analysis, we evaluated the five systems' organisation, structure, and procedures against the framework for accountability for reasonableness of Daniels and Sabin.[102] This ethical-theoretical framework defines four conditions for achieving legitimate and fair coverage decisions for new treatments. The four legitimacy conditions are: (i) Transparency of the decision process: the process must be fully transparent about the grounds for/rationales behind a decision; (ii) Relevance of the decision criteria: the decision must rest on reasons that all stakeholders can accept as relevant to meeting health needs fairly given the resource constraints; (iii) Revisability of decisions: decisions should be revisable in light of new evidence and arguments; and (iv) Enforcement/regulation: there must be some kind of regulation guaranteeing the previous three conditions.

Although this framework has been criticized,[104,105] empirical evidence suggests that priority-setting processes that fulfil the conditions for accountability for reasonableness

are perceived as being legitimate and fair.[101,106-108] Without making any value judgements, we evaluated each country's achievement regarding these legitimacy conditions. There is a conceptual distinction between assessment, appraisal and decision.[78] Our evaluation mainly focuses on appraisal.

Third, based on our legitimacy evaluation, we developed a policy tool that can improve transparency and relevance of the drug reimbursement decision-making process in all countries. We unravelled the decision-making process in smaller pieces and identified questions that all systems seem to address to a certain extent, more or less explicitly. After that, we assigned appraisal criteria currently used either explicitly or implicitly to each of the defined questions. This process led to a five-question decision framework, including a set of relevant criteria for each question. Our developed framework provides a tool to structure the decision process, allows reconstruction of the decision process, can improve consistency across decisions, and provides a tool to increase transparency of the appraisal process. Finally, to illustrate the application of our framework, we described how each country addresses the questions and uses the criteria of our framework.

## Results

### Evaluation of the four conditions for accountability for reasonableness

*Condition 1. Transparency*

Although all five systems seem to use similar criteria, the actual role of the criteria in the decision-making process is often not transparent. Assessment reports are usually made public, except in Austria, where evaluation reports (for outpatient drugs) are not published. However, the appraisal process, which leads to an advice or a decision, is rarely made public, although variations exist. The minutes of the French expert committee's meeting are published, including the main points of discussion, the voting results, and a motivated advice. The Belgian system publishes the initial assessment report, applicant responses, and the committee reactions to these responses, whereas the eventual (provisional) advice is withheld. Both countries conceal confidential information upon applicants' request if deemed justifiable by the expert committee. Dutch assessment and appraisal reports are available online and include main points of discussion. Appraisal committee meetings are open to the public. Sweden publishes the final reports online after deliberation with the manufacturer; confidentiality issues stated by the latter are concealed. Noteworthy is that in Sweden pharmaceutical companies can withdraw their case before the final reimbursement decision has been made, in which case no report is published, a guarantee of confidentiality at the cost of transparency.

*Condition 2. Relevance of the decision criteria and rationales*

Involvement of all stakeholders affected by a decision is thought to facilitate accountability for reasonableness because it increases the likelihood that the rationales adopted will be relevant and acceptable.[107,109] This presumes, though, that all stakeholders understand the decision problem and recognise the choices that have to be made to meet the different health care system objectives; that is, they must be aware that resources are limited and fair choices have to be made within such a resource-constrained context.[109] All systems ensure stakeholder involvement either through direct representation of stakeholders in the expert committee (Belgium and Austria) or through consultation of stakeholders by the expert committee in cases where this committee consists of scientific experts (Sweden, The Netherlands, and France). Only the Swedish expert committee has a patient representative as committee member.



*Condition 3. Revisability*

Revisability is most important in case of (high) uncertainty about the estimates of efficacy, effectiveness, cost-effectiveness, or budget impact, or if relevant evidence is still being developed. Austria is the only country that has no system of systematic revisions, although ad hoc revisions can be initiated. Belgium and The Netherlands have a revision procedure for specific drug classes, which can occur only once after the initial decision and within a window of 1.5 to 3 years (Belgium) or 4 years (The Netherlands). France revises all positive decisions every 5 years. In addition to an ongoing revision of all enlisted drugs before 2002, Sweden decides on a case-by-case basis whether a decision requires revision after a certain number of years. In all countries, depending on the reassessment results revisions can have consequences, such as delisting or a change in the level of reimbursement or the level of restrictiveness of the reimbursement condition. After revision, reimbursement conditions might become more or less restrictive than during the period of temporary reimbursement, depending on the reassessment results. In The Netherlands, the first revisions of expensive inpatient drugs that were conditionally reimbursed for a period of 4 years, are discussed now. Yet, their consequences are still unknown.

*Condition 4. Enforcement*

All countries legally instituted a designated national reimbursement agency. These agencies fall under ministerial responsibility and are audited or certified by external (parliamentary) committees. However, in all countries, little self-evaluation of the system is performed on the process and outcomes. (Parts of) reimbursement processes are monitored only on an ad hoc basis. The outcome is mainly monitored on pharmaceutical expenditure. All countries have formal appeal procedures for reimbursement decisions, although there is a variety in how and for what reasons appeal is possible. All countries

but Austria allow applicants to appeal against a decision on procedural grounds to an administrative court. In Austria, the Independent Pharmaceutical Commission acts as an appeal court for both procedural and content issues.

### Decision framework for the transparent use of relevant decision criteria

The results of our in-depth analysis of drug reimbursement systems showed that all countries use similar criteria in their decision-making process, including severity of disease, added therapeutic value, cost-effectiveness, budget impact, and uncertainty of evidence. However, systems lack transparency about how they deal with each of

**Table 4.2** Key Questions and relevant criteria for increasing the transparency of drug reimbursement appraisal processes

| Decision | Question | Relevant criteria |
|---|---|---|
| *Medical, therapeutic and/or societal need* | Does the product target a medical, therapeutic and/or societal need? | *Medical need:*<br>- Life-threatening condition<br>- Severe symptoms<br>*Therapeutic need:*<br>- Effective alternative treatment available<br>*Societal need:*<br>- High prevalence<br>- Disease leads to health inequalities<br>- Distance from an acceptable baseline health level |
| *Preparedness to pay for a particular indication* | Are we, as a society, prepared to use public resources to pay for a treatment to improve this particular indication? | - Personal responsibility<br>- Affordable out-of-pocket |
| *Preparedness to pay for a particular treatment* | Are we, as a society, prepared to use public resources to pay for this particular treatment, given that we are prepared to pay for a treatment to improve this indication? | - Safety and efficacy of the treatment compared to alternative treatment(s)<br>- Quality and uncertainty of the evidence regarding safety and efficacy<br>- Curative, symptomatic, or preventive<br>- Therapeutic value<br>- Significance of health gains |
| *Preparedness to pay more than an alternative* | Given that we are prepared to pay for this treatment using public resources, are we prepared to pay *more* than the best alternative treatment? | - Added therapeutic value<br>- Potential savings elsewhere<br>- Quality and uncertainty of the evidence regarding effectiveness<br>- Acceptability of co-payments<br>- Rarity of disease |
| *Willingness to pay: price and reimbursement basis* | How much more are we *willing* to pay out of public resources for this particular treatment? | - Added therapeutic value<br>- Incremental costs<br>- Budget impact / ability to pay<br>- Cost-effectiveness ratio<br>- Medical, therapeutic and societal need<br>- Limits to cost sharing<br>- Quality and uncertainty of evidence |

the criteria in their appraisal process and how their relative importance was judged. By unravelling the decision-making process, we were able to identify five key questions that all systems seem to address to some extent more or less explicitly. We assigned appraisal criteria currently used without much transparency or even implicitly to each of the defined questions. Table 4.2 displays our developed decision framework.

**Five key questions in decision making**

*Question 1: Is there a medical, therapeutic and/or societal need for this indication?*
A pharmaceutical is valuable in as far as it meets a specific need, be it medical, therapeutic, and/or societal.[110] The evaluation of 'need' in a specific disease is essentially relative, that is, compared with other indications that need treatment. Medical and therapeutic needs are functions of disease severity and treatment necessity, respectively. The more severe a disease and the less effective alternative treatments or the fewer the available alternatives, the higher the medical and therapeutic need.[111] Need also relates to societal objectives, such as reducing health inequalities. Medical, therapeutic, and societal need can collectively refer to the societal objective of equitably maximising health or well-being.[112-114]

In the literature, suggestions to operationalize need criteria have mostly been in terms of disease severity. Examples of approaches include 'fair innings',[115] 'severity of illness',[116] 'proportional shortfall',[117] and 'rule of rescue'.[118] By taking available treatment alternatives into account when determining disease severity (i.e., disease severity given current treatment options), medical and therapeutic need are addressed simultaneously. Measures to draw conclusions about societal needs, however, remain necessary.

All countries in our study have operationalized need during some phase of their decision-making process. Austria considers societal need when assessing added therapeutic value: drugs benefitting the majority of patients are classified higher in the added therapeutic value classification than those benefiting a subgroup. France defines need in a particular disease area relative to other needs in the health care sector through the assessment of the 'medical service rendered' (SMR), which is determined by disease severity, level of efficacy relative to adverse effects, the drug's place in therapeutic strategy (particularly with regard to treatment alternatives), treatment properties (preventive, curative, or symptomatic) and public health benefit. As such the SMR addresses medical, therapeutic and societal need. An insufficient SMR leads to a negative reimbursement advice.

Other countries only appear to operationalize medical need, The Netherlands formally do so during the appraisal process using disease severity based on the proportional shortfall definition.[119,120] Sweden uses medical need and solidarity as one of the three main principles for priority-setting in health care, which is further defined by various levels of disease severity: life-threatening diseases, disease prevention, and less severe acute and chronic diseases.[107] Belgium uses necessity of treatment to determine the level of reimbursement, ranging from necessary for life-threatening diseases to symptomatic treatment. The relative weight of medical need vis-à-vis other needs is in all countries unclear.

Although rarity of a disease was also mentioned by interviewees from all countries as important to decision making, whether rarity as such determines need, or the fact that often no alternative treatment exists for a severe disease that happens to be rare, is unknown.[121]

*Question 2: Is society prepared to pay with public resources for a treatment that will improve the indication in question?*
Preparedness to pay is independent of ability to pay and product price, a feature that differentiates 'preparedness' from 'willingness'. Before discussing preparedness to pay, policymakers should determine whether society is prepared to pay for anything that would improve the indication of the treatment under consideration. Preparedness to pay is independent of a particular treatment's need, cost, or effectiveness but might depend on the causes of the disease (e.g., unhealthy or risky behaviour), the characteristics of the population groups affected by the disease (e.g., their socioeconomic status) or the nature of the outcome (e.g., relief of a headache). The answer might be *"Yes, if…",* in which case preparedness to pay is subject to conditions.

Although the preparedness to pay out of public resources is not necessarily strictly linked to the medical, therapeutic and societal need, we found that both judgements are in practice frequently considered equal. This indicates that society believes that treatments for high needs should be able to rely on public funding, regardless of, for instance, personal responsibility. Therefore, countries operationalise this question similarly to the needs question, meaning they are in principle prepared to pay for treatments for high medical, societal, or therapeutic needs.

*Question 3: Do we want to pay for this product out of public resources?*
Societal willingness to pay for the treatment under consideration, given its characteristics, may depend on the effectiveness and therapeutic value of the treatment compared with alternative treatments and whether it concerns a curative, symptomatic, or

preventive treatment. It can also depend on the burden of the costs of a treatment, for example, for a relatively cheap treatment such as paracetamol, the administration costs of reimbursement would be higher than the treatment itself.

All countries evaluate the therapeutic value of each individual drug to evaluate whether the drug should be reimbursed and thus paid for by society. This question is often considered in combination with preparedness to pay (i.e., question 2).

*Question 4: Do we want to pay more for the drug compared with the comparator?*
Whether society wants to pay more for a drug than its comparator depends on the product's added societal value, which depends on its added therapeutic value, potential savings effected elsewhere in the health care sector, and the quality and certainty of the evidence on these two criteria.



All countries but Austria use internal reference pricing to determine the reimbursed price for products with equivalent therapeutic value, meaning society is not willing to pay more for the drug than other products with equivalent therapeutic value. Added therapeutic value can be decomposed in several elements; increased efficacy and/or effectiveness and safety get the highest weight in all countries. A drug judged to have added therapeutic value is likely to be reimbursed at a higher price. Although improvement in comfort, ease of use and applicability are mentioned as determinants of added therapeutic value, they are in practice rarely sufficient for a product to be reimbursed at a higher price.

*Question 5: How much more is society willing to pay with public resources for this treatment?*
Societal willingness to pay depends on societal value. This value is determined by all previous criteria and is independent of price. In practice, it is difficult to measure societal value in monetary terms. Therefore, in a supply-driven context, where pharmaceutical companies decide what, when, and at what price to launch a drug, policymakers will in practice have to consider whether the price requested by the company is reasonable given its societal value.

For this purpose, cost-effectiveness and budget impact are used as decision criteria, strongly depending on the previously described value-criteria. While cost-effectiveness is traditionally seen as a criterion for assessing efficiency, it only does so when health maximisation is the main objective and a threshold value for the incremental cost-effectiveness ratio (ICER) is defined. All our countries deny using an ICER threshold value or a threshold range, thus confirming the observation of previous studies that the ICER has limited weight in the appraisal process.[122,123] Instead of being a criterion for technical

efficiency, the ICER can also be used as an instrument or measure to judge the acceptability of an intervention's cost, given its societal value. This requires the weighing of all value-criteria against each other. In this respect, all five countries appear to be willing to pay more for a unit of health gained in case of more severe diseases.

It is difficult to define a priori the relative weight of each criterion because decision makers and stakeholders might want to give different weights in different situations. For example, therapeutic value may get more weight when no alternative treatment is available. Interviewees from all countries reported a higher willingness to pay for drugs for rare diseases for which no alternative treatment exists (therapeutic need). This may then suggest the acceptability of a higher cost-effectiveness ratio. The Netherlands also reports a higher willingness to pay for more severe diseases and France and Belgium apply a lower level of cost sharing to drugs for more severe diseases. Table 4.3 illustrates examples how our countries operationalize the relationship between separate value-criteria and the (additional) willingness to pay.

**Table 4.3** Illustrations of the relationship between value-criteria and societal willingness to pay

| Value criterion | How the value criterion influences willingness to pay |
|---|---|
| Medical, therapeutic, and societal need | In *Sweden* priority-setting principles state that persons in greatest medical and therapeutic need should get the highest priority. The Swedish expert committee refers to "marginal utility", which is further defined as "if no alternative treatment exists, cost should be reasonable"; "reasonable", however, remains undefined.[124]<br>In *the Netherlands*, the reimbursement agency recently suggested a threshold range for the incremental cost-effectiveness ratio (ICER), where willingness to pay varies within that range depending on disease severity (medical need).[120] The Dutch minister has neither confirmed the range nor endorsed an ICER threshold.<br>In *Austria*, medical need is implicitly considered in the evaluation of the therapeutic benefit, which in turn is strongly related to the price. |
| Added therapeutic value | *Belgium* and *the Netherlands* apply a binary outcome (i.e., yes or no) for added therapeutic value, thereby not relating societal willingness to pay to the degree of added therapeutic value.<br>*France* and *Austria* classify the degree of added therapeutic value in five and six categories, respectively. Societal willingness to pay is defined in function of the added value category.<br>*Sweden* uses the ICER to determine an acceptable price of a product, thereby directly relating societal willingness to pay to the degree of added therapeutic value. |

Additional criteria helping policymakers to assess the acceptability of a requested price are budget impact and mechanisms for cost sharing. Although these criteria are not value-criteria, they cannot be considered independently. Table 4.4 illustrates how budget impact and cost sharing can modulate decisions and give incentives to install measures to stimulate value-based medicine.

**Table 4.4** Illustrations of how criteria for judging the acceptability of a requested price take value-criteria into account and can shape drug reimbursement decisions

| Criteria for appraising the acceptability of a requested price | Relation with value criteria |
|---|---|
| Budget impact | Budget impact is a decision criterion in *all countries* but Sweden. In *Sweden* regional county councils are responsible for the financing and implementation of decisions. The same county councils are also responsible for clinical guidelines, which include financial incentives that stimulate the usage of preferred drugs.<br>*France* and *Belgium* use price-volume agreements (financial risks-sharing agreements).<br>For statins, *Belgium* has defined a "first choice treatment" (reimbursed without conditions) and a "second choice treatment" (subject to the condition that the first choice failed to benefit a patient).<br>*All countries* use financial incentives to influence utilisation (co-payments, co-insurance, deductibles). |
| Cost sharing | *Belgium* uses "necessity of treatment" (medical need) to define the level of cost sharing. A negative correlation has been observed between the level of cost-sharing and the added therapeutic value of drugs,[123] indicating that products for more severe diseases are more likely to be considered of added therapeutic value.<br>In *France*, medical need is one of the criteria determining the SMR rating (Service médical rendu), which determines the level of cost-sharing.<br>Both *France* and *Belgium* use the medical needs criterion to define the level of co-insurance or co-payment, which increases affordability of the most necessary treatments. |

Finally, uncertainty of evidence may impact upon the appraisal of the value-criteria as well as on the societal willingness to pay. For example, uncertainty about the added therapeutic value in daily clinical practice might lead expert committees to lower their estimate of the added therapeutic value, advise restricted reimbursement, or deny reimbursement altogether. They could also make a temporary reimbursement decision or negotiate a lower price.[23] Our five countries deal with uncertainty in budget estimates by the implementation of one or more of the following measures: financial risk sharing agreements, price negotiations, cost sharing, and conditional reimbursement.

## Discussion

An in-depth analysis of five European drug reimbursement systems showed that these systems use similar criteria in their drug reimbursement decision processes. The relative importance attached to each of the criteria may vary, but the implicit questions posed during a decision-making process are similar. Our study shows that there is room for improving the transparency and relevance of decision criteria, two legitimacy conditions.

Empirical evidence suggests that the four legitimacy conditions defined by the Daniels and Sabin framework actually improve perceived legitimacy, fairness, and quality of decision-making, but should be used flexibly.[101,106-108,125,126]

The public payer is continuously faced with the dilemma of simultaneously ensuring equitable access to high quality health care and sustainability of the health care system. The challenge for policymakers is therefore to find a publicly acceptable balance between the objectives. This is pursued by considering and weighing several criteria in the decision-making process. Added therapeutic value, being the most prominent criterion in decision making in all countries, addresses the quality of care objective. Disease severity, also important in decision making in all countries, reflects the equity objective. Cost-effectiveness addresses the objective of efficiency (maximising health with a given amount of resources). It is a reimbursement criterion in all countries, be it only for revisions in France. No country, however, uses a fixed ICER threshold value; even threshold value ranges seem unacceptable. Budget impact, which also reflects the sustainability objective, is considered in all countries either at the national or at the regional decision level. Although all countries have a more or less open-ended pharmaceutical budget, reimbursement can still be denied for budgetary reasons. Disease rarity, a frequently mentioned decision criterion, reflects the equity objective of systems: patients with rare diseases should have equal chances of affordable treatment. It gives companies the opportunity to set high prices and remain somewhat inflexible in price negotiations.

While all these criteria are relevant, their relative importance and how they shape the final decision often remains unclear. This can result in differences in accountability of the systems: the lower the transparency of both formal and informal criteria, the less accountable the system.

Our relatively small sample of countries, not necessarily representative for Europe, could be seen as a limitation to our study. However, important differences were observed in structure, organisation and procedures of the drug reimbursement systems, supporting the external validity of our study. A thorough understanding of the explicit and implicit processes taking place during a drug reimbursement decision process required an intensive search for formal as well as informal information. Such an analysis was therefore only feasible in a small number of countries within a reasonable period of time.

Our conclusions could be extended by reviewing actual reimbursement dossiers in these countries. For example, judging the consistency of decision making requires detailed comparison of reimbursement dossiers. Furthermore, different appraisal processes and reimbursement criteria can still produce comparable results in terms of drug expendi-

tures, health gains and equity. It would be worthwhile to further investigate to what extent these crucial outcomes are sensitive to country differences in reimbursement policy.

We believe, however, that transparency of the drug reimbursement decision process can be improved in all countries by using an explicit decision framework. Our developed framework provides a first provisional tool to structure the decision-making process, it can support the justification of decisions, and is a tool for defining and making explicit the societal choices which currently often remain implicit. Crucial is the societal acceptability of the decision criteria. Proper justification of the reimbursement advice or decision, with a sufficiently differentiated reflection on the multiple considerations taken into account during the appraisal and decision-making processes and with a clear statement on the final position taken on each key question, ensures transparency and enhances trust in the system. No system can define a general rule applicable to decisions in all situations,[122] but the decision process can be reconstructed by providing an explicit answer to each crucial question.



## Conclusions

To reach accountability for reasonableness and thus ensure a legitimate drug reimbursement process, any democratic political system has the obligation to be transparent, use societally-relevant rationales in decision making, allow revisability of decisions in the light of new evidence, and enforce the three previous conditions. Many systems currently lack transparency, especially in the use of appraisal criteria. The appraisal process could benefit from using an explicit decision framework specifying the social choices and decisions made during the appraisal process as well as the criteria on which the choices and decisions are based. This would improve accountability and coherence between decisions, and, in turn, enhance legitimacy of societal decision making on drug reimbursement.

## Acknowledgements

We would like to thank Gérard De Pouvourville, John Hutton and Willy Palm for their helpful comments on a preliminary version of our project report. We are also grateful to Frans Rutten, Ton de Boer, Adri Steenhoek, Elly Stolk, Erik Schokkaert and Carine Van de Voorde for their advice during our research project.

This study was partly performed in the context of the Escher project (T6-202), a project of Top Institute Pharma, Leiden, The Netherlands.

# 5

# A comparative study of the role of disease severity in drug reimbursement decision making in four European countries

With Elly Stolk, Tessa Scharringhausen, Anthonius de Boer, and Marc Koopmanschap

Submitted for publication

## Abstract

Considerations beyond (cost-) effectiveness are important in reimbursement decision making. We assessed the importance of disease severity in drug reimbursement decision making in Belgium, France, The Netherlands and Sweden.

We investigated scientific literature and policy documents and conducted three interviews in each country (four in The Netherlands) with persons involved in drug reimbursement.

Disease severity, as part of the 'need principle', is an important consideration; especially in case of high severity. The Netherlands operationalised disease severity using the proportional shortfall approach. Sweden uses categories to give an indication of the severity. In The Netherlands and Sweden, severity only implicitly plays a role in the decision whether to reimburse a drug, whereas in Belgium and France it also explicitly plays a role in determining the willingness to pay out of public sources. Interviewees acknowledged that quantitative information, besides a qualitative description of the disease, may provide additional information. None of them, however, considered such information to be of decisive importance.

Although disease severity is important in drug reimbursement decision making in all four countries, all seem to struggle in making its actual role explicit. Belgium and France are most explicit by using severity in setting reimbursement levels; all four countries could, however, improve the transparency of its actual importance relatively to the other criteria in the decision making.

## Introduction

Considerations beyond (cost-) effectiveness and safety are important in reimbursement decision making.[103,127-131] One such consideration is disease severity or any other operationalization of 'medical need'. Valesco-Garrido et al.[132] found the 'need' aspect to be one of the most considered criteria in nine European countries. It seems that drugs that target more severe diseases more easily obtain reimbursement; this reflects the equity objective of most health care systems. On the other hand, health care interventions may not be reimbursed for diseases of low severity as it indicates a limited 'need'. Many stakeholders and analysts seem to agree on the fact that the severity of the disease may play a role in reimbursement decision making. However, the way disease severity is used is often not transparent and may differ across countries.[15,130]

Previous studies investigated the importance of societal concerns, for example, by eliciting policymakers' preferences,[133,134] or by eliciting public preferences regarding the social value of the Quality Adjusted Life Year (QALY),[135-137] preferences for severity of illness,[138-140] for end-of-life treatment,[141,142] or for the rule-of-rescue.[143] These studies provided insights into the importance of the severity of the disease. Most studies are, however, mainly theoretically based. Consequently, there are still many unsolved issues regarding the implementation in actual practice. A comparison across countries on the role of disease severity in reimbursement decision making may provide insights into actual policymaking. It may, therefore, facilitate learning probabilities how to incorporate this criterion in health care resource allocation as many countries seem to experience difficulties in doing so.

We investigated the role of disease severity and its operationalization in drug reimbursement decision making in four European countries: Belgium, France, The Netherlands and Sweden. We explored concepts of disease severity and assessed how information on disease severity is used in everyday drug reimbursement decision making. This paper addresses the following questions: (i) is disease severity considered in drug reimbursement decision making; (ii) if considered, how or by which method or indicator is it presented; and (iii) in what way is this information used in practice and how does it affect decision making.

## Context

All four countries have a national drug reimbursement agency. Within the agency, a technical department prepares the assessment and preliminary reports. An independent

pharmaceutical expert advisory committee assesses and appraises the evidence and is responsible for advising the final decision maker (i.e., the minister of health). Notably, the ministers hardly ever deviate from the advice.[103] In Sweden, the expert advisory committee also makes the final decision. Only The Netherlands has a separate appraisal committee, which has an advisory function. All four countries use drug effectiveness and cost-effectiveness (in France only recently -2012- for new drugs) as reimbursement criteria. This paper focuses on the role of disease severity in decision making. For more detailed information on the four systems regarding their (other) reimbursement criteria, reimbursement process, and their Health Technology Assessment (HTA) models, we refer to Franken et al.[103] and Le Polain et al.[15] Table 5.1 provides an overview of the four drug reimbursement systems.

**Table 5.1** Summary characteristics of the four drug reimbursement systems

| | **Belgium** | **France** | **The Netherlands** | **Sweden** |
|---|---|---|---|---|
| **System characteristics** | | | | |
| National reimbursement agency | INAMI/RIZIV | HAS | CVZ | TLV |
| Pharmaceutical expert advisory committee | CTG/CRM | CT | WAR-CG | PBB |
| Appraisal committee | n/a | n/a | ACP | n/a |
| Final decision maker | Minister | Minister | Minister | PBB |
| Expert committee report publicly available | Yes | Yes | Yes | Yes (No, if applicant withdraws request) |
| **Assessment and appraisal criteria** | | | | |
| Medical, therapeutic and societal need | Yes | Yes | Yes | Yes |
| Added therapeutic value | Yes | Yes | Yes | Yes |
| Cost-effectiveness | Yes | Yes (for new drugs since 2012) | Yes | Yes |
| Explicit threshold for cost/ QALY | No | No | No | No |
| Budget impact | Yes | Yes | Yes | No |
| Other | Price and reimbursement basis | Public health, treatment properties, disease severity | Disease severity and rarity, own responsibility, accessibility, societal affordability, public health | All effects on a person's health and quality of life, human value, need and solidarity |

*ACP = AdviesCommissie Pakket; CRM = Commission de Remboursement des Médicaments; CT = Commission de la Transparence; CTG = Commissie voor Tegemoetkoming Geneesmiddelen; CVZ = College voor Zorgverzekeringen; HAS = Haute Autoritée Santé; INAMI =Institut National d'Assurance Maladie-Invalidité; n/a = not applicable; PBB = Nämnden för läkemedelsförmåner; QALY = Quality Adjusted Life Year; RIZIV = Rijksinstituut voor Ziekte- en Invaliditeitsverzekering; TLV = Tandvårds- och läkemedelsförmånsverket; WAR-G = Wetenschappelijke AdviesRaad Geneesmiddelen*

## Concepts of disease severity

Need principles are commonly discussed in academic literature concerning health care rationing decisions.[112] A drug is valuable when it fills a specific need; this may be a medical, therapeutic, or societal need.[110] These needs depend on factors such as treatment necessity and disease severity (medical need), the availability and the effectiveness of alternative treatments (therapeutic need), and the prevalence of the disease and inequalities in health (societal need).[15] As such, disease severity is part of the 'need principle'; the more severe a disease, the higher the (medical) need.

A large number of empirical studies acknowledge the concept of 'severity' as prioritising principle.[140] Several approaches to determine 'who are the worst off' are described in the literature to operationalize the concept of severity. According to the 'fair innings' approach, everyone is entitled to some 'normal' span of health achievement.[115] This implies that treatments for patients who did not yet had their fair innings are valued higher than treatments for patients who have had their fair share. The latter ones are, according to Williams,[115] 'living on borrowed time'. This approach considers life time health achievement including the quality as well as the length of a life, thus also including past health (losses). The 'severity-of-illness' approach prioritises persons with the worse off initial condition based on the severity of the initial health state as well as the expected (prospective) health in case no treatment is available.[144,145] This approach emphasises the need to rescue persons with a severe condition (e.g., facing immediate death); the approach does not consider past health. In contrast to the previous two 'absolute' worse off approaches, the 'proportional shortfall' approach, considers the worse off in relative terms. This approach bases the need on the proportion of health lost due to the disease as compared to the expected health (i.e., level of health and remaining life expectancy) without the disease.[146] Finally, the rule of rescue' approach prioritises identifiable individuals facing avoidable death, regardless of the costs.[118] It should be noted that, if used for priority setting, all approaches can identify a group of persons (or an individual) who are the worst off. The last approach, however, deviates because it concerns identifiable individuals who are worst off no matter what disease, whereas the others identify the worst off per disease and concern a measure of loss. Therefore, only the first three approaches facilitate a numerical expression of the severity of a specific disease that could be used in reimbursement decision making at the national level.

## Methods

We compared the role of disease severity and its operationalization in drug reimbursement decision making in Belgium, France, The Netherlands and Sweden. The selection of these countries was based on our previous research on European drug reimbursement systems.[15,103,130,147] All four countries have established HTA agencies using reasonably comparable reimbursement processes.

To obtain insight in the role of disease severity in actual decision making, we first evaluated the reimbursement processes and criteria. Second, we explored scientific literature describing concepts of disease severity. Third, we conducted face-to-face interviews; three in each country (four in The Netherlands because of the existence of the appraisal committee). Fourth, using data triangulation we combined the information from the literature, policy documents and the interviews to assess the role of disease severity and its operationalization in drug reimbursement decision making.

All interviews were tape-recorded. Citations were reported anonymously and were translated by the authors if the interview was not in English (i.e., interviews in The Netherlands and Belgium were in Dutch, interviews in France and Sweden were in English). The selection of interviewees was based on their specific involvement in drug reimbursement. They were either a representative of the reimbursement agency or an expert from the pharmaceutical expert (advisory) committee or the appraisal committee. In each country, we interviewed at least one person with an economic background and one person with a medical or pharmacy background. Interviews were semi-structured; questions addressed topics such as explicit versus implicit use of disease severity, operationalization(s) of disease severity, reimbursement levels, qualitative versus quantitative information on disease severity, published versus not published reimbursement information, and assessment and appraisal considerations that might be related to or interact with disease severity (e.g., age, child vs. adult, fair innings, past health, future health, absolute vs. relative health status, health gain, quality of life, rarity of the disease, availability of alternative treatments, end-of-life treatment, rule of rescue, medical need, life-style, own responsibility, necessity to insure, and personal vs. societal affordability).

All interviewees had the opportunity to comment on the final draft of the paper. Although the reflections made are based upon information from the interviews, it should be noted that the reflections are strictly ours and do not necessarily reflect the views of the interviewees.

## Results

### Is disease severity considered in drug reimbursement decision making?

All interviewees acknowledged considering the severity of the disease in drug reimbursement decision making, especially in case of high severity. An interviewee replied *"yes definitively, oh yes"* to the question whether disease severity played a role.

### How or by which method or indicator is disease severity presented?

In all four countries reimbursement information contains a qualitative description of the disease and, besides that, most often contains information on survival, progression free survival, and sometimes on quality of life. This information is prepared by the reimbursement agency and presented to the pharmaceutical expert committee, and in The Netherlands to the appraisal committee. Also published reimbursement information most often contains a qualitative description of the disease.



The Swedish agency indicates whether the disease is 'highly', 'moderately' or 'less' severe. This information is always available for the pharmaceutical expert committee, but not always available in the published memorandum. In The Netherlands, the proportional shortfall approach is presented to operationalize the concept of disease severity in policy documents.[120] However, numerical outcomes based on the proportional shortfall approach were until now not available for the pharmaceutical expert committee, and only recently a few times available for the appraisal committee as well as in published reimbursement information (i.e., for the first reassessments of expensive inpatient drugs). Moreover, members of the appraisal committee have indicated that they preferred to be informed with more information than only the proportion shortfall calculations, for example, also information based on the fair innings approach.

In France, disease severity is one of the five Service Médical Rendu (SMR) criteria; the other SMR criteria are efficacy and adverse events, place of the drug, availability of alternative treatment, treatment properties and public health benefit. Disease severity is, however, no criterion of the new proposed French Relative Therapeutic Benefit (RTB) criterion, which is proposed to replace the SMR and the Amélioration du Service Medical Rendu (ASMR) criterion (both the price and reimbursement decision will be based on the RTB criterion). Furthermore, France has special regulations for thirty serious and chronic diseases (i.e., the Affections de Longue Durée (ALD) List), certain irreplaceable and costly drugs (e.g., cancer or AIDS), costly drugs for diseases that constitute a progressive or disabling disorder with a previous treatment period over six months (i.e., the so-called '31st disease'), and multiple diseases of over six months (i.e., the so-called '32nd disease').[15]

Belgium uses reimbursement categories which reflect the necessity of the drug, and thus partly reflect disease severity. The categories are category A (and Fa) for vital drugs for life-threatening diseases (e.g., diabetes and cancer), category B (and Fb) for therapeutic significant drugs for non-life threatening diseases (e.g., antibiotics), category C for therapeutic less significant drugs for systemic treatment –symptomatic treatments–, category Cs for chronic illnesses, and category Cx for contraceptives and antispasmodics.[148]

**In what way is information on disease severity used in practice and how does it affect decision making?**

All interviewees acknowledged that disease severity affects drug reimbursement decision making. An interviewee stated *"if you talk about disease severity, you have of course the implicit as well as the explicit weighing."* Interviews revealed that disease severity often plays an implicit role and sometimes an explicit role in the consideration whether or not society is willing to pay for a treatment.

*The role of disease severity in decsion making whether or not to reimburse a drug*

Disease severity or, overarching, the need principle can be found in all four countries' policies as prioritising principle. All share similar system objectives: equitable access, quality of care and system sustainability.[103] For example, the Swedish Health and Medical Service Act (1982:763) emphasises equal access to health services on the basis of need and a vision of equal health for all. Accordingly, the three Swedish prioritising principles are human value, need and solidarity (i.e., those in greatest need of health care should be given priority access to care), and cost-effectiveness.[107] Article R163–1.6.3 of the French Code de la Se´curite´ Sociale' states that drugs should not be reimbursed in case of absence of severity of the disease they address.[149] In The Netherlands, necessity (i.e., whether the severity of the illness or the care needed justifies solidarity) is one of the four criteria for determining the basic benefit package (including drugs).[120] In Belgium, it is obliged by law to position new technologies relatively to other available technologies in medical practice.

Effectiveness, safety, and cost-effectiveness are formal and explicit reimbursement criteria. All interviewees, however, agreed that, implicitly, the more severe a disease, the higher the chance to obtain reimbursement and/or a higher level of reimbursement. As indicated by interviewees *"the severity of the disease balances the benefit risk ratio"* and *"it is an issue for discussion for drugs that just go with the narrowest of margins."* Although disease severity is a formal appraisal criterion in The Netherlands, drugs are not often discussed in the appraisal committee, so far only reassessments of expensive inpatients drugs have been discussed in the appraisal committee. Even though the French SMR

classification seems straightforward, interviewees acknowledged being more lenient in case of high(er) severity. A French interviewee stated *"The importance of severity might be very high, you have the same type of evaluation, assessment of the efficacy and effectiveness, but since in one case it is a very severe disease and you want to give something to the patient, it is reimbursed, and it has a level of SMR that opens reimbursement and in the other case, I can tell the patient is not happy about that."* Regarding the effect of the list of diseases for full reimbursement, another French interviewee stated *"When a new drug is about one of these diseases so, we are authorised to be very rigorous …… our thinking is only about the level of efficacy …… we are sure that when you give 15% the patient will be full reimbursement, but if we give a bad score the pricing will be difficult."* This seems to imply that the French committee not only aims to ensure reimbursement for severe diseases, but also, simultaneously, aims to influence the price to ensure value for money.

Seven out of thirteen interviewees (i.e., four Dutch, two Swedish and one Belgian interviewee(s)) were familiar with at least one of the disease severity concepts. Confronting interviewees with elements from these concepts revealed that interviewees did not consider age as criterion in the decision making. Some even mentioned that it would be illegal to use age as criterion. However, all acknowledged being more lenient towards drugs for children. Past health was also not seen as criterion. Consequently, the 'fair innings' approach does not seem to fit well in actual practice. Also 'the rule of rescue' approach does not fit in actual practice because the agencies make national decisions and, therefore, cannot consider identifiable individuals. However, media attention in specific (individual) cases may lead to (ad-hoc) decisions based on 'the rule of rescue' when policymakers are under societal pressure. Moreover, decisions on (ultra) orphan drugs can also be influenced by 'the rule of rescue' considerations. Interviewees stated that the absolute gain in health or quality of life (future health effects) is most important, and also the current (absolute) health status is important. This reasonably fits with Nord's 'severity-of-illness' approach. Even though it is Dutch policy to use the 'proportional shortfall' approach, interviewees stated *"The proportional shortfall is one way of doing it, but this may not be appropriate in all cases, that depends on the type of disease, sometimes it can be better to use other methods to quantify the severity, or use more methods at the same time and discuss what is most relevant for the specific case,"* and *"but the description of the disease next to it also remains important."*

All interviewees indicated that they would like to have, at least, a qualitative description of the disease. Only six out of thirteen stated to appreciate the availability of numerical expressions of disease severity. However, such figures were only appreciated if presented additional to a qualitative description of the disease, and, if appropriate, additional to an incremental cost-effectiveness ratio. An interviewee explained *"if you*

*compute all information including disease severity into one cost-effectiveness estimate, then it will become very opaque, something we will not do so quickly, it is more that we come to a decision by putting cost-effectiveness, disease severity and other arguments side by side."* Other interviewees expressed concerns regarding the use of quantitative information on disease severity in decision making: *"you think is it possible, is it easy to do it in a reliable way?"* and *"cost-effectiveness, now it is accepted in principle, but still, …… medical people are still suspicious about it, …… the speed in these developments goes too fast, the rest of the society is not there"* and *"using numbers between one and zero is at odds with the complexity of the reality,"* and *"they (i.e., expert committee members) would prefer more qualitative and not too formal information, with possibility of discussing, …… they would prefer to feel more free."*

Although all interviewees agreed that the severity of the disease affects the decision whether or not to reimburse a drug, none of them could indicate its relative importance compared to other decision criteria. All interviewees agreed that it remains a balancing exercise. One interviewee explained *"Another discussion is our willingness to pay, at the moment, we maybe sometimes; I personally do think that we are willing to pay more for severe disease compared to for example an erectile dysfunction."* Although formally using the proportional shortfall approach in The Netherlands, the relative importance of disease severity remains unclear. Similarly, the classification used in Sweden (i.e., low, moderate, high severe) does not explicate how important disease severity is relatively to other criteria in the decision making. A Swedish interviewee stated *"in some dossiers it is (i.e., published severity scoring), we probably could and should publish it more frequently those statements, but it is available in some, …… if it is relevant …… you still have a balancing there (i.e., in the public memorandum), but it is not as clear."* Nevertheless, previous research (i.e., a comparative study of Dutch and Swedish published reimbursement information[147]) found that in four out of eleven cases disease severity was explicitly mentioned. Although the relative importance of disease severity was not stated, all four drugs obtained reimbursement, three of them concerned indications in cancer.

Regarding other decision criteria, a Belgian interviewee indicated that budgetary impact may even be more important than disease severity. Interviewees acknowledged that the 'end-of-life' criterion is mostly visible in treatments in cancer, thus concerning treatments that most often obtain reimbursement. Nevertheless, all agreed that even in such cases a minimal absolute gain, without clearly indicating what kind of minimum would still be acceptable, was still important. Interviewees did not consider the fact whether the treated disease was self-inflicted or someone's 'own fault.' On the other hand, especially in case of low disease severity and low treatment costs, interviewees agreed considering the necessity to reimburse a treatment and/or whether treatment costs could be a

person's own responsibility. Furthermore, most interviewees agreed that rarity as such is not (that) important; most also indicated a correlation between severity, rarity and the availability of alternative treatments. As explained by an interviewee *"that is difficult, I think if all three elements are there, in case it concerns a rare disease, which is also severe and no other alternative treatment is available, that scores high, we are more lenient in such a situation."* Another interviewee stated *"disease severity was number one …… now we are not thinking the same way, it is very rare, it is not a frequent situation that there is no other drug …… we are reluctant to give favourable opinion when the drug is not very very, they are all efficacious because they have market authorisation of course, but the level is sometimes very very thin."* This shows that disease severity correlates with the 'need principle', in this case a relation between the medical need and the therapeutic and societal need.

*The role of disease severity in determining the level of reimbursement*

Once drugs are granted reimbursement, they are fully reimbursed in The Netherlands and Sweden. In contrast, Belgium and France use levels of reimbursement which are partly based on disease severity (see Table 5.2). All Belgian and French interviewees agreed that the reimbursement levels partly reflect the severity of the disease. In France, the reimbursement levels (i.e., 65%, 30% and 15%[150]) depend on the SMR level. Besides that, drugs are fully reimbursed for patients with a disease included on the ALD list, 31st and 32nd diseases, or drugs classified as being irreplaceable and costly. Interestingly, disease severity is no criterion of the new proposed RTB criterion, whereas it was for the SMR, implying that, in the future, the reimbursement level will not (partly) reflect disease severity anymore. Nevertheless, a French interviewee stated *"if you have a severe*



**Table 5.2** Levels of reimbursement and cost sharing mechanisms

| | Belgium | France | The Netherlands | Sweden |
|---|---|---|---|---|
| **Reimbursement level** | | | | |
| Level of reimbursement | 100%; 75%; 50%; 40%; 20% | 100% (ALD, 31st and 32nd diseases); 65%; 30%; 15%; 0% (SMR level) | 100% | 100% |
| Basis for the level of reimbursement | Category of treatment necessity: A (Fa), B (Fb), C, Cx and Cs | SMR level: Important, Moderate, Weak, Insufficient | n/a | n/a |
| Cost sharing | Product specific co-insurance | Product specific co-insurance; prescription fees | General health care deductible | Drug specific co-payment |

*Category A (and Fa): vital drugs for life-threatening diseases (e.g., diabetes and cancer); B (and Fb): therapeutic significant drugs for non-life threatening diseases (e.g,. antibiotics); C: therapeutic less significant drugs for systemic treatment –symptomatic treatments-; Cs: chronic illnesses; Cx: contraceptives and antispasmodics;*
*ALD: Affections de Longue Durée; SMR: Service Médical Rendu*
*n/a : not applicable*

*disease you might admit a less important difference than in a non-severe disease"* and *"for a non-severe disease it is not justified to have a high level of reimbursement."*

Belgium'ms reimbursement levels (i.e., 100%, 75%, 50%, 40%, and 20%) depend on the reimbursement categories (i.e., the therapeutic necessity of the drug category)[148] which partly reflect the severity of the disease. Category B contains by far the largest group of drugs followed by Category A (i.e., 78.3% and 19.3% in 2009, respectively[151]). The decision base for the reimbursement level hardly ever gives rationale for discussion. As a Belgian interviewee stated: *"the category is usually not a discussion because the companies know very well, by analogy, where to place their product."*

## Discussion

We investigated the role of disease severity and its operationalization in drug reimbursement decision making in Belgium, France, The Netherlands and Sweden. As expected, the severity of the targeted disease is an important, but often implicit, consideration in drug reimbursement decision making; the more severe the disease the higher the chance to get the treatment reimbursed (and at a higher reimbursement rate). In The Netherlands and Sweden disease severity only implicitly plays a role in the decision whether or not to reimburse a drug, whereas in Belgium and France it also explicitly plays a role in determining the percentage of costs society is willing to pay out of public sources for a specific treatment targeting a more or less severe condition. About halve of the interviewees (7 out of 13) were known with at least one of the concepts of disease severity. However, all four countries seem to struggle in operationalizing the concept and making the actual role of disease severity explicit.

A limitation of our study is that we only conducted thirteen face-to-face interviews in four countries. Expert committee meetings are, however, not public; therefore it is impossible to attend such meetings and observe the deliberation in actual practice. Nevertheless, we conducted interviews with persons involved in everyday decision making and observed many similarities across the four countries. We therefore believe that saturation was achieved and that our study provides important insights into the role of disease severity in actual drug reimbursement decision making in Belgium, France, The Netherlands and Sweden and yields valuable lessons for policymakers and Health Technology Assessment (HTA) researchers.

National reimbursement agencies are responsible for appraising whether a drug is worth paying for by society and are thus accountable to society. According to Daniels

and Sabin,[102] a fair and legitimate prioritising procedure must satisfy four conditions: (i) transparency of the decision-making process; (ii) relevance of the decision criteria; (iii) revisability of the decision in light of new evidence and arguments; and (iv) enforcement of the existing criteria. Our study shows that disease severity is a relevant decision criterion. However, our study also shows that policymakers experience difficulties in employing the concept of disease severity explicitly, and also the importance of disease severity relatively to other decision criteria is not transparent.

The governance structure to safeguard legitimacy of reimbursement decision making may differ in our four countries. The Swedish, Dutch and French system are more information –assessment– driven compared to a more deliberation –process– driven system in Belgium in which stakeholders are part of the discussion.[15] However, our study illustrates that all four countries struggle in making the role of disease severity explicit. Even though The Netherlands has a separate appraisal committee and formally uses the proportional shortfall approach and Sweden classifies the severity in categories, it remains unclear how important disease severity is relatively to other criteria in the decision making.



Using levels of reimbursement, which depend on the severity of the disease, as in Belgium and France, can be a legitimate way of incorporating disease severity into the decision how much society is willing to pay out of public sources for treating a specific condition. However, other countries may be averse against the use of co-payments, and using levels of reimbursement may therefore be seen as inappropriate or inequitable. In such cases, the severity of the disease can only play a role in the decision whether or not to reimburse the health care intervention as in The Netherlands and Sweden.

Interestingly, other studies[129,142,152] found strong support for 'fair innings' arguments in the general public. In contrast, our study revealed that age is not important for policymakers and that using age as criterion as advocated by 'fair innings' arguments may even be against the law in some countries. Nevertheless, interviewees indicated to be more lenient in case of treatments targeting children. Our study also revealed that past health is not important, but that absolute future health gains as well as the current health status are important in decision making. This fits with 'severity of illness' arguments. Although identifiable individuals cannot be considered in national decision making, media attention and societal pressure may appeal to 'the rule of rescue' arguments. Cookson[153] analysed eleven potentially relevant justifications for the NICE 'end-of-life' premium; he concluded that none of them provides sufficient ethical justification. Interestingly, our interviewees indicated that even in case of end-of-life treatment a minimal absolute gain is still important. Such findings have been previously advocated by Kvamme et al.[154]

Another important finding of our study is that policymakers prefer to have, at least, qualitative information on the severity of the disease and, above all, prefer to maintain their discretionary decision power and implicitly weigh all decision criteria. It remains the question why policymakers experience such difficulties in explicating their decision base. It could be that using an explicit criterion and instrument not only leads to favourable decisions at one end of the scale (i.e., high severity leads to reimbursement), but also may enforce making and/or explicating unfavourable decisions at the other end of the scale (i.e., low severity leads to no reimbursement). The latter one may be a less comfortable position. Nevertheless, we believe that policymakers could enhance the transparency of the actual role of disease severity and/or underlying considerations (e.g., future health, past health, age, end-of-life considerations and medical, therapeutic and societal needs) in their decision making. This can be achieved by using classifications (as in Sweden) or numerical expressions of disease severity (as in The Netherlands), but, more importantly, more information on the weighing exercise of the criteria could be published. For example, published information can describe that because of the high severity of the disease (e.g., remaining life expectancy of 6 months), the absolute gain of two months in life expectancy is acceptable. Policymakers may (re)consider the use of additional HTA evidence; for example, evidence derived from the disease severity concepts, and/or evidence from multi criteria decision analysis (MCDA) including evidence from obtained (public) preferences (e.g., social value of the QALY,[135-137] preferences for severity of illness,[138-140] end-of-life treatment[141,142]). This may help to improve consistency in the deliberation and increase the legitimacy of their decision making.

On the other hand, it is important that academic and HTA researchers realise that HTA and/or MCDA evidence will most likely not be decisive and only be informative to policymakers. Our study reveals that policymakers may not be interested in explicitly weighing QALYs. Many policymakers do not prefer to use quantitative information on disease severity and prefer to implicitly weigh all criteria and, above all, prefer to maintain discretionary decision power. To further advance evidence based and legitimate decision making, it is, however, important to continue developing the theoretical concepts of disease severity and investigating different ways of operationalization. Information on and the use of (operationalization of the) disease severity approaches may provide additional quantitative information on the severity of the disease in a more structured way. Therefore, (HTA) researchers and policymakers should closely work alongside each other, and make use of each other's expertise to fine tune the most useful way to operationalize disease severity in order to enhance the legitimacy of societal decision making.

## Conclusions

Our study showed that the severity of the disease is an important consideration in drug reimbursement decision making, especially in case of high severity. However, all four countries seem to struggle in making its actual role compared to other criteria explicit. Belgium and France are explicit by using the severity of the disease in setting reimbursement levels. However, all four countries could improve the transparency of the actual importance of disease severity relatively to other criteria in the decision making.

## Acknowledgements

The authors greatly appreciate the valuable input of all interviewees.

A draft version of this chapter has been discussed at the 2014 LowLand Health Economists' Study Group (LoLaHESG). The authors would like to thank Prof.dr. Johan Polder for providing valuable feedback on this chapter.

This study was performed in the context of the Escher project (T6-202), a project of Top Institute Pharma, Leiden, The Netherlands.

# 6

# Unravelling drug reimbursement outcomes: A comparative study of the role of pharmacoeconomic evidence in Dutch and Swedish reimbursement decision making

With Fredrik Nilsson, Frank Sandmann, Anthonius de Boer, and Marc Koopmanschap

PharmacoEconomics 2013; 31(9):781-797

## Abstract

To sustainably manage equitable access to effective drugs, many developed countries have established a national system to determine whether drugs should be reimbursed. Our objectives were (i) to investigate the role of pharmacoeconomic evidence in Dutch and Swedish drug reimbursement decision-making; and (ii) to determine the extent to which appraising the importance of full economic evaluations relative to other evidence is a transparent process.

We investigated all Dutch and Swedish drug reimbursement information published in the period January 2005 to July 2011. After categorising all the reimbursement applications and decisions in published data sources, we selected all dossiers –in both countries– that included a full economic evaluation (i.e., cost-effectiveness and/or cost-utility analysis) and then investigated how the evidence was appraised for its societal value.

In The Netherlands, only 35% of the 118 applications on List 1B (i.e., claiming added therapeutic value) were found to include pharmacoeconomic evidence. In all cases where drugs received a 'no' decision, combined with an evaluation that they were of similar (n = 7) or added (n = 5) therapeutic value, we found that the pharmacoeconomic evidence had been judged insufficiently robust. We also found that in 21% of the 'yes' decisions, combined with an evaluation of similar (n = 2) or added (n = 2) therapeutic value, the pharmacoeconomic evidence had been judged insufficiently robust. In Sweden, we found that drugs that received a 'no' decision (n = 39) had been judged either not cost effective (74%) or not supported by sufficiently credible data (26%). Nearly all drugs that received a 'yes' decision (n = 252) had been judged cost effective (92%). However, of all these judgements 53% were based on a price comparison and 10% on a cost-minimisation analysis; only 33% were based on a full economic evaluation. More economic evaluations were available in Sweden than in The Netherlands (97 vs. 31, respectively), mainly due to the numerous exemptions from pharmacoeconomic evidence in The Netherlands (65%). Dossiers for only 11 drugs included a full economic evaluation in both countries; of these, the reimbursement decisions differed for four drugs. Appraisal elements were reported only descriptively; their actual influence on the final decision remained unclear. In four dossiers, the (high) severity of the treatable disease was explicitly mentioned in both countries; three of these were identical and related to indications in cancer.

Both countries publish drug reimbursement information. Therapeutic value appears to be the most decisive criterion; the relative importance of full economic evaluations is more modest than would generally be expected, especially in The Netherlands. Although the assessment process is reasonably transparent, both countries could make the appraisal process more transparent by more explicitly showing the actual role of each different (societal) criterion in their decision making.

## Introduction

Globally, increasing expenditures are placing health care systems under continual pressure. In the past two decades, in order to manage equitable access to effective drugs while controlling expenditure, many developed countries have established agencies to determine whether the costs of new drugs should be reimbursed. For each new drug, these agencies are responsible for assessing the evidence (i.e., quantification of evidence) of its therapeutic value and appraising (i.e., the process of valuing assessment outcomes and weighing them against other criteria) whether it warrants being paid for by society.

Previous publications in this area have provided a global overview of the various reimbursement criteria used in different countries;[81,86,87,155,156] compared reimbursement outcomes for specific groups of drugs;[157-159] compared health technology assessment (HTA) recommendations in England, Scotland, Sweden, Canada and Australia;[160] compared the contribution of economic analysis in the French and Scottish systems;[161] and quantitatively analysed reimbursement outcomes (for example, UK National Institute for Health and Clinical Excellence [NICE][162,163] or Australian Pharmaceutical Benefits Advisory Committee [PBAC][164] decision making). However, reimbursement assessment and appraisal processes and their outcomes have seldom been subjected to in-depth cross-country comparison combined with analysis of countries' evidence bases and consideration of potential differences between their reimbursement policies.

A recent study described five European drug reimbursement systems and comparatively analysed their objectives, institutions, processes, formal reimbursement criteria, and output and real-life outcomes.[15,103,130] To further investigate the extent to which policy differences result in differences in reimbursement outcomes and, more specifically, to determine the role of pharmacoeconomic evidence in decision making, this present research compares Dutch and Swedish drug reimbursement decisions in the period January 2005 to July 2011. We used a two-pronged approach. First, we retrospectively analysed all reimbursement dossiers published during that period in the two countries, thereby investigating the reimbursement decisions and the assessment of the underlying evidence. Second, for all dossiers that included a full economic evaluation in both countries, we determined the extent to which the decision-making process was transparent, by exploratively analysing the evidence base, the assessment, the appraisal, and the relative importance of the reported decision criteria.

## Context

### Drug reimbursement in The Netherlands

In The Netherlands, The Health Care Insurance Board (College voor Zorgverzekeringen [CVZ]) is legally responsible for managing entitlements in the basic benefit package, including drugs. A technical department prepares assessment reports, which are then evaluated by the Scientific Pharmaceutical Advisory Commission (Wetenschappelijke Adviesraad Geneesmiddelen [WAR]; formerly Commissie Farmaceutische Hulp [CFH]). On the request of the CVZ Board of Directors, reimbursement applications may also be evaluated by the Appraisal Commission (Adviescommissie pakket [ACP]), though this seldom occurred in the first few years following the commission's establishment in 2008. In appraising the reimbursement request, this commission considers a broader societal perspective before advising the CVZ Board. The Board then forwards the definitive reimbursement advice to the Minister of Health, who makes the final decision. Regarding pharmacoeconomic evidence, the CVZ and the WAR formally advise only on its robustness. The advisory reports are published online in Dutch and include the assessment, summaries of the evidence, and reasons for the advice.

Application requirements differ depending on the type of drug. Applications for expensive inpatient drugs require information on therapeutic value, a cost prognosis and a research plan for outcomes research. Outpatient drugs fall under the drug reimbursement system (Geneesmiddelen Vergoedingssysteem [GVS]), which consists of List 1A (i.e., groups of therapeutically interchangeable drugs), List 1B (i.e., non-interchangeable drugs and/or drugs with added therapeutic value), and List 2 (i.e., drugs restricted in reimbursement). For drugs on List 1A, the reimbursement level is limited to the group's historically determined average product price. Drugs on List 1B are fully reimbursed. Applications for List 1A only require pharmacotherapeutic evidence; applications for List 1B also require a budgetary impact estimation and, since 2005, pharmacoeconomic evidence.

### Drug reimbursement in Sweden

Sweden's Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket [TLV]) is legally responsible for deciding, by means of a value-based pricing approach as well as taking a societal perspective, whether outpatient drugs are to be included in the pharmaceutical benefits scheme. County councils are responsible for procurement and reimbursement of inpatient drugs. The Agency's Reimbursement Application Unit prepares assessment reports; these are evaluated by the Agency's Pharmaceutical Benefits Board (PBB), which makes the final reimbursement decision. The Agency's Director-General makes decisions, without consulting the PBB, on requests for

price changes, new strengths, generic drugs, new package sizes, and parallel-imported drugs. Decisions are published online in Swedish and include summaries of the evidence and reasons for the decisions.

Application requirements differ depending on the type of application. The two main application categories are (i) generics, and (ii) new pharmaceuticals (including important new indications). The requirements are most extensive for new pharmaceuticals; the procedure for generics is simpler. If a drug is judged to be clinically similar to the comparator, a cost-minimisation analysis or price comparison may suffice. If the manu-

**Table 6.1** Summary of characteristics of the Dutch and Swedish drug reimbursement system

| | **The Netherlands** | **Sweden** |
|---|---|---|
| **System characteristics** | | |
| National reimbursement agency | CVZ | TLV |
| Pharmaceutical committee | WAR (and ACP) | PBB |
| Scope of national agency | Outpatient and expensive inpatient drugs | Outpatient drugs |
| Final decision maker | Minister of Health | PBB |
| Implementation of the outcome | Positive list | Positive list |
| Reimbursement reports publicly available | Yes, CVZ's advisory report to minister | Yes, TLV's decision including summary of evidence and motivation of decision |
| **Assessment and appraisal criteria (national level)** | | |
| Medical, therapeutic and societal need | Yes | Yes |
| Added therapeutic value | Yes | Yes |
| Cost-effectiveness | Yes | Yes |
| Budget impact | Yes | No |
| Explicit threshold for cost/ QALY | No | No |
| Other | Disease severity and rarity, own responsibility, accessibility, societal affordability, public health | All effects on a person's health and quality of life, human value, need and solidarity |
| **Re-evaluation/ revision** | | |
| Re-evaluation/ revision of previous decision | Yes, expensive inpatient drugs after 4 years | Yes, decided case-by-case |
| | Outpatient drugs only ad-hoc (from 2013 case-by-case) | Therapeutic areas in which drugs do not appear to be used in a cost-efficient way |

*CVZ = Health Care Insurance Board [in Dutch: College voor Zorgverzekeringen]*
*WAR = Scientific Pharmaceutical Advisory committee [in Dutch: Wetenschappelijke Adviesraad Geneesmiddelen]*
*ACP = Appraisal committee [in Dutch: Adviescommissie Pakket]*
*TLV = Dental and Pharmaceutical Benefits Agency [in Swedish: Tandvårds- och läkemedelsförmånsverket]*
*PBB = Pharmaceutical Benefits Board [in Swedish: Nämnden för läkemedelsförmåner]*

facturer claims an added therapeutic value, evidence is required on that value, as well as on cost-effectiveness.

County councils are responsible for implementing national (compulsory) reimbursement decisions; each county can establish its own drug therapeutic committee and own therapeutic guideline. Outpatient drugs are fully reimbursed after a maximum co-payment level has been reached. Table 6.1 provides an overview of the characteristics of the Dutch and Swedish drug reimbursement systems. It should be noted that the requirements that both countries place on pharmacoeconomic evidence are similar, as is their adoption of a societal perspective (e.g., inclusion of productivity costs) in the decision making. For more detailed information on these two systems and their HTA models, see Franken et al.[103] and Le Polain et al.[15]

## Methods

We investigated whether and how the Dutch and Swedish national agencies responsible for deciding on reimbursement of new drugs actually used pharmacoeconomic evidence in their decision making. We also assessed the extent to which the process of establishing its importance relative to other criteria was transparent. These two countries were selected because, in both cases, their responsible agencies formally use drug effectiveness and cost-effectiveness as reimbursement criteria. For both countries, we investigated all reimbursement dossiers published in the period January 2005 to July 2011. It is important to emphasise that we only used published documentation. Our analysis started in 2005 because that was the first year in which pharmacoeconomic evidence was required for reimbursement decision making in The Netherlands.

Reimbursement decisions and the reasons given for these decisions were investigated as follows. First, for new drugs, we compared all the reimbursement applications and decisions and sorted them into categories. For those categories in which it was reasonable to expect that pharmacoeconomic evidence had been used (i.e., outpatient drugs with an added therapeutic value claim in The Netherlands, and all new pharmaceuticals in Sweden), we investigated the reimbursement decisions to determine whether pharmacoeconomic evidence had been included, and if so, we identified the type of evidence. If no pharmacoeconomic evidence had been included, we retrieved any reported underlying reasons.

Second, to determine the extent to which establishing the importance of full economic evaluations –relative to other criteria considered in the decision making– was a trans-

parent process, for a group of selected drugs we exploratively analysed the evidence-base, the assessment, the appraisal, and the reported relative importance of the decisive criteria. We selected those dossiers that, in both countries, provided a full economic evaluation (i.e., cost-effectiveness, cost-benefit and/or cost-utility analysis[165]); dossiers including only a price comparison or cost-minimisation analysis were not selected. Dossiers on vaccines were also excluded.

To analyse the assessment, we compared the Dutch and Swedish evaluations of the pharmacotherapeutic and pharmacoeconomic evidence and retrieved other factors that may have influenced the decision. Regarding pharmacotherapeutic evidence, we extracted data on therapeutic value, comparator, efficacy, adverse effects, other clinical effects, and pharmacotherapeutic uncertainty, as well as information on underlying studies. Regarding pharmacoeconomic evidence, we extracted data on the type of economic evaluation, incremental effects, incremental costs, type of comparison (i.e., indirect vs. direct), costs per intermediate outcome, costs per life-year gained (LYG), costs per quality-adjusted life-year (QALY), and pharmacoeconomic uncertainty, as well as information about the economic evaluation (e.g., applied model). To allow easy comparison, Swedish kronor (SEK) were converted to Euros (€) using the European Central Bank's average exchange rate over the period January 2005 until July 2011 (€1.00 = SEK9.5379).

Finally, as both countries adopt a societal perspective in their decision making, in order to analyse the appraisal of a drug's societal value, we extracted reported information on consultation with stakeholders, product price, expected number of eligible patients, budgetary impact, added therapeutic value, cost-effectiveness, medical and therapeutic need, as well as any other reported societal criteria (e.g., disease severity, human value, public health, accessibility, affordability, solidarity, rarity, alternative treatments, and incentive for innovation).

## Results

### Reimbursement applications and decisions in The Netherlands

Between January 2005 and July 2011, The Netherlands published 311 advisory reports for new drugs: 58 expensive inpatient and 253 outpatient drugs (see Figure 6.1). For general outpatient drugs (n = 186), 68 were List 1A and 118 were List 1B applications. Of the List 1A applications (n = 68), four were subsequently placed on List 1B, 60 on List 1A, and four were rejected for reimbursement. Of the List 1B applications (n = 118), 63 were subsequently placed on List 1B, 29 on List 1A, and 26 were rejected for reimbursement.

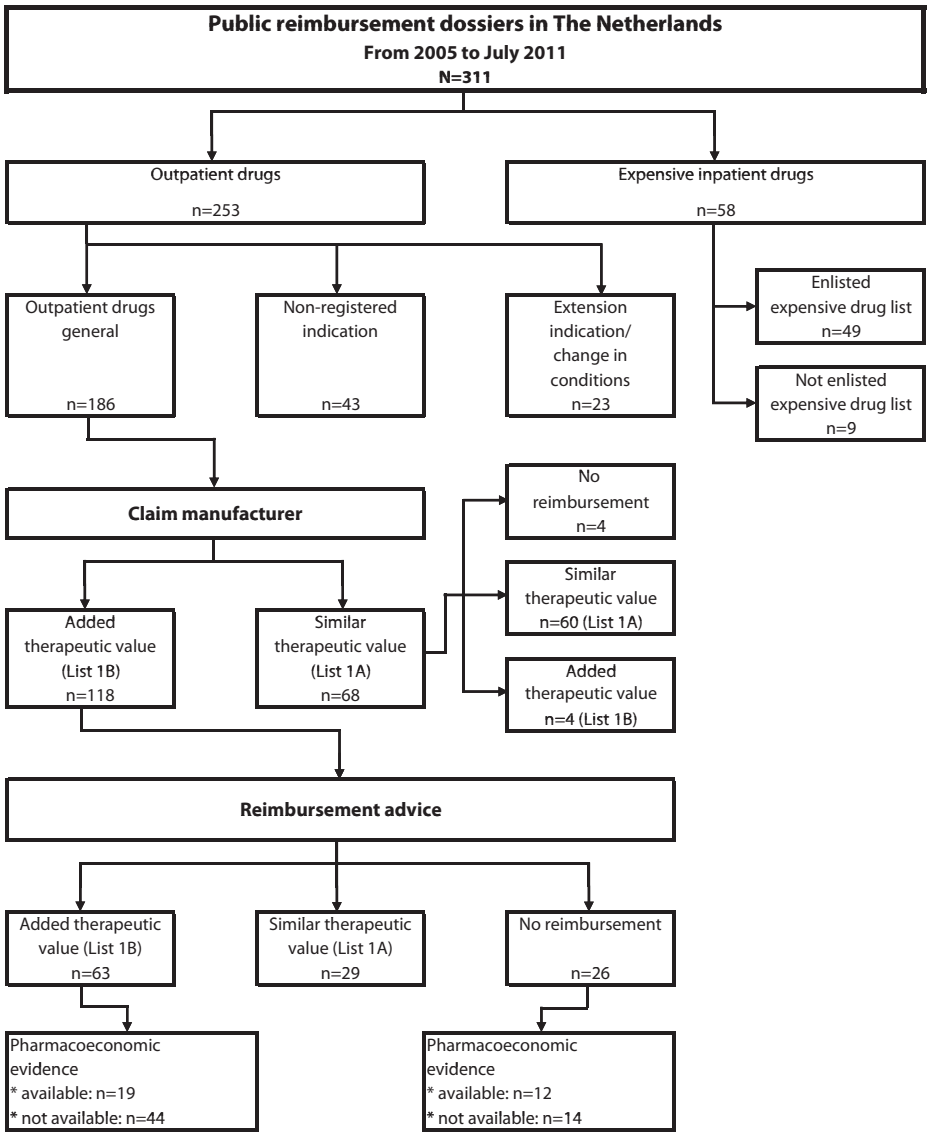


**Figure 6.1** Overview of reimbursement applications and outcomes in The Netherlands

**Table 6.2** Pharmacoeconomic evidence and its evaluation in The Netherlands

| **Application** | **List 1B** | **List 1B** |
|---|---|---|
| **Reimbursement advice** | **no listing** | **List 1B** |
| Total number of reimbursement dossiers | 26 | 63 |
| **No pharmacoeconomic evidence available** | 14 (54%) | 44 (70%) |
| Exemption due to orphan status [a] | 1 | 24 |
| Exemption due to being a HIV drug [a] | | 7 |
| Exemption due to estimated budget impact < €500,000 [a] | | 4 |
| Small number of expected eligible patients [a] | 1 | 1 |
| Similar/ lower therapeutic value and higher costs [b] | 12 | |
| Similar therapeutic value, no higher costs [b] | | 1 |
| Added therapeutic value and budget neutral or lower costs [b] | | 3 |
| Application before 01-01-2005 (evaluated in 2005) | | 4 |
| **Pharmacoeconomic evidence available** | 12 (46%) | 19 (30%) |
| Type of pharmacoeconomic evidence | | |
| *Cost-minimisation analysis* | 1 | 3 |
| *Cost-effectiveness analysis* | 3 | 3 |
| *Cost-utility analysis* | 4 | 4 |
| *Cost-effectiveness & cost-minimisation analysis* | | 1 |
| *Cost-effectiveness & cost-utility analysis* | 4 | 8 |
| CE plane and acceptability curve available | 3 | 9 |
| Either CE plane or acceptability curve available | 3 | 2 |
| Evaluation pharmacotherapeutic and economic evidence | | |
| *Similar therapeutic value* | 7 | 4 |
| Pharmacoeconomic evidence | | |
| Sufficiently robust | | 2 |
| Insufficiently robust | 7 | 2 |
| *Added therapeutic value* | 5 | 15 |
| Pharmacoeconomic evidence | | |
| Sufficiently robust | | 7 |
| Reasonably robust | | 4 |
| Moderately robust | | 2 |
| Insufficiently robust | 5 | 2 |

*List 1B: not interchangeable within a cluster/ added value*
*CE plane: Cost-Effectiveness plane; HIV: Human Immunodeficiency Virus*
[a] *Exempted as described by Borst 2000, Cheung 2011 and van der Meijden 2011*
[b] *Exempted as described by documentation of the Dutch Health Care Insurance Board (not publicly available)*

In The Netherlands, pharmacoeconomic evidence is only evaluated and published if a drug is judged to have added therapeutic value and is placed on List 1B. Therefore, none of the original List 1B applications that were placed on List 1A (n = 29) included pharmacoeconomic evidence. Interestingly, such evidence was included in only 19 of the 63 'yes' (30%) and 12 of the 26 'no' (46%) decisions (List 1B applications). The relevant regulations[19,166,167] formally exempted certain drugs, such as orphan drugs (n = 25), and HIV drugs (n = 7) (see Table 6.2). Available pharmacoeconomic evidence was presented as a cost-minimisation analysis (n = 5), a cost-effectiveness analysis (n = 19) and/or a cost-utility analysis (n = 20).

With regards to all 12 'no' decisions that included pharmacoeconomic evidence, that evidence had been evaluated as insufficiently robust and the judgement was that introduction of the drug would increase the pharmaceutical budget. Five of these drugs had been judged to be of added therapeutic value and seven to be of similar therapeutic value, yet they received a 'no' decision, for reasons that were not explicitly stated. Three applications concerned two vaccines; their dossiers discussed an inconsistency with policy, as vaccines should be part of the national vaccination programme (i.e., a separate financial budget). Notably, one drug (i.e., ivabradine) received a 'yes' decision a year later on the basis of new pharmacoeconomic evidence.

Of the 19 'yes' decisions that included pharmacoeconomic evidence, 15 of the drugs had been judged to be of added therapeutic value, and introduction of the drug would increase the pharmaceutical budget. Interestingly, in the case of four dossiers (two indicating added and two indicating similar therapeutic value), the drugs obtained a 'yes' decision (21%) even though the pharmacoeconomic evidence had been judged to be insufficiently robust.

Table 6.2 provides an overview of the publication of this pharmacoeconomic evidence and its evaluation in The Netherlands.

**Reimbursement applications and decisions in Sweden**

Between January 2005 and July 2011, Sweden's TLV made 291 decisions on new pharmaceuticals: of these, 39 were rejected for reimbursement and 252 were accepted for reimbursement (see Figure 6.2). Of the accepted applications, 46 were restricted in reimbursement (e.g., in most cases to a subgroup of patients); of these 46 applications, 15 dossiers also included either conditions on evidence development (n = 9) or time restrictions (n = 6). A total of 22 were granted 'general' reimbursement conditional on evidence development.

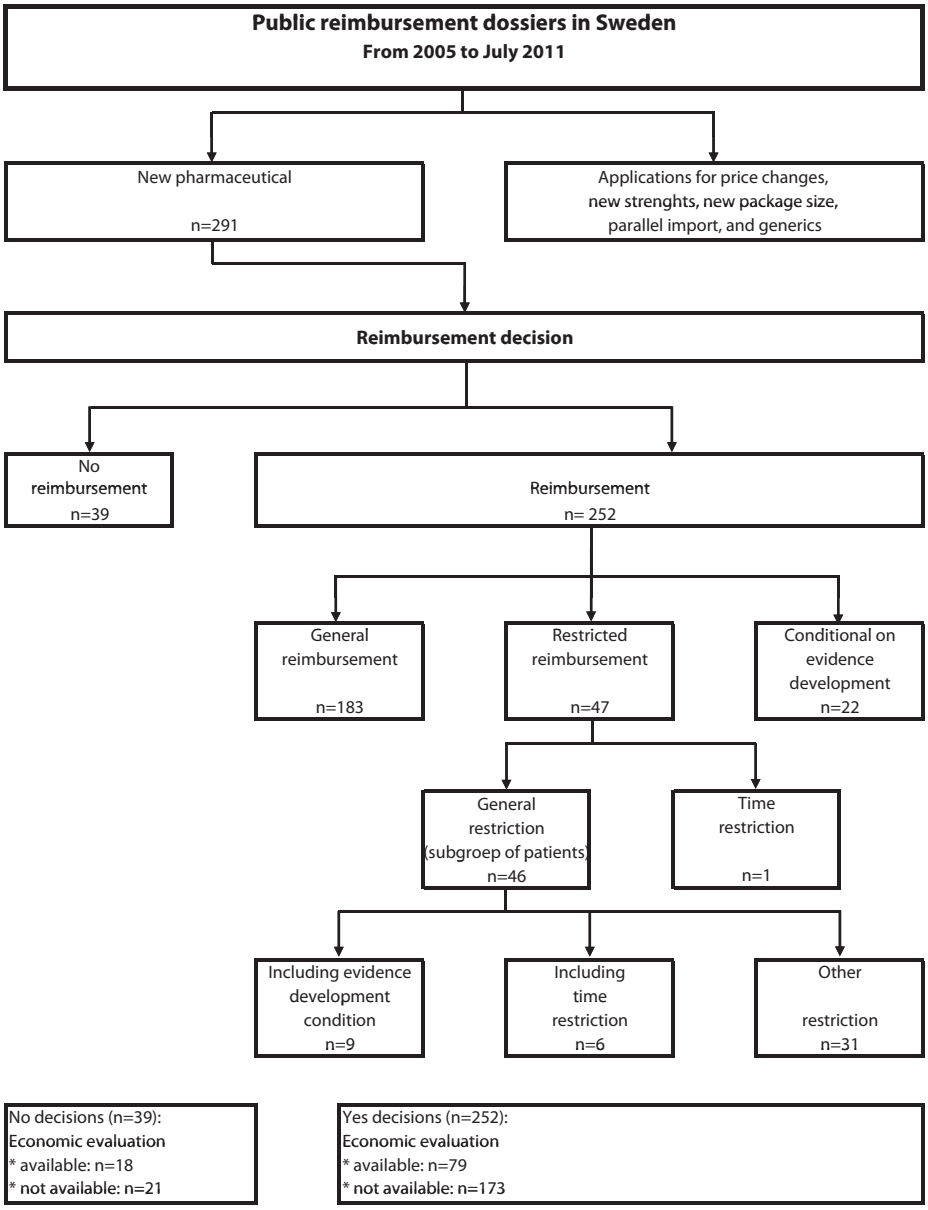


**Figure 6.2** Overview of reimbursement applications and outcomes in Sweden

Nearly two thirds of all these decisions relied on price comparisons (53%) or cost minimisation analyses (10%); only 33% were based on a full economic evaluation. In ten cases, the decision was not based on any economic evidence and was therefore classified as 'other'; most of these decisions referred to high disease severity or lack of alternative treatments.

Regarding the underlying reasons for the 'yes' decisions (n = 252), 60% of the drugs had been judged to be similar in clinical effect to the comparator and to be cost effective (i.e., having similar or lower costs); 31% were accepted for reimbursement on the basis of a full economic evaluation. A total of 91% of the full economic evaluations were part of judgements in the 'better' clinical effects category, 4% in the 'similar' category, and 3% in the 'worse' category. Three drugs were judged to be 'worse' in effect than the comparator but cost effective due to a sufficiently lower price. Interestingly, eight 'yes' decisions were not based on any pharmacoeconomic analysis, though the drugs were accepted for reimbursement because the costs were judged to be reasonable. The 11 decisions classified as 'other' refer to a high disease severity combined with a lack of alternative treatments; one decision was based on a health economic evaluation developed for British conditions.

Regarding the underlying reasons for the 'no' decisions (n = 39), 51% of the drugs were judged similar in clinical effect to the comparator but not cost effective (i.e., higher price or costs); 46% were rejected as not cost effective based on a full economic evaluation. Thirty-three percent of the full economic evaluations were part of judgements in the 'better' clinical effects category, 11% in the 'worse' and 6% in the 'similar' category. In ten decisions (26%), either the clinical or the economic evidence was judged as lacking in credibility. Table 6.3 provides an overview of the reimbursement outcomes and type of pharmacoeconomic evidence in Sweden.

**Comparative analysis of a selection of drugs**

We found that the reimbursement dossiers for only 11 drugs included a full economic evaluation (i.e., cost-effectiveness and/or cost-utility analysis[165]) in both countries. In seven of these cases, the reimbursement decision was similar, whereas in four cases it differed. Notably, all four differing decisions were restricted or conditional in at least one of the countries. Table 6.4 provides an overview of the results of this comparative analysis.

*Pharmacotherapeutic evidence (11 cases)*

All the dossiers that we investigated included information on pharmacotherapeutic evidence, and all but one (Sweden: inhaled insulin) included a description of the under-

**Table 6.3** Reimbursement outcomes and type of pharmacoeconomic evidence in Sweden

| **Total number of applications (new pharmaceuticals)** | 291 | | | |
|---|---|---|---|---|
| Type of pharmacoeconomic evidence available | | | | |
| Economic evaluation | | 97 | (33%) | |
| Cost minimisation analysis | | 29 | (10%) | |
| Price comparison | | 155 | (53%) | |
| Other | | 10 | (3%) | |
| **Yes reimbursement decision** | 252 | (87%) | | |
| General reimbursement | | 183 | (73%) | |
| Type of pharmacoeconomic evidence | | | | |
| *Economic evaluation* | | | 42 | |
| *Cost minimisation analysis* | | | 16 | |
| *Price comparison* | | | 120 | |
| *Other* | | | 5 | |
| Restricted reimbursement | | 46 | (18%) | |
| Including time restrictions | | | 6 | |
| Including evidence development conditions | | | 9 | |
| Type of pharmacoeconomic evidence | | | | |
| *Economic evaluation* | | | 23 | |
| *Cost minimisation analysis* | | | 8 | |
| *Price comparison* | | | 11 | |
| *Other* | | | 4 | |
| Time restricted reimbursement | | 1 | (0.4%) | |
| Type of pharmacoeconomic evidence | | | | |
| *Economic evaluation* | | | 1 | |
| Conditional on evidence development | | 22 | (9%) | |
| Type of pharmacoeconomic evidence | | | | |
| *Economic evaluation* | | | 13 | |
| *Cost minimisation analysis* | | | 3 | |
| *Price comparison* | | | 5 | |
| *Other* | | | 1 | |
| **No reimbursement decision** | 39 | (13%) | | |
| Type of pharmacoeconomic evidence | | | | |
| Economic evaluation | | | 18 | |
| Cost minimisation analysis | | | 2 | |
| Price comparison | | | 19 | |
| **TLV judgement** | | | ***Judgement based on a full economic evaluation*** | |
| Yes reimbursement decision | 252 | | 79 | (31%) |
| Similar clinical effects and cost effective | | 152 | | 3 |
| Better clinical effects and cost effective | | 78 | | 72 |
| Worse clinical effects but cost effective | | 3 | | 2 |
| No economic analysis but reasonable costs | | 8 | | |
| Other | | 11 | | 2 |
| No reimbursement decision | 39 | | 18 | (46%) |
| Similar clinical effects but not cost effective | | 20 | | 1 |
| Better clinical effects but not cost effective | | 7 | | 6 |
| Worse clinical effects and not cost effective | | 2 | | 2 |
| No credible economic evidence nor clinical effects | | 10 | | 9 |

**Table 6.4** Comparison between The Netherlands and Sweden of eleven drugs

| Drug/ country (decision date) | Decision | | | Pharmacotherapeutic evidence | | | | Pharmacoeconomic evidence | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Restrictions (RE) & conditions (CO) | | | Comparator | | Judgement therapeutic value | | Type of evidence | Comparator | | Cost/QALY (€) | | Judgement cost-effectiveness | |
| | | RE | CO | C1 | C2[a] | C1 | C2 | | C1 | C2[a] | C1 | C2 [a] | C1 | C2 |
| **Rimonabant** | | | | | | | | | | | | | | |
| NL (Mar 07) | N | | | D+E | SIB | = | = | CEA+ CUA | D+E | | € 20,926 | € 22,538 | NS | |
| SW Nov 06) | Y | Y[c d] | Y[f g] | SIB and ORL | Advice & lifestyle | = | + | PC+ CUA | reductil & xenical | Advice & lifestyle | | | Similar CE | Reas CE |
| **Ticagrelor** | | | | | | | | | | | | | | |
| NL (May 11) | Y | Y[c e] | N | CLO | PRA | + | n.r. | CEA+ CUA | CLO | | € 7,965 | | Suff | |
| SW (Jun 11) | Y | N | N | CLO | PRA | + | + | CUA | CLO | PRA | € 15,727 | € 18,348 | CE | CE |
| **Exenatide** | | | | | | | | | | | | | | |
| NL (Sep 07) | N | | | NPH INS | | = | | CEA+ CUA | NPH INS | INS glargin | € 17,979 | € 11,142 | NS | |
| NL (Jan 09) | Y | Y[c] | N | NPH INS | | + | | CEA+ CUA | NPH INS | | € 5,231 | | Suff | |
| SW (Jun 07) | Y | N | Y[f] | PL (+MET and/ or SU) | INS analogues | + | + | CUA | INS analogues | | € 13,315 | | CE | |
| **Varenicline** | | | | | | | | | | | | | | |
| NL Mar 08) | N | | | Nicotin replacer | BUP | = | = | CEA+ CUA | BUP | Nicotin replacers | € 1,520 | € 1,720 | NS | |
| SW (Jul 07) | Y | Y[e] | Y[f] | BUP | PL | + | + | CUA | BUP | | € 10,484 | | CE | |
| **Inhaled insulin** | | | | | | | | | | | | | | |
| NL (Feb 07) | Y | Y [c] | N | SC INS | | + | | CUA | DMI INS | DMII mix INS | € 44,596 | € 35,184 | NS | |

**Table 6.4** Comparison between The Netherlands and Sweden of eleven drugs (continued)

| Drug/ country (decision date) | Decision | | | Pharmacotherapeutic evidence | | | | Pharmacoeconomic evidence | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Restrictions (RE) & conditions (CO) | | | Comparator | | Judgement therapeutic value | | Type of evidence | Comparator | | Cost/QALY (€) | | Judgement cost-effectiveness | |
| | | RE | CO | C1 | C2[a] | C1 | C2 | | C1 | C2[a] | C1 | C2 [a] | C1 | C2 |
| SW (Oct 06) | Y | Y[c d] | Y[f] | SC INS DMI | SC INS DMII | = | = | CUA | SC INS DMI | SC INS DMII | Very low | Cost saving to € 62,906 | CE | CE |
| **Sitagliptin** | | | | | | | | | | | | | | |
| NL (Jun 07) | Y | Y[c] | N | SU derivates | thiazolidine-diones | = | = | CEA+ CUA | ROS | | Dom | | NS | |
| SW (Jun 07) | Y | N | Y[f] | PL + MET/ PIO | SU+ MET | + | + | CUA | SU | ROS | € 15,727 | Dom | CE | CE |
| **Ivabradine** | | | | | | | | | | | | | | |
| NL (Apr 06) | N | | | LA nitrates | calcium antagonist | + ^ | + ^ | CEA | Revasc (admission) | | Dom | | NS | |
| NL (Jun 07) | Y | Y[c] | N | LA nitrates | calcium antagonist | + | + | CEA | Revasc (time CABG/ PCI) | | Dom | | Mod Suff | |
| NL (Apr 11) | Y[b] | Y[c] | N | LA nitrates | calcium antagonist | + ^ | + ^ | n/a | | | | | | |
| SW (Dec 07) | N | | | Revasc | | < | | CUA | Revasc | | € 26,211 to € 57,665 saving per QALY lost | | Not CE | |
| SW (Dec 08) | N | | | No tx (refractory) | | < | | CUA | No tx | | | | No credible data | |
| **Methylnatrexone** | | | | | | | | | | | | | | |
| NL (Oct 08) | Y | N | N | PL | | + | | CEA+ CUA | Supportive care | | € 33,464 | | Suff | |
| SW (Feb 09) | Y | N | N | Laxatives | | + | | CUA | Laxatives | | € 38,373 | | CE | |

**Table 6.4** Comparison between The Netherlands and Sweden of eleven drugs (continued)

| Drug/ country (decision date) | Decision | | | Pharmacotherapeutic evidence | | | | Pharmacoeconomic evidence | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Restrictions (RE) & conditions (CO) | | | Comparator | | Judgement therapeutic value | | Type of evidence | Comparator | | Cost/QALY (€) | | Judgement cost-effectiveness | |
| | | RE | CO | C1 | C2[a] | C1 | C2 | | C1 | C2[a] | C1 | C2 [a] | C1 | C2 |
| **Dasatinib** | | | | | | | | | | | | | | |
| NL (Mar 07) | Y | N | N | HD IMA | | + | | CEA+ CUA | HD IMA | | Dom | | NS | |
| SW (Marc 07) | Y | N | N | No tx | | + | | CUA | IMA | No tx | € 6,836 | € 70,246 | CE | CE |
| **Erlotinib** | | | | | | | | | | | | | | |
| NL (Mar 06) | Y | N | N | DOC | PEM | + | + | CMA+ CEA | DOC | Supportive care | | | Reas Suff | |
| SW (Oct 05) | Y | N | N | DOC | No tx | = | + | CUA | DOC | | Similar | | CE | |
| **Bortezomib** | | | | | | | | | | | | | | |
| NL (Sep 07) | Y | Y[c] | N | THL & DXM | LEN & DXM | + | + | CEA+ CUA | THL & DXM | LEN & DXM | € 34,267 | -€ 290,158 | Mod Suff | |
| SW (Feb 07) | Y | N | N | DXM | | + | | CUA | HD THL | HD DXM | € 40,890 | € 65,528 | CE | CE |

*BUP* bupropion, *CA* calcium antagonist, *CABG* coronary artery bypass grafting, *CE* cost effective, *CEA* cost-effectiveness analysis, *CLO* clopidogrel, *CMA* cost-minimisation analysis, *CUA* cost-utility analysis, *C1/2* comparator 1/2, *D&E* diet and exercise, *DMI/II* diabetes mellitus type I/II, *DOC* docetaxel, *Dom* dominant, *DXM* dexamethasone, *HD* high-dose, *IMA* imatinib, *INS* insulin, *LA* long acting, *LEN* lenalidomid, *MET* metformin, *Mod* moderately, *N* no, *NA* not applicable, *NL* The Netherlands, *NPH* neutral protamine Hagedorn, *n.r.* not reported, *NS* not sufficient, *ORL* orlistat, *PC* price comparison, *PCI* percutaneous coronary intervention, *PEM* pemetrexed, *PIO* pioglitazone, *PL* placebo, *PRA* prasugrel, *QALY* quality-adjusted life-year, *Reas* reasonably, *Revasc* revascularisation, *ROS* rosiglitazon, *SC* subcutaneous, *SIB* sibutramine, *Suff* sufficient, *SU* sulfonylurea drugs, *SW* Sweden, *THL* thalidomid, *tx* treatment, *Y* yes

[a] Only the first two comparators are reported [b] 'Yes' to reimbursement (similar to 2007 application), but no extension of indication

*RC* Restrictions & CO conditions: [c] Specific indication/subgroup of patients; [d] In time; [e] Other; conditions: [f] Additional data collection; [g] Marketing and/or information

Therapeutic value: +, added value; =, similar value; <, lower value; ^, only for subgroup of patients

lying studies. All the Dutch dossiers included detailed information on underlying studies (stating, at least, name of study, patient numbers, follow-up, and literature references), whereas only three Swedish dossiers included such detailed information. Comparators were always reported and, in eight cases, at least one was the same. Regarding the best therapeutic value judgement, irrespective of comparator, Sweden reported nine added, one similar and one lower; The Netherlands reported eight added and three similar in this respect. In three cases (i.e., varenicline, inhaled insulin, sitagliptin) using similar comparators, the therapeutic value judgement differed. First, varenicline received a 'no' decision in The Netherlands, based on similar therapeutic value and insufficiently robust economic evidence. In contrast, it received a 'yes' decision in Sweden, based on added therapeutic value and cost-effectiveness of treatment, though the decision was conditional on additional data collection. Second, inhaled insulin obtained a 'yes' decision in both countries based on similar therapeutic value and cost-effectiveness of treatment in Sweden, but based on added therapeutic value and insufficiently robust economic evidence in The Netherlands. Both countries restricted reimbursement to a subgroup of patients; in addition, Sweden made the decision temporary, requiring additional data. Third, sitagliptin received a 'yes' decision in both countries, based on added therapeutic value and cost-effectiveness of treatment in Sweden, but based on similar therapeutic value and insufficiently robust economic evidence in The Netherlands. The Dutch agency restricted reimbursement to a subgroup of patients, whereas the Swedish agency made the decision conditional on data collection.

*Pharmacoeconomic evidence (11 cases)*

All the dossiers that we investigated included pharmacoeconomic evidence and information on input variables. In all dossiers, the comparators were similar. In Sweden, incremental effects and costs were separately reported in six and four dossiers, respectively. Costs per QALY were reported for nine drugs; one dossier only reported *"similar cost-effectiveness"*. Cost savings were reported in two cases. In The Netherlands, all dossiers included information on incremental effects and costs. Costs per LYG and per QALY were reported in six and seven dossiers, respectively. Five did not report costs per QALY or LYG due to 'dominance'.

Both countries require uncertainty analysis using, for example, cost-effectiveness planes and acceptability curves. In Sweden, uncertainty was only descriptively reported (e.g., specification of uncertain items; judgement: high or low uncertainty). In contrast, all Dutch dossiers included detailed information on uncertainty, such as sensitivity analyses outcomes (all dossiers), cost-effectiveness planes (ten dossiers) and acceptability curves (eight dossiers).

In Sweden, ten drugs that received 'yes' decisions had been judged cost effective (incremental cost-effectiveness ratios [ICERs] ranged from €6,836 to €40,890/QALY, and *"similar"* to the comparator, having *"very low"* incremental costs, one case reported a range from cost-saving to an additional cost of €62,906/QALY). One drug that received a 'no' decision had been judged not cost effective, with savings of €26,211 to €57,665 per QALY lost (i.e., south-west quadrant of the cost-effectiveness plane).

In The Netherlands, of the nine drugs that received 'yes' decisions, in three cases the pharmacoeconomic evidence had been judged to be sufficiently robust (ICERs ranged from €5,231 to €33,464/QALY), in one reasonably robust (ICER: 'dominant'), in two moderately robust (ICERs: 'dominant' and €34,267/QALY), and in three insufficiently robust (ICERs: 'dominant' [n = 2] and €44,596/QALY). Of the four drugs that received 'no' decisions, in each case the pharmacoeconomic evidence had been judged insufficiently robust (ICERs ranged from ' dominant' to €20,926/QALY). Due to the limited number of dossiers investigated, it was not possible to establish a relationship between the quality of economic evidence and the final decision.

*Appraisal elements (11 cases)*
All Swedish dossiers included information on consultations with county councils. Although it is Dutch policy to obtain the opinions of various stakeholders (e.g., patient, professional, and health insurer associations), these were only reported for six drugs. None of the Dutch dossiers were evaluated by the appraisal committee (ACP).

Budgetary consequences including the expected number of eligible patients were reported in all Dutch dossiers. Although all Swedish dossiers included the product price, none included information on budgetary consequences (budgetary impact is not a decision criterion in Sweden).

Appraisal elements were only descriptively reported; their actual influence on the final decision was not clarified. Most dossiers (10 of the 11 drugs in both countries) included a description of the disease. Although no quantifications of disease severity were incorporated, the fact that the disease in question was (highly) severe was reported in four cases in both countries. In three cases, the drug was identical, and all three concerned 'yes' decisions for indications in cancer. Other reported appraisal elements were public health reasons, new treatment mechanism, medical need for additional treatment options, limited treatment possibilities, fatality of disease, comparators not reimbursed, reduced risk factors of severe diseases, and necessity to insure (i.e., whether the illness or the care needed justified solidarity).

*Similar decision (seven cases)*
Seven drugs received a 'yes' decision in both countries. The therapeutic value judgement (added value) was the same in four cases even though three used different comparators. In Sweden, all four were judged to be cost effective treatments, whereas in The Netherlands the pharmacoeconomic evidence was judged sufficiently (n = 2), moderately (n = 1), or insufficiently (n = 1) robust. The other three decisions, while using similar comparators, differed in therapeutic value judgement (i.e., similar vs. added value). In Sweden, all three were judged cost effective, whereas in The Netherlands the pharmacoeconomic evidence was judged reasonably (n = 1) or insufficiently (n = 2) robust. In all economic evaluations, at least one of the comparators was the same.

*Diverging decision (four cases)*
For four drugs, the reimbursement decision differed.

First, in The Netherlands in 2007, rimonabant received a 'no' decision based on similar therapeutic value, insufficiently robust pharmacoeconomic evidence (€20,926/QALY), and increase of the pharmaceutical budget. The Dutch dossier reports that two comparators (sibutramine and orlistat) for the same indication (i.e., overweight and obesity) are also not reimbursed. In Sweden in 2006, using similar comparators, rimonabant obtained a 'yes' decision based on added therapeutic value (comparator 2), as well as on similar (comparator 1) and reasonable (comparator 2) cost-effectiveness (no values reported). However, reimbursement was restricted not only in time but also to a subgroup of patients and was made conditional on obligatory collection of additional data and provision of marketing information. It should be noted that, at the time of the rimonabant decision, the two comparators (sibutramine and orlistat) were reimbursed in Sweden. In 2009, European Medicines Agency (EMA) market authorisation was withdrawn; the drug was subsequently delisted.

Second, in The Netherlands in 2007, exenatide first received an initial 'no' decision for the entire diabetic population, based on similar therapeutic value, insufficiently robust pharmacoeconomic evidence (€17,979/QALY), and increase of the pharmaceutical budget. In 2009, it received a 'yes' decision, with reimbursement restricted to a subgroup of patients (i.e., patients for whom –a combination of– other diabetic treatments were inappropriate or had insufficient effects, and who had a body mass index [BMI] ≥ 35 kg/m2), sufficiently robust pharmacoeconomic evidence (€5,231/QALY), and increase of the pharmaceutical budget. In Sweden in 2007, using similar comparators, exenatide received a 'yes' decision conditional on additional data collection. It had been judged to be of added therapeutic value and to be cost effective (€13,315/QALY). Having been incorporated in a 2009 review of the diabetic therapeutic class, it is still reimbursed,

though with restrictions similar to those placed on other diabetic products (i.e., patients for whom –a combination of– other diabetic treatments were inappropriate or had insufficient effects). Re-evaluation of the requested additional data has been put on a prioritisation list. It should be noted that, due to the additional BMI restriction, the Dutch decision is more restrictive.

Third, in The Netherlands in 2007, varenicline received a 'no' decision, based on similar therapeutic value, insufficiently robust pharmacoeconomic evidence (€1,520/QALY), and increase of the pharmaceutical budget. The dossier reports that the comparator (bupropion) is also not reimbursed for the indication of smoking cessation; it also indicates the costs of varenicline treatment compared with smoking, explicitly stating that 'own responsibility' and 'co-payments' in smoking cessation are to be discussed in a follow-up report. It should be noted that the manufacturer subsequently initiated a court case, though this did not alter the decision. In Sweden in 2007, using similar comparators, varenicline received a 'yes' decision based on added therapeutic value and cost-effectiveness (€10,484/QALY). Reimbursement was restricted to second-line treatment combined with motivational therapy and was made conditional on obligatory marketing information. It should be noted that the comparator bupropion is also reimbursed in Sweden.

Fourth, in The Netherlands in 2006, ivabradine for the indication of stable angina pectoris received an initial 'no' decision, based on added therapeutic value, insufficiently robust pharmacoeconomic evidence ('dominant'), and increase of the pharmaceutical budget. In 2007, it received a 'yes' decision based on the same pharmacotherapeutic evidence combined with new pharmacoeconomic evidence (using a different comparator), which was judged to be moderately robust ('dominant'). Reimbursement was restricted to a subgroup of patients. In 2011, the manufacturer's request to loosen this restriction was denied. In Sweden, first in 2007 and then in 2008, ivabradine received two 'no' decisions. Both were based on lower therapeutic value (using a different comparator than in the Dutch dossiers); the first was also based on an actual lack of cost-effectiveness, the second on insufficient evidence of cost-effectiveness (saving of €26,211- €57,665/QALY lost; comparator in 2007 similar to that used in the Dutch economic evaluation).

## Discussion

Using official information on drug reimbursement decisions that were published in The Netherlands and Sweden in the period January 2005 to July 2011, we investigated the role of pharmacoeconomic evidence in reimbursement decision making. We also deter-

mined the extent to which the decision-making process was transparent, by comparing the evidence base, the quantitative assessment, the comparative appraisal, and the reported decision criteria.

As expected, we observed differences in processes and outcomes due to reimbursement policies. First, whereas the Dutch agency evaluates both outpatient and expensive inpatient drugs, the Swedish agency only evaluates outpatient drugs. Second, because information on budgetary impact is not a criterion at the national level in Sweden, none of the Swedish dossiers that we investigated included such information. Third, the Swedish agency judges whether a drug is cost effective and incorporates that judgement in its final reimbursement decision, whereas the Dutch agency only evaluates the pharmacoeconomic evidence for its robustness and accordingly advises the Minister of Health, who makes the final decision. Fourth, in Sweden, consultation with stakeholders is limited to county councils, whereas the Dutch policy is to obtain the opinions of a wide range of stakeholders. Finally, Swedish decisions may be made conditional on the obligatory provision of additional data where outcomes are uncertain, whereas in The Netherlands this practice did not occur for outpatient drugs during the period of our analysis. It should be noted that, at the time of writing, The Netherlands is planning to expand reimbursement conditional on evidence development to specific groups of outpatient drugs in 2013.[168]



We also expected to observe less use of pharmacoeconomic evidence in The Netherlands than in Sweden, partly because, in the former, new drugs on List 1A (i.e., groups with therapeutically interchangeable drugs) are reimbursed similarly to therapeutically equivalent existing drugs and are therefore not evaluated for cost-effectiveness. However, surprisingly, only 35% of the dossiers for Dutch List 1B drugs (i.e., non-interchangeable drugs and/or drugs with added therapeutic value: 'no' decision 46%; 'yes' decision 30%) included pharmacoeconomic evidence. In contrast, Sweden requires that all new drugs be evaluated for cost-effectiveness. Only 3% of the 291 Swedish applications included no economic evidence. However, only 33% included a full economic evaluation. Eighty percent of the Swedish economic evaluations were part of judgements in the category 'better' clinical effects. It should be noted that Sweden's system of comparing the prices of drugs is similar in its impact on the reimbursement costs to the Dutch clustering system of grouping therapeutically equivalent drugs on List 1A.

Regarding the role of pharmacoeconomic evidence in decision making in everyday practice, the Swedish 'no' decisions were based on a judgement that either cost-effectiveness (74%) or credible economic evidence (26%) was lacking; 92% of the 'yes' decisions were based on a judgement of cost-effectiveness. Even though, in principle, the Swedish

agency explicitly judges whether a drug is cost effective, it is important to realise that, in practice, most of these judgements were merely based on price comparisons (53%). Forty-six percent of the 'no' and 31% of the 'yes' decisions were based on a full economic evaluation. In one of the dossiers, cost saving at the expense of QALYs lost was judged cost-ineffective. Strikingly, most of the Swedish decisions that were not based on any economic evidence were made in 2005 or 2006.

In The Netherlands, all 'no' decisions were based on a judgement that the pharmacoeconomic evidence was insufficiently robust, combined with a judgement of either similar (n = 7) or added (n = 5) therapeutic value. More surprisingly, the same was true for 21% of 'yes' decisions. Of the drugs that received a 'yes' decision, 79% had been judged of added therapeutic value, whereas 21% had been judged of similar therapeutic value to the comparator. It should be noted that most of the unfavourable judgements regarding the robustness of pharmacoeconomic evidence in The Netherlands occurred in the early years of this requirement (2005-2007). All but one of the Dutch 'yes' decisions (nine of ten) made in the period 2008-2011 were based on a judgement that such evidence was sufficiently robust.

We found the agencies in both countries to be insufficiently transparent in two senses: (i) in not explaining how they determined the importance –for the decision– of full economic evaluations relative to other evidence; (ii) in not stipulating a threshold range of cost-effectiveness. Furthermore, when dossiers only report that a drug is 'dominant', as in The Netherlands and Sweden, or 'similarly cost effective', as in Sweden, the actual and relative differences in effects and costs are unclear.

As so few of the investigated dossiers published during a 6.5-year period actually included full economic evaluations (i.e., 31 in The Netherlands and 97 in Sweden) that were comparable, we could only include 11 drugs in our in-depth comparative analysis. This is the main limitation of our research, and the conclusions we draw from this analysis should therefore be regarded as tentative rather than definitive. Whereas economic evaluations play a more important role in Sweden (i.e., 80% of drugs with an added therapeutic value claim –combined with a premium price– included an economic evaluation) than they do in The Netherlands (i.e., 35% of list 1B applications), our results suggest that therapeutic value is the most decisive criterion, which is consistent with the findings of previous studies.[15,103,162,163]

Another limitation of our study is that we were only able to investigate published information, which may not necessarily reveal all arguments and reasoning that led to a particular decision. Both agencies are responsible for appraising whether a drug warrants

society paying for it and are thus accountable to society. That is why we investigated the transparency of their societal decision making and the reported underlying reimbursement criteria. According to Daniels and Sabin[102], a fair and legitimate prioritising procedure must satisfy four conditions: (i) transparency of the decision-making process; (ii) relevance of the decisive criteria; (iii) revisability of the decision in light of new evidence and arguments; and (iv) enforcement of the existing criteria. Our data show that both The Netherlands and Sweden generally report –albeit in varying detail– their assessment of the pharmacotherapeutic and pharmacoeconomic evidence, including references to underlying studies. However, in Sweden, drug manufacturers can withdraw their application, in which case no data are published. This practice guarantees confidentiality at the cost of transparency.[103,130]

At first glance, Sweden, in only reporting a single cost-effectiveness estimate, may appear to be less transparent than The Netherlands regarding decision-makers' uncertainty. However, the large number of such estimates in The Netherlands may not necessarily improve decision-makers' understanding of the impact of uncertainty nor reduce its consequences. The Sweden agency's practice of requesting additional data on specific items where there is (high) uncertainty contributes to the legitimacy of its decision-making process by fulfilling both the transparency and the revision requirement as defined by Daniels and Sabin. Even though the Swedish agency makes the final reimbursement decision, whereas its Dutch counterpart merely advises the Minister of Health, the former agency could consider publishing a greater amount of standardised and/or quantified information on uncertainty in order to enhance its transparency to the general public. For example, it could achieve this by publishing outcomes of sensitivity analyses or (descriptions of) cost-effectiveness planes or acceptability curves. The Swedish agency could also consider the advantages of more frequently reporting incremental effects and costs separately. In accordance with research results on the limited utility of small increases in life expectancy reported by Kvamme et al.,[154] we assert that, even when a drug is judged to be cost effective, it is still important to consider whether the actual health gain is significant; for example three months of survival rather than just one week.

Although the Dutch system currently does not yet fulfil Daniels and Sabin's revision requirement, it may do so if temporary decision making (including requirements on outcomes research) is extended to specific groups of outpatient drugs in 2013, as planned. Irrespective of this possible improvement, Dutch decision makers could consider emulating their Swedish counterparts in requesting additional data on specific uncertain items, rather than requiring general 'outcomes research'. Recently, Dutch decision makers (i.e., CVZ, ACP and the Minister of Health) experienced difficulties in enforcing the

consequences of first revisions for expensive inpatient drugs (i.e., omalizumab, infliximab, and ranibizumab). After four years of data collection, the revisions were judged to include insufficient data. If the reimbursement of drugs were to be made conditional on evidence development, the consequences of evaluating new evidence (e.g., delisting, or altering the reimbursement price or reimbursement restrictions) would be enforceable. The Swedish policy option of revising a decision increases the probability of a 'yes' decision. In this sense, that system can be regarded as being more lenient than the Dutch system, in which decisions on outpatient drugs are rarely revised.

Surprisingly, although the contrary is often held to be true,[169,170] pharmacoeconomic evidence does not seem to play an important role in The Netherlands. Dutch policymakers could reconsider the rationale behind the exemptions for pharmacoeconomic evidence (65%). Swedish dossiers show that, for example, in the case of drugs that have orphan status, manufacturers still seem to be able to provide information on value for money. Similarly, other European agencies (e.g., in Scotland, England and Wales) do not exempt pharmacoeconomic evidence on the sole grounds that a drug has orphan status or is indicated for HIV treatment.

Although both countries reported on consultation with stakeholders, it was not clear whether, and if so, how this influenced the decision making. Regarding the role of other appraisal criteria, in both countries, dossiers for three drugs (i.e., dasatinib, erlotinib, bortezomib) reported that the drug in question was indicated for a disease of (high) severity (i.e., cancer); all three received a 'yes' decision for reimbursement. Interestingly, while in three out of 11 cases (i.e., varenicline, inhaled insulin, sitagliptin) similar comparators were used, the therapeutic value judgement differed. All three drugs received a 'yes' decision in Sweden, whereas one (i.e., varenicline) received a 'no' decision in The Netherlands. Regarding varenicline, other arguments that may have played a role are 'life-style', 'own responsibility', and/or budgetary impact. In all these 'yes' decisions, reimbursement was restricted to a subgroup of patients and/or was made conditional on additional data collection. This may indicate that both countries were uncertain about the drug's actual value, or that, in The Netherlands, budgetary impact was a concern. It should be noted that the published dossiers did not reveal whether, and if so, how these or other arguments were decisive in the decision-making process.

Furthermore, two of the four different decisions (i.e., rimonabant and varenicline) concerned drugs for which life-style arguments may have played a role in the decision-making. If the Dutch agency's restrictiveness was due to such arguments, that may have reflected their 'necessity to insure' criterion (i.e., whether the illness or the care needed justifies solidarity). Other possible explanations are that the Swedish agency,

on the grounds of value-based pricing, considered the drugs to be cost effective, or that considerations of budgetary impact led to the 'no' decision in The Netherlands. Budgetary impact may also have played a role in the Dutch agency's decision to place a BMI restriction on a third drug, exenatide. It is important to realise that, for rimonabant and varenicline, the decision-making context differed: comparator drugs with the same indications were reimbursed in Sweden and not in The Netherlands. The reimbursement decision may have differed because agencies chose to be consistent with previous decisions.

Finally, the Swedish value-based pricing system may provide pharmaceutical companies with a better incentive for appropriate price setting than the Dutch system, in which prices are determined independently of reimbursement decisions.

## Conclusion

Both The Netherlands and Sweden publish information relating to their decisions on drug reimbursement. Their assessment processes are reasonably transparent: both report –albeit in varying detail– information on underlying studies, therapeutic value and cost-effectiveness. However, both could improve transparency of the appraisal. Therapeutic value appears to be the most decisive criterion; the relative importance of full economic evaluations is more modest than would generally be expected, especially in The Netherlands. The actual role of other (societal) criteria is more or less implicit and thus not sufficiently transparent. Therefore, both countries could improve legitimacy of their decision-making process by more explicitly showing the actual role of each different (societal) criterion in the appraisal process.



## Acknowledgements

This study was partly performed in the context of the Escher project (T6-202), a project of Top Institute Pharma, Leiden, The Netherlands.

# 7

# Health economic evaluations in reimbursement decision making in The Netherlands: Time to take it seriously?

With Marc Koopmanschap and Adri Steenhoek

German Journal for Evidence and Quality in Health Care [In Press]

## Abstract

Health technology assessment already informed Dutch policymaking in the early 1980s. Evidence of health economic evaluations is, however, only systematically used in drug reimbursement decision making. Outpatient drugs with an added therapeutic value and expensive specialist drugs require evidence from an economic evaluation. Due to many exemptions, however, the availability of evidence of health economic evaluations remains rather low.

The Dutch reimbursement agency suggested a cost-effectiveness threshold range depending on the severity of the disease (i.e., €10,000–80,000 per Quality Adjusted Life Year), but it was never confirmed nor endorsed by the Ministry of Health. It is highly questionable whether health economic evaluations currently play a role in actual Dutch reimbursement decision making.

Although the requirements exist in policy procedures, recent cases show that Dutch policymakers experience great difficulties in putting restrictions on reimbursement based on evidence from health economic evaluations. The near future will show whether the need increases to base decisions on societal value for money, and whether Dutch policymakers show the courage to take health economic evaluations seriously.

## Introduction

Health Technology Assessment (HTA) became a policy tool in The Netherlands in the early 1980s.[171] Although the contrary is often held to be true,[172] we believe, however, that the role and impact of evidence from a health economic evaluation (HEE) still remains limited in actual health care decision making. To arrive to this conclusion, this article firstly describes the Dutch reimbursement system and the applied priority setting principles. Since HEEs are mainly used in drug reimbursement decision making, we focus on the drug reimbursement process, the applied criteria and the requirements regarding HEEs. After that, we assess HEE evidence regarding its availability and quality. Finally, we evaluate the impact of HEE in actual decision-making practice.

## Health care reimbursement and priority setting in The Netherlands

The Dutch health care system is based on a social health insurance system, funded by public and private sources. Since 2006, all residents are obliged to enrol to universal basic health insurance which is provided by competing health insurers. Adults pay a flat rate premium, a subsidy scheme relieves financial burden for lower incomes. Supplementary health insurance is privately offered on a voluntary basis. The Health Insurance Act (*Zorgverzekeringswet, Art. 63–66*), Health Insurance Ordinance (*Besluit Zorgverzekering, Art. 2.8*) and the Health Insurance Decree (*Regeling Zorgverzekering*) establish the legal basis of the reimbursement system.

Since the early 1980s, Health Technology Assessment (HTA) became a policy tool in The Netherlands, mainly due to the increasing numbers of new expensive technologies.[171] HTA-studies already informed policy making in the 1980s, for example on liver and heart transplantation, and screening for breast cancer.[173,174] In 1991, the Dutch Committee on Choices in Health Care explicated HTA as a priority setting tool by suggesting a funnel, the funnel of Dunning, to determine the basic benefit package. The Dunning funnel has four decision criteria: the technology should be necessary, effective, cost effective, and affordable (i.e., individuals cannot bear the responsibility for the actual costs).[175] Ever since, the Dutch government continued to express the potential importance of the role of HTA in reimbursement decision making.

The Health Care Insurance Board (*College voor Zorgverzekeringen*; CVZ; from April 2014 *Zorginstituut Nederland*), an independent government funded agency, has the responsibility to advise the minister of Health care, Welfare and Sports regarding the entitlements of the basic benefit package. It is CVZ's mission to *"safeguard and develop the*

*public preconditions for the health care insurance system, so that Dutch citizens can obtain their right to health care."*[176] To ensure this mission, CVZ's guiding principles regarding the entitlements of the basic benefit package are: quality, accessibility and affordability.[177] As inspired by the Dunning funnel, CVZ uses four priority setting principles, namely necessity, effectiveness, cost-effectiveness, and feasibility.[177]

Consequently, cost-effectiveness is one of the four formal priority setting principles for the basic benefit package. This criterion is, however, not systematically used across the entire benefit package, but mainly used for decision making concerning drugs. Regarding non-pharmaceuticals, only in a few cases considerations of cost-effectiveness may have been taken into account (e.g., smoking cessation programs, and severe dyslexia in children).[178] Therefore, the next part of the article will only focus on decision making for drugs.

## Drug reimbursement and health economic evidence requirements

The aim of the Dutch drug reimbursement system is to guarantee safe and efficient pharmaceutical care according to individual patient's need in concurrence with scientific standards.[179] The CVZ has the legal responsibility to advise the minister whether or not a drug should be included in the basic benefit package and thus funded from public sources.

Briefly, the reimbursement procedure is as follows (see Figure 7.1). The applicant submits a reimbursement request. Based on the application file, CVZ's secretariat prepares an assessment report. This report is then, at least once, evaluated by the Scientific Pharmaceutical Advisory Commission (*Wetenschappelijke Adviesraad Commissie Geneesmiddelen*, WAR-CG; formerly *Commissie Farmaceutische Hulp*, CFH). Members of the WAR-CG have expertise in various medical disciplines, pharmacology, health sciences, and (health) economics; two representatives of the ministry attend WAR-CG meetings as observers. During this procedure, CVZ sends preliminary reports to the manufacturer and relevant stakeholders such as physicians, physician associations, patient associations, health care insurers, and hospital associations. Stakeholders have maximally two weeks to put forward their comments. Reimbursement files may, in case of perceived societal aspects, also be evaluated by the Appraisal Committee (*Adviescommissie pakket*, ACP), but this seldom occurred since the committee's establishment in 2008. Besides three CVZ directors, the ACP consists of experts in social security, health care, health insurance, medical ethics, HTA, and one patient representative. Finally, CVZ's Board of Directors forwards the reimbursement advice to the minister of health, who makes the final decision.

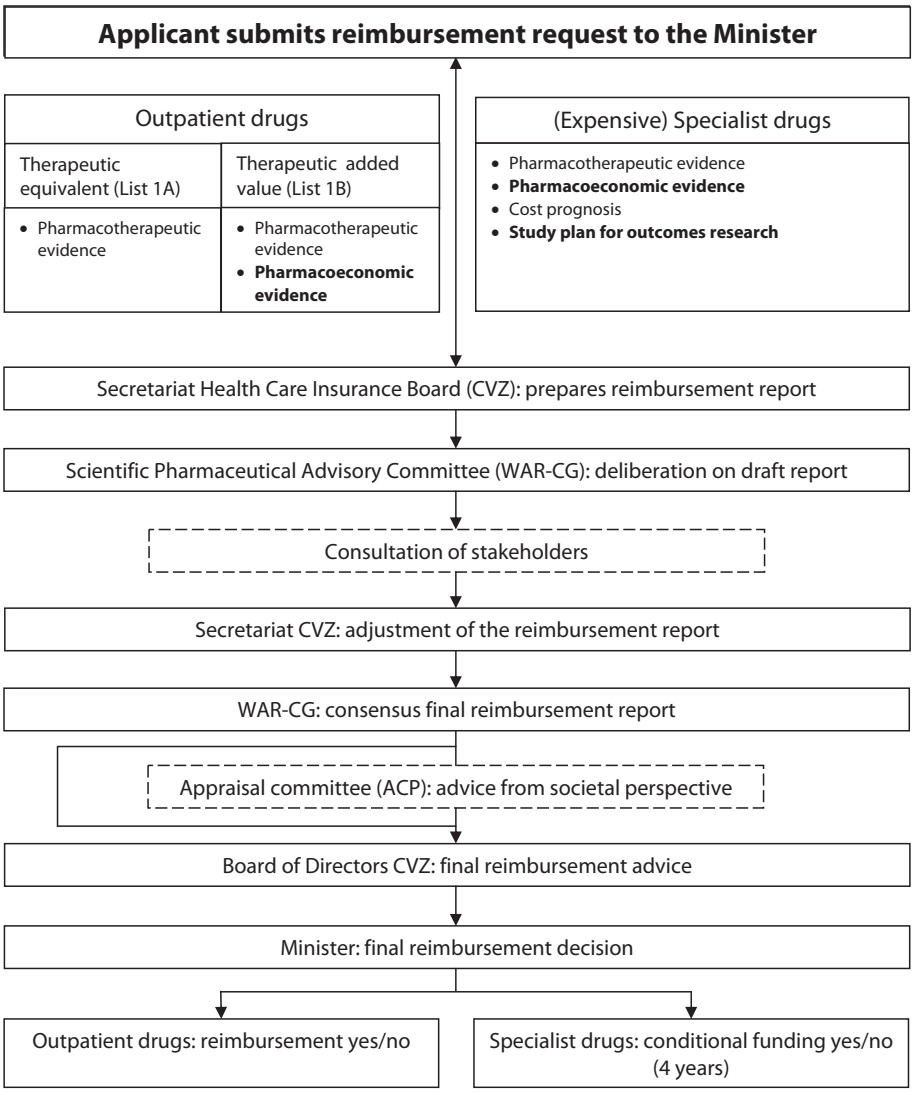


**Figure 7.1** Flowchart of the drug reimbursement procedure in The Netherlands

The CVZ applies, without a formal hierarchy, the following assessment and appraisal criteria: medical need, (added) therapeutic value (including effectiveness and adverse effects), cost-effectiveness, feasibility, necessity to insure, budget impact, disease severity, rarity of the disease, own responsibility, accessibility, social affordability, and public health benefit.[120] The reimbursement pathway and the requirements differ for outpatient and (expensive) specialist drugs (before 2013 called expensive inpatient drugs). Outpa-

tient drugs either belong to List 1A (i.e., groups of therapeutically interchangeable drugs for which the reimbursed price is based on the average price of the group) or List 1B (i.e., drugs with added therapeutic value for which the reimbursed price is not subject to a referenced limit). Besides evidence on pharmacotherapeutic value, only applications for List 1B require pharmacoeconomic evidence. Specialist drugs (i.e., provided through hospitals and an estimated budgetary impact ≥ €2.5 million[180]) fall under a conditional funding policy if uncertainty exists about cost-effectiveness and appropriate use in everyday clinical practice; initial applications require, besides pharmacotherapeutic and pharmacoeconomic evidence, a cost prognosis, and a study plan for outcomes research. After four years of temporary funding, a reassessment is performed based on additional collected data. Since 2012, several groups of expensive outpatient drugs (e.g., TNF alpha blockers) have been transferred to the specialist drug list. The ministry expects that (groups of) hospitals can successfully negotiate for lower prices of these drugs, and that these drugs will be used appropriately if prescribed by medical specialists.

Information on cost-effectiveness through a HEE is thus only required for applications for List 1B and specialist drugs, and, for the latter, reassessments after four years. Since 2002, CVZ encouraged submitting HEE evidence for List 1B applications; in January 2005, it became a mandatory requirement. Since the introduction of the coverage with evidence development policy in 2006, outcomes research including HEE evidence requirements became mandatory for specialist drugs.

Regarding outpatient drugs, the manufacturer is responsible for the application and thus for submitting HEE evidence. Although the manufacturer is not the applicant party for specialist drugs, the manufacturer is in most cases involved in the execution of HEEs. CVZ does not conduct its own HEE as in some other countries (e.g., the United Kingdom); they base their reimbursement advice on the submitted evidence in combination with other available information. The manufacturer is free to conduct their own HEE, or can outsource this to, for example, a HTA organisation or consultancy firm.

Table 7.1 provides an overview of the Dutch method guidelines for pharmacoeconomic research.[181] For more detailed information on the Dutch system, we refer to Franken et al.[103] and Le Polain et al.[15]

## The availability of health economic evidence

The Dutch system seems straightforward; namely, an internal price referencing system determines the reimbursed price in case a drug is therapeutically comparable to other

**Table 7.1** Dutch guidelines for pharmacoeconomic research [a]

| Guideline item | Explanation of the requirements for pharmacoeconomic research |
|---|---|
| Study perspective | The economic evaluation should be conducted using a societal perspective. |
| Comparator | The drug should be compared to the standard treatment for the indication in question, or, if not possible, compared to usual care. The comparator treatment can be a pharmaceutical as well as a non-pharmaceutical. |
| Analysis technique | The following techniques should be used in case of added therapeutic value:<br>• Improvement of quality of life: a cost-utility analysis<br>• No improvement of quality of life: a cost-effectiveness analysis<br>And in case of similar therapeutic value:<br>• Cost-minimisation analysis |
| Time horizon | A time horizon is sufficient when it enables a valid and reliable judgement of the effects and costs (modelling techniques can be used). |
| Costing methods | All (direct, indirect, medical, and non-medical) costs should be included; volume and unit costs need to be transparent. |
| Quality of life methods | Quality Adjusted Life Years should be transparently reported using survival data and utilities (by means of patient reported outcomes EQ-5D, HUI, or by direct utility valuation techniques such as TTO, SG, VAS). |
| Modelling techniques | The pharmacoeconomic model should be transparently reported (including model structure, input parameters, assumptions etc.), and preferably based on peer reviewed public publications. |
| Incremental methods | Incremental differences in effects and costs should be separately reported in detail. |
| Discounting | Costs should be discounted at a 4% rate, and effects at a 1.5% rate. |
| Sensitivity analysis | Sensitivity analysis should be conducted using deterministic, probabilistic, and scenario analysis. |
| Expert panels | In case of lack of input parameters, expert panels can be used. A detailed description should be provided regarding the composition of the expert panel and how consensus was reached. |

*[a] Based on Dutch guidelines for pharmacoeconomic research[181]*
*EQ-5D= EuroQol 5 Dimensions; HUI= Health Utility Index; TTO= Time Trade-Off; SG= Standard Gamble; VAS= Visual Analogue Scale*

drugs (List 1A), and HEE evidence is considered in case a drug has added therapeutic value (List 1B). In practice, however, manufacturers are often exempted from submitting pharmacoeconomic evidence. Formal regulations exempt drugs with an orphan status,[182] drugs with an estimated budgetary impact lower than €500,000 per year,[182] and HIV drugs.[19] As a consequence, between January 2005 and July 2011, 65% of List 1B applications were exempted from conducting a HEE;[147] HEE evidence was only available for 35% of List 1B applications. The available HEE evidence consisted of cost-minimisation analyses (n = 5), cost-effectiveness analyses (i.e., costs per life-year gained; n = 19), and cost-utility analyses (i.e., costs per QALY; n = 20).[147] Moreover, 55% of the published dossiers included a description of a cost-effectiveness plane and/or an acceptability curve.[147] Regarding specialist drugs, so far, many initial applications do not include HEE

evidence. Consequently, HEE evidence is often only available for the reassessment after four years of conditional funding.

Furthermore, even if available, the methodological quality of the performed HEEs is often insufficient. Hoomans et al.[169] reported that, between January 2005 and October 2008, only eight out of twenty-one HEEs were consistent with CVZ's guidelines. Franken et al.,[147] however, reported that most of CVZ's unfavourable judgements on the robustness of the HEE evidence for List 1B applications occurred in the earlier years of the HEE requirement (i.e., 2005 to 2007).

## The impact of health economic evidence in actual decision making

Above all, it should be noted that CVZ only advises the minister regarding the robustness of the pharmacoeconomic evidence, thus CVZ does not advise on the actual cost-effectiveness estimate. In The Netherlands, there is no formal threshold value for the incremental cost-effectiveness ratio. Although CVZ suggested a threshold range depending on the severity of the disease (i.e., €10,000–80,000 per QALY); CVZ also stated that this range is only indicative and not predictive because their advice is always based on a balance of different considerations.[120] One could consider that the ACP (i.e., CVZ's committee that evaluates the entitlements of the benefit package from a broader societal perspective) is responsible for advising regarding cost-effectiveness. However, members of the ACP have indicated that the actual implementation of the cost-effectiveness criterion requires further exploration.[178] Importantly, the Dutch minister has never confirmed nor endorsed a cost-effectiveness threshold (range). An important step forward, however, may be that, in 2012, the two governing political parties formally stated their intention to give the cost-effectiveness criterion a statutory basis.[183] So far, no changes have been made in official documents.

Regarding outpatient drugs with HEE evidence, Franken et al.[147] found that, for all List 1B applications that were rejected for reimbursement between January 2005 and July 2011, the HEE evidence had been judged insufficiently robust. Of these twelve drugs, five were considered to have an added therapeutic value. Although unknown, HEE evidence can have played a role in these decisions to reject reimbursement. Interestingly, however, the HEE evidence of 21% of drugs that obtained reimbursement was also judged insufficiently robust.[147] Of these four drugs, two were considered to be of similar therapeutic value and two of added therapeutic value. It is highly questionable if HEE evidence played a role in these decisions to grant reimbursement. Most decisions

that were granted reimbursement concerned drugs with an added therapeutic value judgement (15 out of 19).[147]

The picture of the impact of HEE evidence in actual decision making becomes more complete when scrutinising decisions regarding specialist drugs. Since 2006, outcomes research including HEE evidence became mandatory in the coverage with evidence development policy. Initially, reassessments were scheduled after three years of temporary funding. Within a short time period, this was revised to four years due to expected feasibility issues. Consequently, one could expect that from 2010 onwards reassessments would take place on a regular basis.

However, only a few reassessments have been finalised; most reassessments, even from initial assessments in the years 2006, 2007 and 2008, are still queued up at CVZ. Remarkable, until a reassessment is finalised, funding from public sources continues for these expensive drugs. It seems that Dutch policymakers experience serious difficulties in enforcing the consequences of reassessments. Even after a time period for data collection in everyday practice, decision makers seemed to be embarrassed by the lack of sufficient evidence for the first reassessments. This, however, did not result in delisting of drugs.

In 2012, CVZ concluded that omalizumab for severe asthma had an added therapeutic value compared to other available treatment, but it remained highly uncertain at what costs (i.e., cost-effectiveness uncertainty range: €36,000 to 87,000 per QALY gained).[184] The minister followed CVZ's advice and arranged a pay-for-performance agreement.[184] It is to be evaluated in the near future whether this agreement contributes to cost effective usage of this drug. Experiences from other countries, however, show mixed results.[185,186] For expensive specialist drugs, CVZ currently often advises to set up public-private financed patient registries to lower uncertainty on the (cost) effectiveness in everyday practice. The minister even made reimbursement conditional on the set up of such a registry for an expensive treatment for melanoma.[187]

Furthermore in 2012, reassessments of three specialist orphan drugs for Pompe and Fabry disease resulted in a turbulent episode. Preliminary draft reports were leaked to the press. CVZ considered a negative advice for these drugs because they were considered too expensive relatively to their effectiveness (e.g., Myozyme for classic Pompe disease was estimated to cost 300,000 to 900,000 € per QALY[188]). The opinions of the ACP members were divided in a very well attended public ACP meeting. One absent ACP member put forward his point of view in writing by questioning the sustainability of reimbursing such expensive drugs in the long term and questioning the fairness of disproportionately prioritising orphan diseases.[189] After a considerable amount of public

debate and political pressure, CVZ modified its advice, recommending continuation of funding because of the high severity of the disease in combination with high costs per patient but a relatively low budgetary impact.[188] Remarkably, CVZ did not have any new scientific HTA evidence. It appears, therefore, that CVZ modified its advice only based on an undecided ACP advice and public and political influence. The minister accepted CVZ's advice and made, in 2013, a (confidential) price agreement with the manufacturer.[190] It is, however, debatable whether HEE evidence influenced the price agreement. Due to the extreme high costs per QALY, it must have been impossible agreeing on a price which ensures societal value for money. Consequently, the Dutch society continues paying an extreme high price for an uncertain and probably limited health gain. Because more expensive (ultra) orphan drugs are to be expected in the near future, it is debatable if such decisions are sustainable in the long term.

Consequently, the only conclusion that we can draw is that, so far, HEE evidence does not play a major role in actual Dutch decision making. Although HEE requirements are implemented in reimbursement policy procedures and the technical quality of HEE has increased, it seems that policymakers experience great difficulties putting restrictions on reimbursement based on value for money considerations. It is of course easier for policymakers to make 'happy' decisions that satisfy all stakeholders and do not attract any media attention.

Because a formal cost-effectiveness threshold (range) is absent, it cannot be expected that the CVZ, only having an advisory role, takes the lead in determining the Dutch threshold and advices the minister accordingly, for which they will get blamed for by society and subsequently overruled by the minister. A threshold range for society's willingness to pay for a QALY gained, potentially with orphan drugs at the upper range, could also act as a gatekeeper for conditional funding. If an expensive drug costs, for example, more than 100,000 euro per patient per year, it will, even in case of optimistically estimated health gains, inevitably fail to meet the maximum of any realistic threshold range. Policymakers could make better use of this knowledge at the initial decision and thus deny conditional funding for a too high priced drug in the first place. Moreover, being more stringent at the initial decision may limit the impression that access to that drug is an acquired right. It should be noted that CVZ does not decide on pricing, and has, therefore, no capacity for negotiations on the price. An alternative approach such as value based pricing is therefore not possible for CVZ; only the minister can arrange, since recently, a price agreement.

Besides a threshold range, other criteria may complement sustainable decision making. Importantly for example, the Dutch Commission for the Assessment of Oncological Re-

sources (Commission BOM) suggested a minimal gain of 2 months in life expectancy.[191] This links closely to pleas for only reimbursing interventions that produce at least a non-negligible health gain.[154]

Reimbursement decision making involves balancing different goals, namely access to high quality products in a sustainable manner. The perceived need to make 'unhappy' decisions, especially related to identifiable victims,[118] may not be very high in The Netherlands because pharmaceutical expenditures as percentage of total health care expenditure remained relatively stable over the previous years, mainly as result of other policies (e.g., tendering policy by health insurers). However, it is questionable if this is sustainable in the long-term. There is no incentive to appropriately price drugs according to their actual value; will there be any limit for prices and cost-effectiveness in the future? Moreover, if evidence from HEEs is hardly considered in decision making, the question may arise why HEE evidence requirements exist in the first place. Until now, it seems that access to (potentially) effective drugs is most important in decision making, no matter at what costs. Besides actual using HEE evidence in decision making, explicating a minimal health gain and a cost-effectiveness threshold (range) may be two of the most difficult political aspects. It is debatable, however, what the value is of a statutory basis of the cost-effectiveness criterion if not applied in actual decision making. The near future will show whether the need increases to base actual decision making on societal value for money, and whether Dutch policymakers show the courage to take HEE seriously.

# C

# Handling uncertainty in drug reimbursement decision making

# 8

# Access to expensive cancer drugs in Dutch daily practice: Should we be concerned?

With Hedwig Blommestein, Silvia Verelst, Michel van Agthoven, Peter Huijgens, Carin Uyl-de Groot

The Netherlands Journal of Medicine 2014; 72(4):235-241

## Abstract

The aim of this study was to investigate whether equal access to bortezomib has been achieved under Dutch policy regulations that guarantee equal access to expensive inpatient drugs.

We investigated accessibility to bortezomib treatment at national and regional levels by (i) conducting interviews with stakeholders in the Dutch health care system to explore prescription barriers and (ii) tabulating sales data from 2004–2009 and trial participation rates.

Interviews revealed awareness of the high treatment costs although prescription barriers were not encountered. National use of bortezomib increased slowly (treating 2% of patients in 2004 to 17% in 2009), indicating a long adjustment period. Furthermore, use remains below the rate estimated by the professional association of haematologists (27%). Regional differences were found for both daily practice use (e.g., ranging from 13–27% in 2009) and clinical trial participation (e.g., ranging from 1–12% in 2006).

Our results were somewhat conflicting: interviews did not reveal any prescription barriers, but quantitative methods showed regional differences, signs of underutilisation, and access inequality. Investigating use and accessibility, based on data triangulation, provides valuable feedback which can enhance evidence-based decision making for both physicians and policymakers. This could improve appropriate and efficient use and ensure equal access to expensive drugs.

## Introduction

Increasing health care expenditures may result in limited and unequal access, particularly with regard to new and innovative cancer drugs with high acquisition costs. Policymakers have to make reimbursement decisions considering both rapid and equal accessibility to promising drugs as well as the scarcity of resources. Usually, guaranteeing rapid access means making decisions while available evidence on clinical and cost-effectiveness is limited.[26] One way of dealing with the need for rapid access and limited evidence is the 'coverage with evidence development' policy; reimbursement under the condition that additional research will be conducted.[26]

Such policies have been implemented in several countries for surgical procedures, medical devices and pharmaceuticals.[31] Over the last decade, a coverage with evidence development policy was also initiated in The Netherlands, partly triggered by signs of underutilisation and 'zip code prescribing' of trastuzumab.[192] Early access to expensive inpatient drugs is linked with the obligation to gather data on appropriate drug use and cost-effectiveness in daily practice.[193] Drugs meeting the criteria of added therapeutic value and expected budget impact of at least 2.5 million were temporarily included in the policy of 2006–2012. Four years after inclusion, a reassessment will determine whether or not additional financing should continue to exist. At the time we conducted our study, hospitals received 80% of its acquisition costs if a drug was included.

Currently more than 30, mostly cancer, drugs are included in this policy. One of these drugs is bortezomib, used for treating multiple myeloma (MM). MM is the second most common haematological cancer. The five-year prevalence in Western Europe is 31,056 while the annual age-standardised incidence rate is 3.2 per 100,000 (IARC GLOBOCAN 2008). Bortezomib obtained EMA approval in 2004 by demonstrating superior efficacy compared with chemotherapy for the treatment of advanced MM;[194-196] it was included on the Dutch expensive drug list in 2006. Advances in MM treatment in the past decade significantly increased overall survival (44.8 vs. 29.9 months[197]), which was largely due to the introduction of autologous stem cell transplantation and new therapeutic agents including thalidomide, lenalidomide, and bortezomib.[197,198] While thalidomide is relatively inexpensive, bortezomib and lenalidomide are expensive drugs. Both are incorporated in professional guidelines.[199] However, the orphan status granted to lenalidomide results in 100% reimbursement for lenalidomide compared with an 80% of reimbursement for bortezomib during our study period. Consequently, accessibility might be an issue, especially for bortezomib.

Previous research studied accessibility and use of expensive drugs in The Netherlands;[200,201] however, it remains unclear whether the Dutch policy actually guarantees equal access to expensive inpatient drugs. We investigated whether equal access to bortezomib has been achieved in The Netherlands. We analysed bortezomib use patterns by means of aggregate sales data and conducted interviews to shed light on perceived or real prescription barriers.

## Methods

We took a two-pronged approach. First, seven in-depth interviews were conducted to qualitatively investigate the existence of accessibility issues and prescription barriers. Interviewees were representatives of stakeholders in the Dutch health care system: (i) a representative of the Dutch Health Care Authority (NZa), (ii) a representative of the Health Care Inspectorate (IGZ), (iii) a hospital director of finance, (iv) four haematologists from hospitals varying in size and country location (the North-West, East, South-West, and South). Respondents were selected based on their involvement and knowledge of expensive inpatient drug regulations (NZa and IGZ) or geographical location and type of hospital (haematologists and director of finance). All semi-structured interviews were recorded and analysed according to the steps of Creswell,[202] including transcription, coding, interpretation, and description.

Second, we quantitatively investigated the use of bortezomib in daily practice. Because data on bortezomib use at the individual patient level are not available, we combined Dutch sales data (excluding use in clinical trials) from 2004–2009 from the manufacturer, Janssen Pharmaceutical Companies of Johnson & Johnson with incidence and prevalence data from The Netherlands Cancer Registry.[203] Figure 8.1 provides the flowchart of data used, intermediate and final outcomes and the underlying assumptions.

To estimate the number of treated patients ([A] in Figure 8.1), the number of vials sold was divided by the average number of vials used per patient. The average number of vials per patient (18.24) was based on a Dutch observational study of 72 bortezomib patients treated in daily practice from 2004–2008.[204]

To investigate bortezomib use across regions, we used the regional division of the nationwide Netherlands Cancer Registry distinguishing eight Comprehensive Cancer Centres.[203] Since these regions differ in size, prescription rates were expressed relative to the number of patients per region. We assumed that equal accessibility to bortezomib would be achieved if the proportion of vials used per region was similar to their propor-

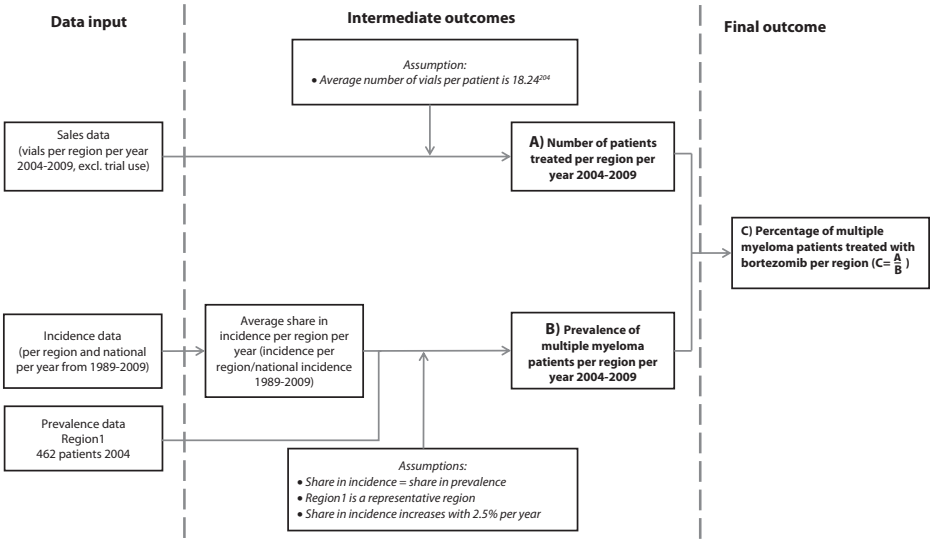


**Figure 8.1** Flowchart of data input, intermediate and final outcomes

tion of national incidence or prevalence. Regional shares in incidence were calculated over the years 1989–2009. For example, the share in incidence in 2009 for Comprehensive Cancer Centre Amsterdam (IKA) was 18.8%. We calculated this percentage by dividing the incidence of IKA (201) by the national incidence (1069).

Because prevalence numbers were only available for IKA (462 patients in 2004) for one year, we estimated other regional prevalence (B) from their relative shares in incidence. Hereby we assumed (i) IKA to be representative for the other regions and (ii) the share in incidence per region is equal to the share in prevalence (e.g., if IKA has 19% of the incidence it will also have 19% of the prevalence), and (iii) an annually increasing prevalence of 2.5% (average annual increase over the years 1989–2009[203]) per year because of rises in incidence.[198] Detailed additional information about incidence and prevalence estimates per year is available from the authors upon request.

To obtain a regionally comparable percentage of treated patients (C), we divided the estimated number of treated patients (A) by the estimated prevalence (B). To put regional percentages in perspective, we compared our computed use with the expected percentage of MM patients eligible for bortezomib treatment as estimated by the Dutch professional association of haematologists (the Dutch-Belgian Cooperative Trial Group for Haematology and Oncology [HOVON]). HOVON estimated that about 1600 patients would be eligible for MM treatment per year. Of these patients, one-third would not qualify for treatment with either bortezomib or lenalidomide due to age, patient's

condition or preferences. As result, 1070 patients are eligible for advanced therapy each year.[205] Since patients treated with bortezomib might also be eligible for treatment with lenalidomide and vice versa, HOVON assumed that the number of patients treated with each drug would be similar (50%). To compare the HOVON estimation with the proportion of patients treated with bortezomib per region, we divided the 535 eligible patients (i.e., 1070 divided by 2) by HOVON's estimated prevalence (i.e., 2000 patients), resulting in an estimation of 27% patients.

Furthermore, since bortezomib was a novel treatment, clinical trials were conducted during our years of investigation. Because MM patients are often included in clinical trials, relatively high or low trial participation could distort our computed daily practice use and identified regional differences. Therefore, we selected the two largest clinical studies including bortezomib during our investigated time period and studied trial participation at the regional level. Calculation methods were similar: we divided the number of patients included in trials by regional prevalence to obtain regional trial participation rates for the years 2005–2009. We then combined trial participation with regional daily practice use to compare similarities and differences across regions.

## Results

### Interview results

Interviewees of the NZA and IGZ did not reveal any accessibility issues for expensive drugs. The IGZ representative, however, admitted that the body had no active role in investigating such issues.

Hospitals regulate financial management in various ways. As a result, it may differ per hospital who is responsible for the budget and who is making the financial decisions. According to the interviewed physicians, their financial department divided the total hospital budget by department, whereas physicians organised the division and implementation of the budget within departments. These assumptions were verified and confirmed by the hospital financial manager. Based on these results, we concluded that in the studied hospitals financial management, of both treatment decisions and organisation of care, was the physicians' responsibility.

Generally, all physicians agreed that access to bortezomib is guaranteed in The Netherlands for patients in need. The existence of strict quantitative restrictions was explicitly denied. Physicians adhered to professional guidelines as far as treatment is concerned, which were frequently mentioned as important. Consultation with colleagues and

patient characteristics also seemed to be important factors in the decision (how) to treat. Apart from some variation immediately after the introduction of bortezomib, respondents believed that all eligible patients had equal access.

The Dutch policy of 2006–2012 aimed to facilitate prescription and guarantee access while maintaining incentive for efficiency. According to haematologists, the effects of this policy were two-sided. An additional budget of 80% facilitated prescription but the remaining 20%, financed from the general hospital budget, could hinder prescription. The policy was therefore perceived as ambiguous: while the government relieved the high financial burden, the remainder still had to be financed from the general hospital budget. The situation stimulated local initiatives to manage access to expensive drugs, resulting in a local expensive drug committee to judge appropriate use and structures for consultations with more experienced physicians. Although expensive drugs were perceived as a high financial burden, according to the respondents, budget played no role in treatment choices.

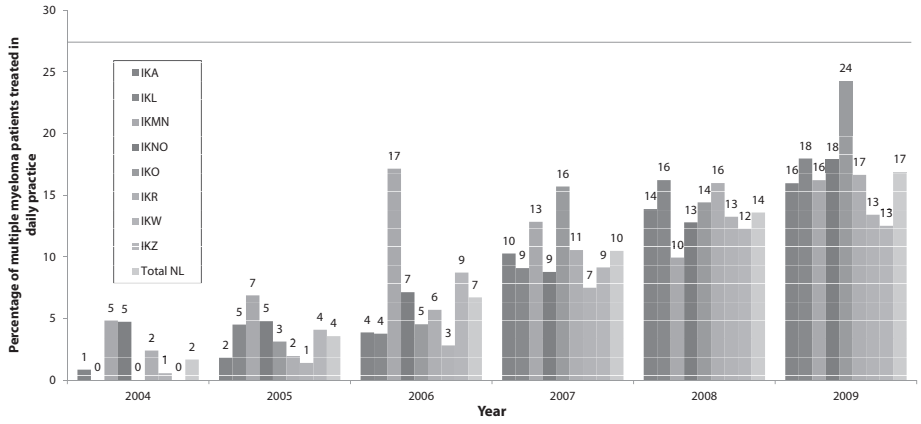


**Figure 8.2** Percentage of multiple myeloma patients treated in daily practice with bortezomib per region from 2004-2009
*IKA: Comprehensive Cancer Centre Amsterdam*
*IKL: Comprehensive Cancer Centre Limburg*
*IKMN: Comprehensive Cancer Centre Netherlands Central*
*IKNO: Comprehensive Cancer Centre North East*
*IKO: Comprehensive Cancer Centre East*
*IKR: Comprehensive Cancer Centre Rotterdam*
*IKW: Comprehensive Cancer Centre West*
*IKZ: Comprehensive Cancer Centre South*

## Data results

*Daily practice use*

Figure 8.2 shows the percentage of patients treated with bortezomib from 2004–2009 irrespective of treatment line. As mentioned in the method section, HOVON estimated 27% of MM patients are eligible for bortezomib treatment in daily practice. This is presented as a horizontal line in Figure 8.2. The figure reveals relatively low use in 2004–2005 for all regions, which was expected since bortezomib was then an innovative treatment and not included on the expensive drug list until 2006. Three regions did not use bortezomib in 2004; all regions used it in 2005. Differences across regions exist in all years with no stable pattern; sometimes regions switched from a high prescription rank in 2005 and 2006 to a low one in 2008. In 2008, two years after inclusion on the expensive drug list, differences between the regions decreased. In 2009, Comprehensive Cancer Centre East (IKO) was the highest prescribing region and Comprehensive Cancer Centre South (IKZ) the lowest, revealing that in one region 24% of patients received bortezomib while in another only 13% received bortezomib. In all the regions the prescription rate was below the 27% of eligible patients as estimated by HOVON.

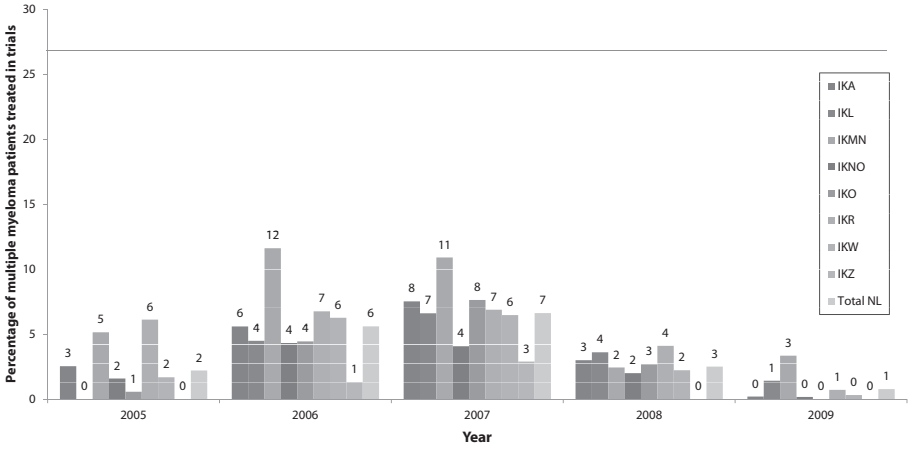


**Figure 8.3** Percentage of multiple myeloma patients treated in clinical trials (HOVON 65 and HOVON 86) per region from 2005-2009
*IKA: Comprehensive Cancer Centre Amsterdam*
*IKL: Comprehensive Cancer Centre Limburg*
*IKMN: Comprehensive Cancer Centre Netherlands Central*
*IKNO: Comprehensive Cancer Centre North East*
*IKO: Comprehensive Cancer Centre East*
*IKR: Comprehensive Cancer Centre Rotterdam*
*IKW: Comprehensive Cancer Centre West*
*IKZ: Comprehensive Cancer Centre South*

*Use in trials*

Figure 8.3 shows the participation in the HOVON 65[206] (phase I/II study) and HOVON 86[207] study (Phase III randomised controlled trial) per region in the 2005–2009 period. We observed different trial participation rates and, as Figure 8.3 illustrates, trial participation increased from 2005–2007, and decreased in 2008 to almost no participation in 2009. A comparison of Figure 8.2 and Figure 8.3 reveals that the percentage of patients treated in trials is lower than daily practice use of bortezomib.

Finally, Figure 8.4 presents the regional percentages of treated patients aggregated over the years 2005–2009. Comprehensive Cancer Centre Netherlands Central (IKMN) had the highest daily practice use and trial use (19% were either treated with bortezomib or included in one of the larger trials); IKZ had the lowest (10%). Figure 8.4 also shows that although differences remain, the fluctuation reduced over time. In general, regions with above average daily practice use also had above average trial participation rates.

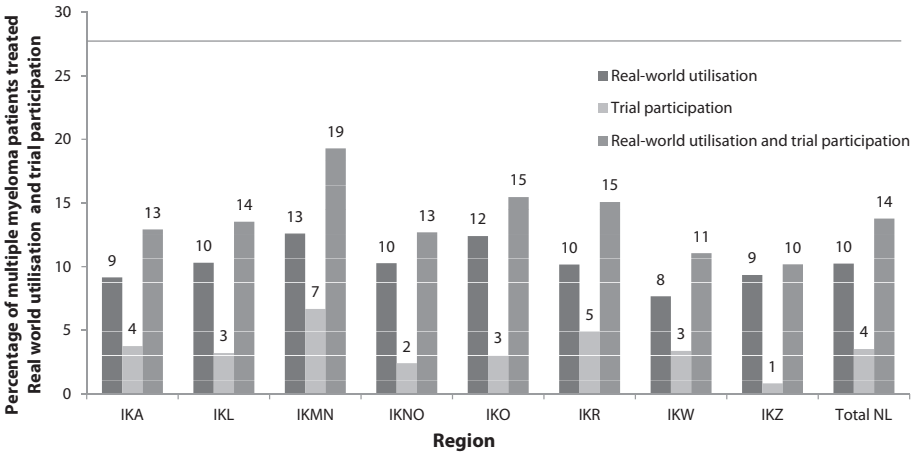


**Figure 8.4** Percentage of multiple myeloma patients treated in daily practice and clinical trials 2005-2009
*IKA: Comprehensive Cancer Centre Amsterdam*
*IKL: Comprehensive Cancer Centre Limburg*
*IKMN: Comprehensive Cancer Centre Netherlands Central*
*IKNO: Comprehensive Cancer Centre North East*
*IKO: Comprehensive Cancer Centre East*
*IKR: Comprehensive Cancer Centre Rotterdam*
*IKW: Comprehensive Cancer Centre West*
*IKZ: Comprehensive Cancer Centre South*

## Discussion

The aim of our study was to investigate whether bortezomib treatment conformed to policy regulations that were designed to guarantee equal access to expensive inpatient drugs in The Netherlands. Interviews revealed that physicians feel some financial pressure but do not experience prescription barriers and believe that access to expensive cancer drugs is guaranteed. In addition, at that time there were no signs of accessibility issues among IGZ and NZa. Our results, however, also showed that (i) after the introduction of bortezomib, it took one to two years before the drug was prescribed regularly in all regions; (ii) the percentage of patients treated is below the expected 27% of eligible patients; and (iii) there are unexplained regional differences.

In order to investigate accessibility issues and compare regional use levels we had to make several assumptions, especially to calculate the percentage of MM patients treated with bortezomib. While the regions defined by the Dutch cancer registry vary in size, population and available hospital facilities, we expect the baseline patient characteristics to be comparable across regions. Since accurate prevalence numbers were unavailable, we assumed prevalence could be obtained from the distribution of incidence after verifying that the regional distribution of incidence was stable over a long period with a maximum deviation of only 3%. Some uncertainty surrounding total prevalence, however, remains.

Although these assumptions influence the percentage of patients treated, we believe our conclusion of low prescription rates will not be effected. Levels of use would only be closer to HOVON's expected use of 27% if the prevalence of multiple myeloma was much lower (i.e., less than 1700 patients). Considering incidence is 1100 patients per year, prevalence of less than 1700 seems highly unlikely.

Nevertheless, the share in incidence per region was remarkably stable confirming a stable division between the regions over time. If prescription rates per region were similar, we expected the regions to be accountable for a similar share in bortezomib as their share in incidence. Therefore, regional variation was definitely established, although violations of our assumptions could enlarge or reduce the differences.

Observed regional variation, in both daily practice and trial use, indicates either differences in prescription behaviour or referral of patients to, for example, more experienced hospitals. Because we used sales data aggregated per hospital, we cannot distinguish between patients living in the region and patients referred to the region. Both causes –prescription behaviour and patient referral– limit accessibility. IKZ may have been

especially sensitive to regional border crossing because it is the only region without an academic hospital. In this region, use and trial participation is low while relatively high numbers are observed in its neighbouring region (i.e., IKMN). Bortezomib administration, however, does not require specialised skills or hospital facilities, implying that expertise may have been a valid reason for referral immediately after the introduction in 2004, but should be of minor importance in subsequent years.

We studied treatment patterns at an aggregated level, hence neglected other treatment options such as thalidomide and lenalidomide. Because thalidomide is relatively inexpensive in The Netherlands, accessibility should not be an issue. Lenalidomide was accepted for reimbursement at the end of 2007 in Dutch daily practice, creating a competitive alternative treatment option for the years 2008 and 2009 in our analyses. However, lenalidomide does not compensate the low levels of bortezomib prescription. In 2007, 75 patients were treated with lenalidomide and this number increased to 452 and 671 in 2008 and 2009, respectively.[205,208]

Regional differences and under-provision have been previously reported in The Netherlands. Large regional differences and under provision of trastuzumab in The Netherlands were, according to the Dutch Breast Cancer Association,[192] mainly due to cost. After the accessibility issues of trastuzumab, the Dutch policy for expensive drugs was revised in 2006. Although bortezomib has been on the market since 2004, it was not until it was admitted to the expensive drug list in 2006 that its use in daily practice doubled compared with the previous year. The increase might indicate that the implemented policy facilitated prescription. Other developments occurred simultaneously however, including changes in professional guidelines that recommended bortezomib in earlier treatment phases. The relatively low use in the first years may have been caused by a long adjustment period of physicians who needed to be familiarised with a new drug.[209,210] Bortezomib was, apart from the re-introduction of thalidomide, the first new innovative treatment option for multiple myeloma patients in four decades. It is important that physicians and policymakers are aware of such lags in the regular use of a new innovative and effective drug. Their implementation should receive more attention to accelerate diffusion by, for example, providing feedback about daily practice use. Groot et al.[200] showed that the use of bortezomib in 2005 was almost three times higher in Sweden and France compared with The Netherlands. Furthermore, Dutch use in 2007 was a little less than 35 mg per 100,000 inhabitants while the European average (Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Switzerland and the UK) was above 50 mg per 100,000 inhabitants.[201] Our results also showed that use was below HOVON's expected rate. Despite financial assistance, use and accessibility issues may thus still exist.

It remains subject to further research whether observed regional differences are due to physician prescription behaviour or referral to more experienced or wealthier hospitals. Differences seem to have decreased compared with previous outcomes of the trastuzumab study in 2005, which might be a result of the changes in the policy regulations. However, we should note that the trastuzumab study analysed patients with breast cancer, whose prevalence is much higher than multiple myeloma.

Wagelaar et al.[211] studied accessibility of two expensive drugs in The Netherlands, bortezomib and trastuzumab, mainly by investigating whether prescription was in accordance with guidelines at the individual patient level. Medical files were examined and interviews were conducted with physicians, members of hospital boards of directors, and patients. They concluded that guidelines were strictly followed and that recommendations by the professional association and patient characteristics determined treatment decisions. Although the budget of 80% was insufficient according to their respondents, accessibility was not an issue. Interestingly, while their results align with our interview results, they are in contrast with our quantitative findings and our research shows that differences in accessibility might not be revealed by using a qualitative research method only.

In 2012, changes in the regulations increased the earmarked budget to full coverage of the 'add-on' diagnoses-related group (i.e., 100% reimbursement of expensive drugs but hospitals and insurers negotiate on the price of the 'add-on'). Although hospital resources remain scarce, this might improve access and reduce remaining regional differences. It will be interesting to closely follow the consequences of this new policy.

We investigated equality in access to bortezomib in the context of Dutch policy regulations for expensive drugs. Use of bortezomib has increased over time although regional differences are still present. We obtained different conclusions using two methods. While interviews did not reveal absolute prescription barriers, regional differences and possibly underutilisation were observed by comparing sales data with incidence and prevalence data. It seems that appropriate drug use and thus also accessibility depends on various factors, regulatory and organisational characteristics of a health care system being two important ones. An evaluation of health policies should therefore be based on mixed methods and data triangulation. Such an evaluation provides insight and valuable feedback that can enhance evidence-based decision making for both health care providers and policymakers. This could improve appropriate drug use and ensure equal access to health care. In the end, efficient and equitable use of scarce resources increases society's benefits from a health care system.

# 9

# Practical feasibility of outcomes research in oncology: Lessons learned in assessing drug use and cost-effectiveness in The Netherlands

With Chantal van Gils, Jennifer Gaultney, Gepke Delwel, Wim Goettsch, Peter Huijgens, Adri Steenhoek, Cornelis Punt, Miriam Koopman, Ken Redekop, Carin Uyl-de Groot

European Journal of Cancer 2013; 49(1):8-16

## Abstract

The aim of this study was to investigate the practical feasibility to develop evidence on drug use and cost-effectiveness in oncology practice.

Feasibility was examined using three Dutch case studies. Each case study investigated the degree of appropriate drug use and its incremental cost-effectiveness. Detailed data were retrospectively collected from hospital records. In total, 391, 316 and 139 patients with stage III colon cancer, metastatic colorectal cancer and multiple myeloma were included in 19, 29, and 42 hospitals, respectively.

The methods used in the case studies were feasible to develop evidence on some aspects of drug use including types of treatments used, dosages, dose modifications, and health care costs. Aspects such as baseline patient characteristics, reasons to start or stop a treatment, and treatment effects were less feasible because of missing values. Despite difficulties to correct for confounding by indication, it was possible to estimate incremental cost-effectiveness by synthesising evidence in two of the three case studies.

It is possible to generate evidence about drug use and cost-effectiveness in oncology practice to facilitate informed decision making by both payers and physicians. This can improve quality of care and enhance the efficient allocation of resources. However, the optimal approach differs between drugs and their indications. Generating high-quality evidence requires active interdisciplinary collaboration. Patient registries can facilitate data collection but cannot resolve all issues. In most circumstances it is inevitable to use data-synthesis to obtain valid incremental cost-effectiveness estimates, but for some indications it will not be feasible to derive a valid and precise estimate.

## Introduction

The increasing number of expensive oncology drugs is making it extremely difficult to strike an optimal balance between ensuring timely access to 'promising' drugs and having sufficient evidence of their comparative benefits and risks. In the last decade, governments have therefore introduced policies linking reimbursement to a requirement for additional data collection.[18,24-26] It has been claimed that the resulting 'schemes' or requirements, such as 'patient access schemes',[27] 'managed entry agreements',[28] 'access/coverage with evidence development',[29,30] comparative effectiveness research,[212,213] or outcomes research,[32,33,214] can result in better evidence by addressing uncertainty arising from the gap between clinical trials and everyday practice.

The initiatives to promote and implement such requirements have initiated a stream of literature, touting its potential as well as highlighting the myriad of methodological challenges to its feasibility.[212,213,215-221] In particular, one of the major concerns raised has been the lack of a randomised controlled setting, which results in problems with internal validity.[219,222-224] Moreover, other challenges have been debated such as the role of data synthesis and modelling,[25,212,223,225] use of existing data sources,[212,213,216,220,226] applicability of outcome measures,[25,33] timeliness,[26,227] and generalisability.[228] However, these issues have mainly been discussed on the basis of theoretical expectations and expert opinion; studies of the practical feasibility are scarce.

Experience with conducting outcomes research has already been gained in The Netherlands as a result of policy regulations for expensive inpatient drugs implemented in 2006. If a drug is included in this policy, hospitals receive an additional ear-marked budget of 80% of its acquisition costs.[229] However, this early access is linked with the obligation to gather data on appropriate drug use and incremental cost-effectiveness.[193,229] In practice, this means that after four years of use, a reassessment will determine whether or not additional financing will continue to exist.

This paper describes our experiences in The Netherlands regarding the practical feasibility of different aspects of outcomes research in oncology. These experiences were based on three different outcomes research studies which examined the feasibility to gather evidence on appropriate drug use and estimate incremental cost-effectiveness of a particular drug.

## Methods

We conducted outcomes research of two expensive drugs for three indications in cancer: oxaliplatin as adjuvant treatment in stage III colon cancer, oxaliplatin as palliative treatment in metastatic colorectal cancer and bortezomib as palliative treatment in relapsed or refractory multiple myeloma. Each of the studies was used to investigate the feasibility to develop evidence on appropriate drug use and to estimate incremental cost-effectiveness. For appropriate drug use, we examined the feasibility to develop evidence in the following areas: types of treatments and regimes (*'treatments'*), dosages and dose modifications (*'dosages'*), baseline patient characteristics (*'patients'*), reasons for choosing a particular treatment and starting or stopping treatment (*'reasons'*), and treatment outcomes (*'effects' and 'costs'*). To investigate the feasibility of data collection, we examined which data were available through existing databases and which data required retrieval from hospital records. For incremental cost-effectiveness, we investigated the feasibility to obtain comparable patient groups, identify treatment comparators, obtain information from literature, and estimate incremental cost-effectiveness. We explored issues with internal validity, data synthesis and modelling, outcome measures and generalisability.

### Description of the case studies

In the two oxaliplatin studies, patients were identified using the population-based registry of the Dutch Comprehensive Cancer Centres. This registry enabled the identification of all Dutch patients who received chemotherapy (2249 stage III colon cancer patients diagnosed in 2005 and 2006; 1957 metastatic colorectal cancer patients diagnosed in 2003 and 2004). Since this registry did not contain all of the required data, we had to contact individual hospitals. Most of the Dutch hospitals (72%) were approached to expedite the data collection and we continued to include hospitals in order of response until the desired number of patients had been reached (stage III colon cancer: n=391; metastatic colorectal cancer n=316). Using hospital records, additional data were retrospectively collected on baseline patient characteristics, known prognostic information, considerations for choosing a treatment, types of treatments, disease free survival, and overall survival. For a randomly selected subgroup (stage III colon cancer: n=206; metastatic colorectal cancer n=130), detailed data were collected on dosage schemes, adverse effects and all hospital resource use. Table 9.1 and Table 9.2 present relevant findings of the oxaliplatin studies.

In the bortezomib study, patients (n=543) were identified using a trial database (HOVON50 study[230]) for first line treatment. We approached Dutch hospitals for data collection for outcomes research and continued to include hospitals until the desired number

**Table 9.1** Summary table oxaliplatin in stage III colon cancer

| | | | | |
|---|---|---|---|---|
| **Study design** | Retrospective observational study | | | |
| Number of patients | | | | |
| Population-based cancer registry | 2249 (diagnosed in 2005 and 2006) | | | |
| Additional/ detailed data collection (hospital records) | 391/ 206 | | | |
| Number of hospitals visited (% of Dutch hospitals) | 19 (17%) | | | |
| **Treatments** | | | | |
| Treatments received in everyday practice | **FL + oxaliplatin** FOLFOX or CAPOX[a] (n=281) | | **FL alone** 5FU/LV or capecitabine (n=110) | |
| 5FU/LV | 48% | | 15% | |
| Capecitabine | 54% | | 85% | |
| **Dosages**[b] | | | | |
| Oxaliplatin dose according to guidelines | 1,020 mg/m2 | | n/a | |
| Percentage of planned dose given | 81% for FOLFOX<br>71% for CAPOX | | n/a | |
| **Patients** | | | | |
| *Prognostic baseline characteristics* | | *% Missing* | | *% Missing* |
| Eligible for pivotal registration trial | 82% | 14% | 63% | 23% |
| Age [median (range)] | 61 (22–82) | | 73 (41–85) | |
| Co-morbid conditions ≥ 2 | 11% | | 25% | |
| Depth of invasion T2-T3 (T4) | 85% (15%) | | 88% (12%) | |
| Nodes involved N1 (N2) | 60% (40%) | | 66% | |
| Abnormal CEA levels[c] | 18% | 14% | 8% | 22% |
| **Reasons**[b] | | | | |
| Reasons for not prescribing oxaliplatin | Hospital policy (18%), advanced age (21%), patient refusal (19%), poor health status (10%), combination of these factors (7%), specific contra-indication (2%), and unknown (23%) | | | |
| Regimes requiring dose modifications | 56% | | 66% | |
| Toxicity requiring hospitalisation | 7% | | 6% | |
| **Effects** | | | | |
| 2-year disease-free survival probability | | | | |
| Eligible [Mean (95% CI)] | 78.4% (72.5%–84.3%) | | 82.8% (72.5%–93.0%) | |
| Ineligible [Mean (95% CI)] | 56.7% (41.4%–72.0%) | | 83.7% (70.5%–96.8%) | |
| **Costs**[b] | | | | |
| Total costs [mean (median)] | €19,639 (€20,230) | | €5,055 (€4,482) | |
| Minimum – maximum | €1708–€60,149 | | €316–€12,127 | |

[a] *FL: Fluoropyrimidines; FOLFOX: oxaliplatin combined with 5FU/LV; CAPOX: oxaliplatin combined with capecitabine*

[b] *Based on a representative subsample of 206 patients*

[c] *CEA: serum carcinoembryonic antigen levels*

**Table 9.2** Summary table oxaliplatin in metastatic colorectal cancer

| **Study design** | Retrospective observational study | | | | | |
|---|---|---|---|---|---|---|
| Number of patients | | | | | | |
| Population-based cancer registry | 1957 (diagnosed in 2003 and 2004) | | | | | |
| Additional/ detailed data collection (hospital records) | 316/130 | | | | | |
| Number of hospitals visited (% of Dutch hospitals) | 29 (26%) | | | | | |
| **Treatments** | | | | | | |
| First-line treatments received in everyday practice | **FL + oxaliplatin** FOLFOX/CAPOX [a] **(n=92)** | | **FL alone** 5FU/LV [a] or capecitabine **(n=198)** | | **(FL +) irinotecan** alone/FOLFIRI/ CAPIRI [a] **(n=26)** | |
| **Dosages** [b] | | | | | | |
| Mean total cumulative dose of oxaliplatin | 1274 mg | | n/a | | n/a | |
| Patients receiving 2nd (3rd) line treatment | 57% (22%) | | 52% (22%) | | 32% (20%) | |
| **Patients** | | | | | | |
| *Prognostic baseline characteristics* | | *% Missing* | | *% Missing* | | *% Missing* |
| Eligible for clinical trial | 85% | | 63% | | 73% | |
| Age [median (range)] | 60(29–81) | | 64(30–92) | | 59(39–73) | |
| WHO performance status ≥ 2 | 13% | 41% | 23% | 39% | 6% | 35% |
| Abnormal lactate dehydrogenase levels | 46% | 13% | 52% | 20% | 47% | 27% |
| Resection of primary tumour | 65% | 4% | 61% | 3% | 58% | 8% |
| **Reasons** [b] | | | | | | |
| Toxicity in first line | | | | | | |
| Requiring hospitalisation | 6% | | 13% | | 6% | |
| Causing termination of treatment | 20% | | 1% | | 20% | |
| **Effects** | | | | | | |
| Overall Survival in months | | | | | | |
| Eligible [Mean (median)] | 18.6 (14.1) | | 14.9 (11.3) | | 27.1 (21.3) | |
| 95% CI | 15.5–21.8 | | 12.8–17.0 | | 18.0–36.3 | |
| Ineligible [Mean (median)] | 18.4 (17.8) | | 11.0 (6.6) | | 9.2 (6.9) | |
| 95% CI | 12.3–24.5 | | 8.2–13.7 | | 3.6–14.9 | |
| **Costs** [b] | | | | | | |
| Total costs [mean (median)] | €27,711 (€23,172) | | €19,236 (€16,208) | | €39,375 (€38,754) | |
| Minimum – maximum | €2,200–€95,118 | | €462–€65,288 | | €10,258–€109,139 | |

[a] *FL: Fluoropyrimidines; FOLFOX: oxaliplatin combined with 5FU/LV; CAPOX: oxaliplatin combined with capecitabine; FOLFIRI: irinotecan combined with 5FU/LV; CAPIRI: irinotecan combined with capecitabine*

[b] *Based on a representative subsample of 130 patients*

**Table 9.3** Summary table bortezomib in relapsed/ refractory multiple myeloma

| | | | | |
|---|---|---|---|---|
| **Study design** | Retrospective observational study | | | |
| Number of patients | | | | |
| HOVON trial database/ detailed data | 543/ 139 | | | |
| Number of hospitals visited (% of Dutch hospitals) | 41 (38%) | | | |
| **Treatments** | | | | |
| Treatments received in everyday practice | **Ever bortezomib** [a] **(n=72)** | | **Never bortezomib (n=67)** | |
| Bortezomib combination therapy | 71% | | n/a | |
| One other drug (dexamethasone) | 80% (58%) | | n/a | |
| Two/ three or more other drugs | 12%/ 7% | | n/a | |
| **Dosages** | | | | |
| Bortezomib total dose compared to pivotal registration trial | 87% | | n/a | |
| Number of bortezomib cycles in everyday practice (pivotal registration trial) | 4 (6) | | n/a | |
| **Patients** | | | | |
| *Prognostic factors at start of 2nd line treatment* | | *% Missing* | | *% Missing* |
| Age [mean (range)] | 57 (34-69) | | 58 (35-68) | |
| WHO performance status 0/ 1/ ≥ 2 | 59%/ 33%/ 8% | 8% | 35%/ 42%/ 22% | 7% |
| Present with neurotoxicity | 49% | 10% | 26% | 3% |
| Serum B2 (mg/l) [mean (range)] | 4 (1.3–16.7) | 64% | 3 (1.1–5.7) | 79% |
| Albumin (g/l) [mean (range)] | 40 (27.0–59.0) | 31% | 38.4 (16.6–52.0) | 37% |
| Haemoglobin (mmol/l) [mean (range)] | 7.5 (5.0–9.5) | 31% | 7.1 (2.1–10.0) | 31% |
| First line HOVON50 experimental TAD arm | 40% | | 27% | |
| Received allogeneic stem cell transplantation | 27% | 1% | 15% | |
| Maintenance treatment | | | | |
| None/ IFNα | 43%/ 21% | | 63%/ 24% | |
| Thalidomide | 36% | | 13% | |
| Best response first line treatment | | 6% | | |
| Complete response/ Partial & minor response | 16%/ 80% | | 10%/ 76% | |
| No change/ Progressive disease | 3%/ 1% | | 4%/ 9% | |
| Time until first progression in months [median (SD) {range}] | 27.6 (12.8) {2.0–57.9} | 1% | 22.4 (15.8) {1.9–61.4} | 1% |
| **Reasons** | | | | |
| Reasons to start a treatment regime | | [b] | | [b] |
| Regimes requiring dose modifications | 52.5% | | n/a | |
| Due to toxicity | 79% | 14% | n/a | |
| **Effects** | | | | |
| Overall Survival from start of relapsed/ refractory disease in months [mean (median)] | 29.5 (33.2) | | 28 (21.6) | |
| Confidence interval (95%) | 25.1–38.8 | | 14.6–50.4 | |
| **Costs** | | | | |
| Total costs [mean (median)] | €81,626 (€72,182) | | €52,760 (€36,882) | |
| Minimum – maximum | €17,793–€229,783 | | €748–€179,571 | |

[a] *In total 25, 35 and 12 patients received bortezomib in 2nd, 3rd, and 4th line or later, respectively*

[b] *Not part of data collection because this was generally not reported in medical records*

of patients who received off-protocol treatment for relapsed or refractory disease had been reached (n=139). Because many patients (49%) were treated in more than one hospital, we had to collect data in 42 hospitals. Using hospital records, detailed data were retrospectively collected on baseline patient characteristics, known prognostic information, types of treatments, dosage schemes, treatment response, time to progression, time till next treatment, adverse effects, survival and all hospital resource use. Table 9.3 presents relevant findings of the bortezomib study.

## Results

### Feasibility to develop evidence on appropriate drug use

Table 9.4 summarises the results regarding the feasibility to develop evidence on different aspects of appropriate drug use.

**Table 9.4** Feasibility to develop evidence on appropriate drug use

| | **Oxaliplatin in stage III colon cancer** | **Oxaliplatin in metastatic colorectal cancer** | **Bortezomib in multiple myeloma** |
|---|---|---|---|
| Feasibility to use existing databases to identify patients | + | + | + |
| Feasibility to obtain a complete dataset using hospital records | + | +/– | +/– |
| Feasibility to develop evidence on: | | | |
| Treatments | + | + | + |
| Dosages | + | + | + |
| Patients | + | + | +/– |
| Reasons | + | + | – |
| Effects: intermediate and final outcomes | + | + | +/– |
| Effects: safety outcomes | +/– | +/– | +/– |
| Costs | + | + | + |

*+ = good; +/– = moderate; – = poor*

*'Treatments' (types of treatments and regimes)*

In all three studies it was feasible to ascertain the types of treatments used and their regimes using data from hospital records. Both oxaliplatin studies showed that patients were treated in a way that was similar to the regimes used in clinical trials and described in professional guidelines.[231] In contrast, the bortezomib study revealed a high degree of treatment variation.[232] More importantly, treatments differed significantly from those described in both the pivotal registration trial and professional guidelines.

*'Dosages' (treatment dosages and dose modifications)*

Details on dosages and dose modifications were well reported in hospital records. However, retrieval of these details required a great deal of time, which significantly reduced the efficiency of data collection. Both oxaliplatin studies showed that the received dosages were comparable to those observed in clinical trials.[231] The bortezomib study showed that patients received lower dosages (13%) and fewer treatment cycles (4 vs. 6) compared to patients in the pivotal registration trial.

*'Patients' (baseline patient characteristics)*

For the oxaliplatin studies, the cancer registry provided information on age, gender, date of diagnosis, disease stage, and tumour location. Additional data on prognostic baseline characteristics required data from hospital records. For the bortezomib study, the HOVON database provided information on age, gender, date of diagnosis, and disease stage. Other baseline characteristics required data from hospital records. In all three studies it was impossible to compile a complete dataset, including prognostic factors. For example, 13% of serum carcinoembryonic antigen levels (stage III colon cancer) and 40% of performance scores (metastatic colorectal cancer), and 71% of serum β2-microglobulin levels (multiple myeloma) were missing. Nevertheless, based on available baseline characteristics, it seemed like patients treated with oxaliplatin[231] and bortezomib[233] were comparable to trial patients.

*'Reasons' (reasons for starting or stopping a treatment)*

The rationale for choosing a particular treatment was often retrievable (74%) from the hospital records in stage III colon cancer. It was possible to determine the most frequent reasons for dose modifications or treatment interruptions in both oxaliplatin studies. In contrast, in the bortezomib study the reasons to start a treatment, reduce its dose or stop a treatment were often not reported.

*'Effects' (health effects)*

In both oxaliplatin studies, the cancer registry only provided survival data, whereas hospital records provided data on disease-free survival, adverse effects, and survival. Similarly, in the bortezomib study, the HOVON database provided survival data and hospital records data on treatment response, adverse effects, and survival. However, in all three studies, treatment responses and adverse effects were often not reported using standardised outcome measures (e.g., Response Evaluation Criteria in Solid Tumours [RECIST], European Group for Blood and Marrow Transplant [EBMT] response criteria, Common Toxicity Criteria [CTC] toxicity grading scale). Although these results could probably be estimated using for instance laboratory test results, this lack of data severely limited a retrospective assessment using outcome measures as treatment response and

time to progression often found in clinical trials. Lastly, hospital records did not provide any standardised data on quality of life (e.g., European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire [EORTC-QLQ-C30], EuroQol Quality of Life Questionnaire [EQ-5D] or Short Form [SF36] Health Survey).

*'Costs' (costs of a treatment)*

In all three studies it was possible to collect data on hospital resource use of individual patients. However, due to feasibility constraints, unit costs for laboratory services were based on a detailed inventory of a subsample of patients. Similarly, in the bortezomib study detailed data collection on concomitant medication was extremely time-intensive. Therefore, detailed data were only collected for a subsample of 18 patients.

## Feasibility to estimate incremental cost-effectiveness

Table 9.5 summarises the results regarding the feasibility to estimate incremental cost-effectiveness.

**Table 9.5** Feasibility to estimate incremental cost-effectiveness

| | Oxaliplatin in stage III colon cancer | Oxaliplatin in metastatic colorectal cancer | Bortezomib in multiple myeloma |
|---|---|---|---|
| Comparability of baseline characteristics between treatment arms | – | +/– | – |
| Feasibility of using data from everyday practice: | | | |
| To correct for bias | – | +/– | – |
| To identify treatment comparator | + | + | – |
| To estimate incremental cost-effectiveness | – | – | – |
| Comparability of eligible everyday practice patients (treated with oxaliplatin/ bortezomib) and clinical trial patients | + | + | +/– |
| Feasibility of data synthesis: | + | + | – |
| To obtain additional data from the literature on: | | | |
| quality of life | +/– | +/– | – |
| efficacy | + | + | +/– |
| effectiveness | – | +/– | – |
| costs | +/– | +/– | – |
| Feasibility to estimate (using data synthesis): | | | |
| Internally valid incremental cost-effectiveness | + | + | – |
| Precise incremental cost-effectiveness | +/– | +/– | – |
| Externally valid incremental cost-effectiveness | + | + | – |

*+ = good; +/– = moderate; – = poor*

*Case 1: Oxaliplatin in stage III colon cancer*

Due to the strong preference of physicians to use oxaliplatin whenever indicated, patients receiving the comparator treatment were significantly different regarding important prognostic factors. To correct for the resulting confounding, different adjustment techniques were applied to the Cox multivariate regression model, such as average covariate adjustment, regression adjustment by propensity score matching, and survival analysis matched on propensity score matching. However, our sample size (n=391) was not powered for this purpose. This, in combination with missing data on prognostic factors, resulted in possibly biased estimates with wide confidence intervals. It was not feasible to estimate incremental cost-effectiveness using only everyday practice data. Therefore, we developed a Markov model to estimate incremental cost-effectiveness. In this model we synthesised effectiveness data from everyday practice with efficacy data from the pivotal registration trial. Patients were categorised 'eligible' or 'ineligible', depending on whether the patients fulfilled the trial eligibility criteria. Ineligible patients (18%) had a worse prognosis compared to eligible patients (82%), but trial patients and eligible case study patients had similar two-year disease-free survivals (80% vs 78%). Effectiveness of the comparator was modelled using trial results. All costs were based on the case study. Applying scenario analyses, incremental cost-effectiveness ratios ranged from €8,247 to €12,289 per quality adjusted life year. Sensitivity analyses of input parameters and model assumptions produced little differences, supporting robustness of the results. Data synthesis resulted in internally valid incremental cost-effectiveness estimates generalisable to Dutch everyday practice.

*Case 2: Oxaliplatin in metastatic colorectal cancer*

As with case 1, patients receiving oxaliplatin were not comparable to patients not receiving oxaliplatin. In this case, the differences in baseline prognosis were less pronounced than in stage III colon cancer, but correction for confounding was hindered by missing values. Although not performed, modelling evidence from the literature with evidence from the case study would have been feasible.

*Case 3: Bortezomib in relapsed or refractory multiple myeloma*

Rapid developments in treatment for multiple myeloma resulted in great heterogeneity. Patients treated with bortezomib were not comparable to other patients regarding prognostic factors. It was impossible to identify a single treatment comparator; more than 10 drugs were given in more than 20 different combinations. Therefore, our comparator included any treatment besides bortezomib. Similar to the oxaliplatin cases, different adjustment techniques were applied to the Cox multivariate regression model to obtain a valid overall survival estimate. However, none succeeded in correcting for the observed confounding. New evidence from extended follow-up and other trials com-

paring different treatments and combinations became available. No information was published on treatment-related costs or quality of life. The great heterogeneity caused by many treatment arms made it impossible to develop a feasible model to estimate incremental cost-effectiveness.

## Discussion

We investigated the feasibility of different aspects of outcomes research in oncology. Our results show that the degree of feasibility depends on both the aspect and treatment indication. To our knowledge, this is the first feasibility study of outcomes research in oncology that is based on empirical evidence.

Based on theoretical expectations and expert opinion, the lack of a randomised controlled setting is one of the major concerns.[219,222-224] As expected, our results confirm that heterogeneity resulted in incomparable patient groups and the inability to correct for confounding. Therefore, it was not possible to estimate incremental cost-effectiveness only using everyday practice data. However, our results also show that it may still be feasible to obtain internally valid and generalisable incremental estimates by synthesising everyday practice data with trial data, provided that everyday practice patients fulfil the eligibility criteria of trials in which these drugs were tested. Furthermore, our results confirm that current databases do not provide sufficient information.[212,213,216,220,226] The need for additional data required the retrieval and scrutiny of hospital records. Regarding applicability of outcome measures,[25,33] our results show that measures used in clinical trials are susceptible to bias due to missing data and the lack of standardisation in their reporting in hospital records. The choice of relevant outcome measures depends on the disease. For example, survival is often the primary outcome measure in oncology, but not in diseases such as rheumatoid arthritis or COPD where quality of life is more relevant. Moreover, timeliness[26,32,227] can differ per drug and disease. While a three-year time frame was sufficient for the oxaliplatin studies, the bortezomib study revealed that treatment advances limited the relevance of the gathered evidence with such a time frame. The challenge of generalisability[228] might be of lesser concern. In our studies, the ability to select representative samples (e.g., by means of the cancer registry), was a key to ensuring generalisability.

The feasibility of outcomes research also depends on its study design. The main limitation of our case studies was the use of retrospective research designs. As a consequence, we faced a great deal of important missing information. A prospective design, using a registry, would offer greater control over data collection as well as the opportunity to

collect data on quality of life. However, in many prospective designs, including registries, data are still retrospectively collected and rely on information provided by others (e.g., physicians, research assistants). Moreover, such a design would still not solve the issue of randomisation. Although a pragmatic trial would be a solution for this, these trials are often impossible due to ethical or feasibility considerations.

Our study was only based on three case studies. However, we believe that our findings can be extended to other oncological diseases. We intentionally selected different indications in cancer reflecting different types of disease populations (small vs. large), expectations regarding practice variation (small vs. large), and relevant outcome measures (intermediate vs. final endpoints).

Our study provides important insight into the implementation of evidence development schemes. We believe that data from everyday practice results in valuable evidence, addressing uncertainties arising from the gap between clinical trials and everyday practice. Above all, it is essential to have a comprehensive understanding of the disease and the treatment effect and this requires interdisciplinary collaboration.

Active interdisciplinary collaboration will result in an enhanced research design focusing on feasible objectives for a particular treatment in a specific indication. It will also reduce problems with missing information and lack of standardisation in reporting. Because current databases do not provide sufficient information, patient registries can offer an opportunity to build new research infrastructures. Although patient registries cannot resolve all issues, if they are used by an active interdisciplinary collaborative research group, they could increase efficiency of data collection and help to reduce issues of generalisability, incomparability of patient groups, missing information, and lack of standardisation in reporting. For orphan drugs, international registries may be the best means to obtain a sufficient amount of evidence (e.g., Pompe Registry[234]). Furthermore, registries can also be used to monitor and improve quality of care beyond outcomes research.

In conclusion, our results show that it is feasible to generate evidence about drug use in everyday oncology practice. For some aspects of appropriate drug use, this will require improvements in reporting in hospital records or compiling data in registries. The feasibility to estimate incremental cost-effectiveness depends on the drug and its indication. We believe that in most circumstances it is inevitable to synthesise data to obtain valid and precise estimates. However, it is essential to realise that for some drugs and indications, it may sometimes be impossible to estimate sufficiently valid and precise incremental cost-effectiveness.

In the end, the generation of more evidence will improve the quality of decisions made by both payers and physicians. This, in turn, can improve quality of care and enhance the efficient allocation of resources and thereby help to ensure long-term sustainability of health care systems.

## Acknowledgements

This study was funded by a grant of the Dutch Health Care Insurance Board (CVZ). The sponsor had no role in the preparation, review or approval of the manuscript. The views and opinions expressed here are those of the authors and do not necessarily reflect those of the funders.

# 10

## Policymaker, please consider your needs carefully: Does outcomes research in relapsed or refractory multiple myeloma reduce policymaker uncertainty regarding value for money of bortezomib?

With Jennifer Gaultney, Hedwig Blommestein, Peter Huijgens, Pieter Sonneveld, Ken Redekop, and Carin Uyl-de Groot

Value in Health 2014; 17(2):245-253

## Abstract

Dutch policy regulations require outcomes research for the assessment of appropriate drug use and cost-effectiveness after four years of temporary reimbursement. We investigated whether outcomes research reduced policymaker uncertainty regarding the question whether the costs are worth public funding.

Our cohort study included 139 patients with relapsed/refractory multiple myeloma who were treated outside of a clinical study; 72 received bortezomib and 67 did not receive bortezomib. Detailed data were retrospectively collected from medical records in 38% of Dutch hospitals.

All patients received second-line treatment; 65%, 40%, and 14%, received three, four, five or more lines of therapy. Neither a specific treatment sequence nor an appropriate comparator could be identified because of large variation in regimes. Kaplan-Meier curves showed an increased overall survival (mean [median] OS 29.5 [33.2] vs. 28.0 [21.6] months) for bortezomib patients (Wilcoxon p=0.01). Total mean costs were €81,626 (range: €17,793–229,783) and €52,760 (range: €748–179,571) for patients receiving bortezomib and patients not receiving bortezomib, respectively. Patients treated with bortezomib, however, were not comparable to other patients despite attempts to correct for confounding. Therefore, it was impossible to develop a feasible model to obtain a valid incremental cost-effectiveness estimate.

It was possible to develop evidence on bortezomib's use, effects and costs in everyday practice. Much uncertainty, however, remained regarding its cost-effectiveness. Policymakers should carefully consider if outcomes research sufficiently decreases uncertainty or whether other options (e.g., finance- and/or outcomes-based risk-sharing arrangements) are more appropriate to ensure sufficient value for money of expensive drugs.

## Introduction

Rising health care expenditures are making it extremely difficult to manage early access to promising innovative, often expensive, drugs while ensuring value for money. Globally, health care systems have therefore introduced policies to reduce initial decision makers' uncertainty regarding the clinical and economic performance of novel drugs. These policies address clinical and/or finance uncertainty, for example by means of finance-based[34,35] or outcomes-based[35,36] risk sharing agreements such as coverage with evidence development schemes[29-31] or outcomes research requirements.[32,33]

Although outcomes research and evidence development requirements increasingly seem an attractive policy option, many unanswered questions remain regarding their actual value[31,35,235] and feasibility.[215,218,236] In The Netherlands, outcomes research requirements were first implemented in 2006 for expensive inpatient drugs. From 2013 onwards, this policy has been extended to specific groups of outpatient drugs. In the Dutch coverage with evidence development policy, early access is linked with the obligation to conduct outcomes research in accordance to guidelines,[193] namely to gather data in everyday practice on appropriate drug use (e.g., patient characteristics, types of treatments, dosages, and dose modifications) and real-world cost-effectiveness. After four years of use, a reassessment will determine whether or not the drug will continue to be reimbursed.[229] Notably, recent Dutch experiences revealed insufficient data to perform a reassessment after four years of outcomes research (i.e., omalizumab, infliximab, and ranibizumab).

In 2006, bortezomib was added on the expensive drug list for relapsed/refractory multiple myeloma, an incurable malignant plasma cell disorder. At the time of the initial reimbursement decision, Dutch policymakers only had information from one pivotal phase III trial,[196] which found bortezomib to be superior to high dose dexamethasone in terms of increased time to progression (6.22 vs. 3.49 months), response rates (38% vs. 18%), response duration (8 vs. 5.6 months), and one-year survival rate (80% vs. 66%). Costs were estimated at €27,432 per treated patient, which was solely based on the price of bortezomib vials; no data on cost-effectiveness were available.[237] Despite favourable trial results, the scarcity in available evidence (i.e., one phase III trial in 669 patients) implied a high degree of uncertainty for policymakers regarding bortezomib's value in everyday practice in terms of real-world effectiveness, health care costs, and cost-effectiveness. Because bortezomib was added on the expensive inpatient drug list, outcomes research needed to be conducted to facilitate a re-evaluation of the initial reimbursement decision.

This article describes our experiences in The Netherlands in performing outcomes research of bortezomib in relapsed/refractory multiple myeloma. We investigated whether outcomes research reduced initial policymaker uncertainty regarding real-world use, effectiveness, health care costs, and cost-effectiveness after data collection in everyday practice. To our knowledge, this is the first study evaluating cost-effectiveness of bortezomib based on real-world data only.

## Methods

### Patient population and data collection

To identify patients who were eligible for bortezomib treatment in everyday practice, we selected our patient population from patients previously enrolled in a clinical trial (HOVON50). The phase III HOVON50 trial enrolled 556 (543 Dutch) patients from November 2001 to June 2005 to investigate the treatment effect of thalidomide in patients aged 18–65 years newly diagnosed with Durie-Salmon stage II/III multiple myeloma.[230] Patients who went off-protocol from this trial regime no longer received protocol-based therapy and were therefore eligible for our outcomes research study because they were treated for relapsed/refractory multiple myeloma in everyday clinical practice.

We approached Dutch hospitals to obtain permission for data collection. We continued to include hospitals until the desired number of patients who received off-protocol treatment for relapsed/refractory disease had been reached. Power calculations (two-sided, $\alpha$=0.05, power= 0.7) of the desired sample size (n>124) were based on differences in response percentages (0.38 vs. 0.18) in the Assessment of Proteasome inhibition for Extending Remissions (APEX) trial.[196] In total, 139 patients were included; 72 received bortezomib and 67 did not receive bortezomib. Because many patients (49%) were treated in more than one hospital, data were collected in 42 hospitals (38% of all Dutch hospitals, and approximately 57% of Dutch hospitals treating haemato-oncology patients). Figure 10.1 shows the flowchart of the patient selection process.

Detailed data for outcomes research were retrospectively collected from hospital records from the time of first relapsed/refractory disease until end of follow-up. Data were collected on baseline patient characteristics, types of treatments and regimes, dosage schemes, adverse effects, treatment response, response rate, time to progression, time till next treatment, survival and resource use.

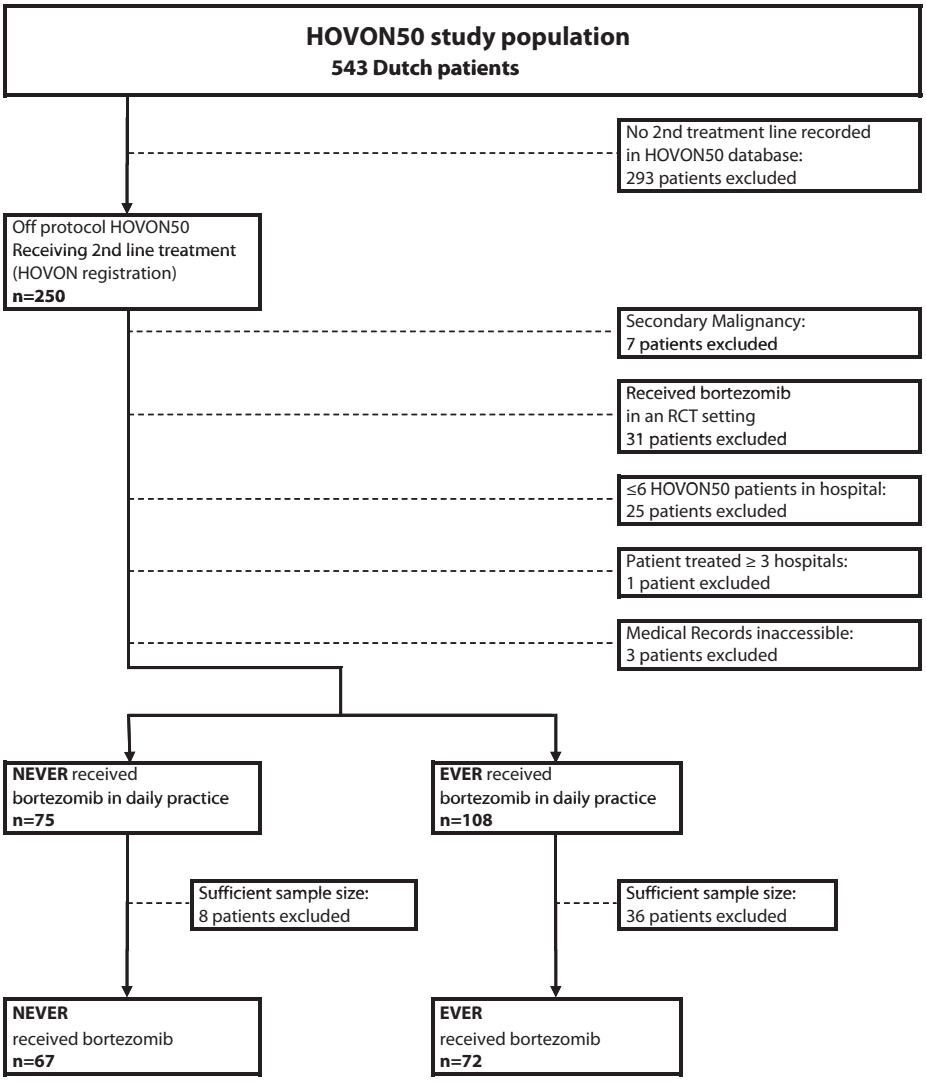


**Figure 10.1** Flowchart of the patient selection process

## Drug use and real-world cost-effectiveness

To assess drug use, we examined baseline patient characteristics, types of treatments received, dosages, and dose modifications. To estimate overall survival (OS) and time to next treatment, Kaplan-Meier curves were computed from start of relapsed/refractory treatment stratified by receipt of bortezomib. Different adjustment methods, such as average covariate adjustment, regression adjustment by propensity scores, and matched analysis, were applied to the Cox multivariate regression model to correct for differences

in baseline characteristics between patients receiving bortezomib and patients not receiving bortezomib.

Treatment costs were computed from a hospital perspective. Costs for individual patients were determined by applying unit costs to individual resource use of the following cost components: outpatient, emergency room, and day-ward visits; hospital admissions; consultations by telephone; radiotherapy; (surgical) procedures; laboratory services; medical imaging services; treatment; and concomitant treatment. One-way sensitivity analyses were carried out by varying the unit costs of hospital visits (inpatient care, outpatient visits, and day-care treatment) between 50% and 150%. Details of the unit costs and cost-analysis are reported elsewhere.[238]

For cost-effectiveness, we investigated the feasibility of obtaining comparable patient groups, identifying treatment comparators, and estimating (incremental) cost-effectiveness.

Statistical analysis was conducted with the statistical software program SAS, version 9.1 (SAS Institute Inc., Cary, NC).

## Results

### Baseline patient characteristics

Missing values on baseline characteristics were common. Low numbers of available prognostic data occurred, for example, for serum β2-microglobulin levels (71% missing), albumin levels (34% missing), performance status (8% missing), and neurotoxicity assessment (6% missing).

Based on available data, baseline characteristics at start of relapsed/refractory treatment differed between patients treated and not treated with bortezomib (see Table 10.1). Significant differences were observed for the proportion of patients presenting with neurotoxicity (p=0.01), WHO performance status (p=0.03), type of maintenance therapy (p=0.01) and time until first progression (p=0.03). As a result, prognosis at start of relapsed/refractory disease varied greatly between both the patient groups.

### Types of treatments received, dosages and dose modifications

Treatment details including type of treatment, dosages and dose modifications were well reported in hospital records. On account of the rapid advances in recent years in treatment options available for multiple myeloma, variation was observed in treatments

**Table 10.1** Baseline characteristics at start of relapsed/ refractory treatment

| Baseline characteristics | Received bortezomib (n = 72) | | Never received bortezomib (n = 67) | | P-value† |
|---|---|---|---|---|---|
| | | % Missing | | % Missing | |
| ***Patient-related characteristics*** | | | | | |
| Age [mean (range)] | 57 (34–69) | | 58 (35–68) | | 0.37 |
| Female | 44% | 0% | 39% | 0% | 0.60 |
| WHO performance status 0/ 1/ ≥2 | 59%/ 33%/ 8% | 8% | 35%/ 42%/ 22% | 7% | 0.03* |
| Albumin (g/l) [mean (range)] | 40.0 (27.0–59.0) | 14% | 37.8 (16.6–52.0) | 16% | 0.11 |
| Serum B2 (mg/l) [mean (range)] | 4.1 (1.3–16.7) | 64% | 3.0 (1.1–5.7) | 79% | 0.20 |
| C-reactive protein (mg/l) [mean (range)] | 11.4 (1–67) | 67% | 31.6 (1–171) | 69% | 0.07 |
| Creatinine clearance (mmol/l) [mean (range)] | 7.0 (1.9–16.0) | 79% | 8.5 (2.3–16.0) | 90% | 0.42 |
| Haemoglobin (mmol/l) [mean (range)] | 7.5 (5.0–9.5) | 31% | 7.1 (2.1–10.0) | 31% | 0.16 |
| Platelet count (x$10^9$/l) [mean (range)] | 213 (10–657) | 1% | 227 (28–828) | 1% | 0.54 |
| Plasma cell infiltration > 50% | 28% | 56% | 19% | 46% | 0.57 |
| Neurotoxicity present | 44% | 10% | 25% | 3% | 0.01* |
| ***Previous treatment-related characteristics*** | | | | | |
| First line HOVON50 experimental TAD arm | 40% | 0% | 27% | 0% | 0.11 |
| Received stem cell transplantation | 26% | 0% | 15% | 0% | 0.14 |
| Maintenance therapy | | | | | |
| None/ IFNα | 43%/ 21% | 0% | 63%/ 24% | 0% | 0.01* |
| Thalidomide | 36% | | 13% | | |
| Best response to first line treatment | | 6% | | 6% | 0.12 |
| Complete response/ Partial response | 15%/76% | | 11%/ 63% | | |
| Minor response/ No change | 5%/3% | | 11%/ 5% | | |
| Progressive disease | 1% | | 10% | | |
| Reason for going off protocol HOVON50 | | 0% | | 0% | |
| Normal completion | 25% | | 13% | | 0.09 |
| Excessive Toxicity | 21% | | 28% | | 0.33 |
| Progression/Relapse | 32% | | 28% | | 0.71 |
| Time until first progression [mean (range) in months] | 27.6 (2.0–57.9) | 1% | 22.4 (1.9–61.4) | 1% | 0.03* |

*†Continuous variables were compared by Kruskall-Wallis test, and either Pearson's chi-sqaure or Fisher's exact test was used to compare categorical variables across all groups; *significant at α=0.05*

received by patients. Table 10.2 shows the number of patients receiving treatment by treatment line. All 139 patients received second-line treatment, 65% received third-line, 40% fourth-line, 14% fifth-line, 6% sixth-line, 2% seventh-line and 1% eighth-line treatment. Because of a large degree of variation, it was impossible to identify a general treatment pattern. Nevertheless, the percentage of patients treated with thalidomide decreased over the lines, whereas lenalidomide usage increased. Of all patients receiving bortezomib, 79% were previously treated with thalidomide, which coincided with Dutch treatment guidelines. Six patients received bortezomib in more than one line.

As Table 10.2 reveals, a combination of treatments was common practice; more than 10 drugs were given in more than 20 different combinations. The most frequent combinations were thalidomide/dexamethasone (n = 57), lenalidomide/dexamethasone (n = 38), melphalan/prednisone (n = 32) and vincristine/adriamycin/dexamethasone (n = 22).

Bortezomib was given as mono-therapy in 29% and as combination therapy in 71% of the administrations. It was combined with one other treatment in 58%, two other treatments in 9% and three or more other treatments in 5% of the administrations. It was most often combined with dexamethasone (41%). Most of the patients were treated in cycle regimes similar to the pivotal registration trial (i.e., APEX trial[196,239]). Patients in everyday practice, however, received fewer treatment cycles (4 vs. 6) as well as lower dosages (13%).

It was not feasible to establish a pattern for dose modifications according to toxicities. Often, no reason for dose modification was reported or physicians only reported that the condition of the patient required a dose modification without describing the reason for poorer condition. In total, 53% of bortezomib regimes required a dose modification. As expected, the most common reported toxicity was neurotoxicity (61%).

**Treatment effects**

Policymakers generally prefer OS and quality adjusted life years (QALYs) as outcome measures in reimbursement decision making[240] Therefore, OS from start of relapsed/refractory disease was used to analyse the treatment effect of bortezomib. Moreover, using either time to progression or progression-free survival as effectiveness measures, which is usual in clinical trials, was deemed inappropriate because physicians seemingly used less strict criteria in comparison to clinical trials, which dictate response criteria.

The mean follow-up duration was 26.0 (SD 14.4) and 21.5 (SD 16) months for patients treated and patients not treated with bortezomib, respectively. At the end of data collection, 37 patients treated with bortezomib and 31 patients not treated with bortezomib

**Table 10.2** Treatments received by treatment line

| | All 139 patients | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Line 2 | | Line 3 | | Line 4 | | Line 5 | | Line 6 | | Line 7 | | Line 8 | |
| **Treatment** | (N = 139) | | (n = 90) | | (n = 55) | | (n = 20) | | (n = 8) | | (n = 3) | | (n = 2) | |
| Bortezomib | 25 | 18% | 35 | 39% | 12 | 22% | 6 | 30% | 1 | 13% | 1 | 33% | 0 | |
| Lenalidomide | 4 | 3% | 14 | 16% | 21 | 38% | 6 | 30% | 5 | 63% | 1 | 33% | 1 | 50% |
| Thalidomide | 73 | 53% | 15 | 17% | 8 | 15% | 3 | 15% | 1 | 13% | 0 | | 1 | 50% |
| Adriamycin | 17 | 12% | 10 | 11% | 4 | 7% | 2 | 10% | 0 | | 2 | 67% | 0 | |
| Vincristine | 11 | 8% | 6 | 7% | 4 | 7% | 2 | 10% | 0 | | 0 | | 0 | |
| Melphalan | 21 | 15% | 7 | 8% | 5 | 9% | 2 | 10% | 1 | 13% | 0 | | 0 | |
| High dose melphalan (HDM) | 9 | 6% | 4 | 4% | 1 | 2% | 0 | | 0 | | 0 | | 0 | |
| Dexamethasone | 80 | 58% | 52 | 58% | 32 | 58% | 9 | 45% | 5 | 63% | 2 | 67% | 0 | |
| Prednisone | 28 | 20% | 12 | 13% | 13 | 24% | 10 | 50% | 4 | 50% | 2 | 67% | 1 | 50% |
| Cyclophosphamide | 14 | 10% | 9 | 10% | 14 | 25% | 6 | 30% | 1 | 13% | 1 | 33% | 1 | 50% |
| Donor lymphocyte infusion (DLI) | 19 | 14% | 11 | 12% | 4 | 7% | 2 | 10% | 1 | 13% | 0 | | 0 | |
| Stem cell transplantation (allo+auto) | 19 | 14% | 7 | 8% | 2 | 4% | 1 | 5% | 0 | | 0 | | 0 | |
| Interferon alpha | 0 | | 2 | 2% | 0 | | 0 | | 1 | 13% | 0 | | 0 | |
| Experimental | 1 | 1% | 1 | 1% | 0 | | 0 | | 0 | | 0 | | 0 | |
| Other | 1 | 1% | 2 | 2% | 3 | 5% | 2 | 10% | 0 | | 0 | | 0 | |
| Totals | 322 | | 187 | | 123 | | 51 | | 20 | | 9 | | 4 | |

were still alive. Kaplan-Meier curves (see Figure 10.2) from start of relapsed/refractory treatment showed a longer mean (29.5 vs. 28.0 months) and median (33.2 vs. 21.6 months) OS for patients receiving bortezomib (Logrank p=0.31; Wilcoxon p=0.01). The crossing of curves might be due to the low number of patients still in follow-up after approximately 36 months (i.e.,14 patients in each group). It could also be related to great heterogeneity within the patient groups or between the groups (i.e., patients groups are incomparable).

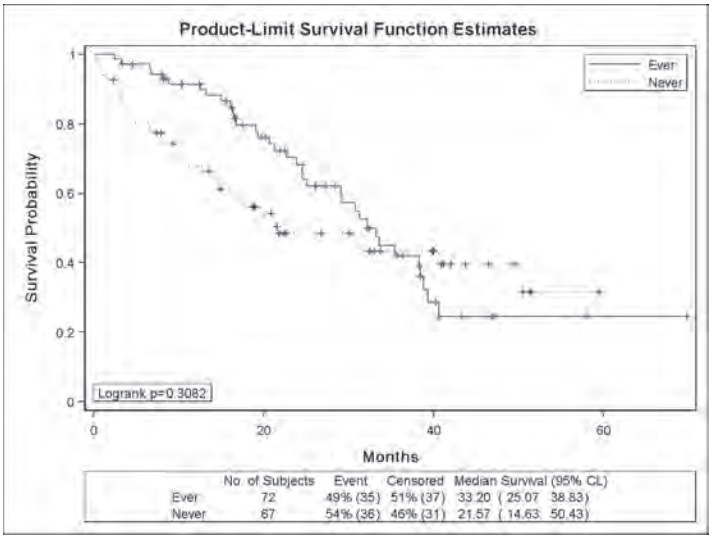


**Figure 10.2** Kaplan-Meier OS curves from start of relapsed/ refractory treatment stratified by treatment with bortezomib

Previous research found that receiving thalidomide as first-line treatment is associated with reduced OS after relapsed/refractory treatment.[241,242] Therefore, we stratified the Kaplan-Meier curves [not shown] by HOVON50 treatment arm (Thalidomide, Adriamycin, and Dexamethasone (TAD) arm vs. Vincristine, Adriamycin and Dexamethasone (VAD) arm). This revealed an increased survival (Logrank p=0.056; Wilcoxon p=0.015) in favour of the non-experimental HOVON50 arm (mean [median] OS for patients treated with bortezomib in the VAD arm 30.9 [33.6] and the TAD arm 27.0 [29.2] months; and for patients not treated with bortezomib in the VAD arm 29.9 [31.1] and the TAD arm 13.7 [15.9] months). Moreover, differences (Logrank p=0.41; Wilcoxon p=0.04) in survival [not shown] were also found between the treatment lines in which bortezomib was administered (mean OS 18.8, 31, 31.6 months, and median OS 24.0, 35.4, 32.5 months for receiving bortezomib in second-, third-, or fourth-line or later, respectively). As Figure 10.3 shows, however, further stratifying all the four groups by HOVON50 treatment arm

resulted in a statistically insignificant effect on OS (Logrank p=0.16; Wilcoxan p=0.08). This was mainly due to the small number of observations in each group (numbers ranged from 5 to 25 in the bortezomib groups).

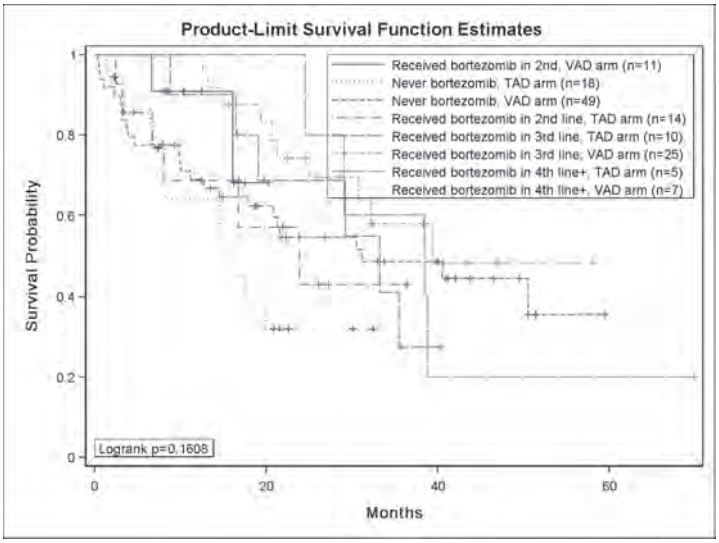


**Figure 10.3** Kaplan-Meier OS curves form start of relapsed/ refractory treatment stratified by treatment with bortezomib, bortezomib treatment line and HOVON50 treatment arm

Despite applying different adjustment techniques (i.e., average covariate adjustment, regression adjustment by propensity scores and matched analysis) to the Cox multivariate regression model, none succeeded in correcting for differences between patient groups. This suggests that residual confounding by indication exists on account of missing information. Consequently, patients receiving bortezomib were incomparable to patients not receiving bortezomib and thus any comparison between the groups would be invalid.

### Treatment costs

Table 10.3 presents the total mean costs for patients treated with bortezomib (n = 72) and patients not treated with bortezomib (n = 67). Total mean costs for patients treated with bortezomib amounted to €81,626 but varied widely between patients (range: €17,793 to €229,783). Active treatment (costs excluding stem cell transplantation €30,733; SD €24,654) was the most important cost driver accounting for 44% of total costs. Bortezomib accounted for 57% and lenalidomide for 35% of the active treatment costs. Total mean costs for patients receiving bortezomib in second-line (n = 25), third-line (n = 35), and fourth-line or later (n = 12) were €53,726, €95,962, and €97,937, respectively. These

**Table 10.3** Total mean costs from start of relapsed/ refractory treatment stratified by treatment with bortezomib

| | Ever received bortezomib (n = 72) | | Never received bortezomib (n = 67) | |
|---|---|---|---|---|
| **Costs (Euro 2009)** | **Mean** | **SD** | **Mean** | **SD** |
| Hospital admissions | | | | |
| *Haematology/ internal/ surgical ward* | € 12,294 | € 13,750 | € 12,168 | € 13,843 |
| *Intensive care unit* | € 607 | € 2,909 | € 2,297 | € 6,071 |
| Hospital visits | | | | |
| *Outpatient* | € 5,676 | € 3,902 | € 3,732 | € 4,023 |
| *Day-care* | € 4,799 | € 2,993 | € 1,132 | € 1,653 |
| *Emergency room visits* | € 160 | € 216 | € 65 | € 111 |
| *Telephone consults* | € 51 | € 65 | € 30 | € 53 |
| Radiotherapy | € 1,971 | € 3,139 | € 1,698 | € 2,623 |
| Surgery | € 766 | € 2,058 | € 1,383 | € 3,035 |
| Diagnostics (e.g., laboratory & radiology) | € 7,497 | € 5,246 | € 6,417 | € 7,264 |
| Concomitant medication | € 13,103 | € 8,855 | € 9,017 | € 8,637 |
| *Acute* | € 2,521 | € 4,882 | € 802 | € 1,683 |
| *Chronic/ prophylactic* | € 10,582 | € 7,591 | € 8,215 | € 8,398 |
| Therapy | € 37,118 | € 28,790 | € 16,496 | € 28,140 |
| *Bortezomib* | € 17,407 | € 11,143 | € 0 | € 0 |
| *Lenalidomide* | € 10,769 | € 18,062 | € 8,923 | € 24,282 |
| *Thalidomide* | € 514 | € 708 | € 818 | € 957 |
| *Dexamethasone* | € 103 | € 127 | € 67 | € 69 |
| *Adriamycin* | € 133 | € 268 | € 39 | € 105 |
| *Vincristine* | € 25 | € 61 | € 6 | € 20 |
| *Melphalan* | € 76 | € 202 | € 67 | € 124 |
| *Prednisone* | € 21 | € 78 | € 7 | € 15 |
| *Interferon alpha* | € 22 | € 134 | € 6 | € 49 |
| *Cyclophosphamide* | € 17 | € 40 | € 9 | € 23 |
| *Donor leukocyte infusions* | € 1,594 | € 3,015 | € 467 | € 1,310 |
| *Stem cell transplantation* | € 3,969 | € 13,313 | € 4,412 | € 10,937 |
| *Other* | € 52 | € 269 | € 0 | € 0 |
| **Total costs** | | | | |
| Mean | € 81,626 | € 47,246 | € 52,760 | € 45,865 |
| Minimum | € 17,793 | | € 748 | |
| Maximum | € 229,783 | | € 179,571 | |

*SD = standard deviation*

differences were most likely because most of the patients (68%) treated in second line were still in follow-up at the time of data collection compared to 54% and 8% of patients in third-line and fourth-line or later, respectively.

Total mean costs for patients not treated with bortezomib amounted to €52,760 and also varied widely between patients (range: €748 to €179,571). The most expensive patients consumed substantially high proportions of their total costs for hospital stays, resource use, and active treatment. Inpatient hospital days (€12,168; SD €13,843) was the most important cost driver (23%), followed by active treatment (€14,821; costs excluding stem cell transplantation €10,409; SD €24,340). Lenalidomide (€8,923; SD €24,282) accounted for 60% of the active treatment costs and 17% of the total costs; stem cell transplant (€4,412; SD €10,937) accounted for 30% of the active treatment costs and 8% of the total costs.

One-way sensitivity analysis by varying the unit costs of inpatient hospital days, day-care treatments, and outpatient visits appeared to have a rather modest effect on the total mean costs. The greatest effect was obtained by varying the unit price for inpatient hospital days (range for patients treated with bortezomib €75,176–€88,076; range for patients not treated with bortezomib €45,527– €59,993).

**Real-world cost-effectiveness**

Because of great differences in baseline prognosis, the inability to correct for these differences, and extensive treatment variation, it was impossible to develop a feasible model to obtain valid and precise incremental cost-effectiveness estimates of bortezomib compared to other treatments. Without the intention to make direct comparisons, however, it was possible to estimate costs per month of survival for patients receiving bortezomib and patients not receiving bortezomib.

The costs from start of relapsed/refractory treatment for patients treated with bortezomib were €2,767 per month of survival (total mean costs: €81,626; mean OS: 29.5 months). Similarly, for patients treated with bortezomib in second-line, the costs were €2,858 per month of survival (total mean costs: €53,726; mean OS: 18.8 months). Costs for patients receiving bortezomib in third-line and fourth-line or later were €3,096 (total mean costs: €95,962; mean OS 31.0 months) and €3,099 (total mean costs: €97,937; mean OS 31.6 months) per month of survival, respectively.

The costs from start of relapsed/refractory treatment for patients not treated with bortezomib were €1,884 per month of survival (total mean costs: €52,760; mean OS: 28.0 months).

## Discussion

Despite favourable findings of bortezomib's registration study, there was a high degree of uncertainty for policymakers whether the high drug costs were worth public funding. Although outcomes research and evidence development requirements globally seem to be popular as well as promising policy options to reduce decision maker uncertainty,[29,32,243] our results show that its actual value might depend on the type of evidence required and type of uncertainty addressed.

The reimbursement decision was based on one phase III trial.[196] No data were available on long-term survival and health care costs besides the price of bortezomib vials. Consequently, policymakers were uncertain of bortezomib's effects, costs and cost-effectiveness in everyday practice. The registration trial compared bortezomib with high-dose dexamethasone. In contrast, outcomes research showed that treatment in clinical practice was far more heterogeneous. Although real-world patients received fewer treatment cycles (4 vs. 6) as well as lower dosages (13%) compared with trial patients, time to progression (6.8 vs. 6.22 months) and response rates (complete response: 8% vs. 6%; very good, partial, and minimal response: 55% vs. 41%) seemed reasonably similar. One-year survival rate, however, was lower in everyday clinical practice (66% vs. 80%). (A detailed comparison between our real-world patients and trial patients is reported elsewhere[244]). Furthermore, outcomes research showed detailed health care costs beyond the price of bortezomib itself. Thus, outcomes research provided valuable information on types of treatments received, which patients received or did not receive bortezomib, dosages, dose modifications, (overall) survival, treatment costs, and costs per month of survival. Hence, outcomes research reduced initial policymaker uncertainty about bortezomib's use, effects and costs in everyday practice.

Outcomes research, however, did not reduce the uncertainty of the societal value of bortezomib compared with other treatments. Because of extensive treatment variation, it was not possible to identify appropriate treatment comparators. Furthermore, as expected, our results confirm previous concerns[219,222-224] that great heterogeneity and a lack of randomisation in everyday practice resulted in incomparable patient groups. Although other observational studies successfully used the propensity score matching technique,[245-249] essential prerequisites,[250] such as large patient numbers and consistency in comparator, were missing in our study. Despite applying different adjustment techniques to the Cox multivariate regression model, none succeeded in correcting for differences between patient groups mainly on account of small patient numbers, extensive treatment variation and missing data. Consequently, we concluded that it was impossible to compare patients receiving bortezomib and patients not receiving

bortezomib; any comparison between the groups would be invalid. Therefore, a feasible model to estimate real-world incremental cost-effectiveness of bortezomib compared to other treatments remains to be demonstrated. Only with a feasible model it would be useful to perform uncertainty analysis of input parameters to report the uncertainty surrounding the outcomes (e.g., stochastic, parameter, and structural uncertainty[251]).

At the time of reassessment, policymakers could, besides our outcomes research results, make use of published literature providing information from various studies describing the efficacy of bortezomib as mono-therapy or combination therapy as well as describing the efficacy of other new multiple myeloma therapies. Only a few cost studies[252-254] and economic evaluations[255-257] were published in relapsed/refractory multiple myeloma. Previous cost-studies, however, were based on conventional therapies,[252] did not apply micro-costing techniques,[253] nor provided information on the use of novel agents.[254] Previous economic evaluations were based on synthesising data and expert opinions and did not use patient level data.[255-257] Therefore, our results provided, to our knowledge, the first results based on real-world data only.

We believe that our results illustrate the value of outcomes research as well as its challenges and thus provide important lessons for policymakers. We acknowledge that we base our conclusions on one outcomes research study in multiple myeloma. Therefore, our conclusions might not be generalisable to outcomes research for all other drugs. Our findings, however, regarding missing data, incomparability of patients, treatment heterogeneity due to rapid treatment advances, and low patient numbers are most likely generalisable to other drugs in comparable diseases. A limitation of our study was that we used the HOVON50 population to select patients who received bortezomib outside of a clinical trial. Many patients, however, received bortezomib within a clinical trial. Consequently, we might have induced selection bias. This is, however, partly a consequence of only using everyday practice data. Even if we would have increased our sample size, outcomes research may be infeasible for a low prevalence disease. Another limitation was the use of a retrospective research design. Because of this, we faced a great deal of important missing information and we could not collect data on quality of life. Although we believe that a prospective design, using a registry, would offer greater control over patient selection and data collection, a registry will not resolve all issues as shown by four Dutch registries for patients with cancer.[258] Population-based registries might however enable the selection of sufficient numbers of similarly treated patients and reduce issues with generalisability, missing information, and lack of standardisation in reporting in hospital records. Furthermore, registries can also be used to monitor and improve quality of care beyond outcomes research. If medical records are to be used

for data collection, however, it is important to emphasise that there is a high need to improve reporting of clinical data.

The survival of multiple myeloma has improved in the past decade in which new innovative drugs became available.[197,259] The question arises, however, why new innovative drugs need to be so expensive.[260,261] In the case of bortezomib, outcomes research was probably not the best option to reduce policymaker uncertainty regarding the initial question of whether the additional costs are worth public funding because the data on relative outcomes were invalid and thus could not resolve the questions about value for money from a societal perspective. It was however useful in generating real-world evidence on clinical outcomes and the costs of the patients receiving the drug in daily practice, which can be used to better manage the allocation of public funding within a particular disease. In contrast to the Dutch policy, the manufacturer and the UK department of health agreed, after first receiving a negative advice from the National Institute for Health and Care Excellence, on a performance-based response-rebate scheme for bortezomib.[262] Although outcomes research has been the only option in The Netherlands, the Dutch minister announced the implementation of risk-sharing arrangements from 2013 onwards.[168]

Several taxonomies exist that classify different risk-sharing arrangements,[18,34,235,243,263] but many issues related to various arrangements are known. For example, monitoring issues, administrative burden, and time-consuming procedures for filling claims unfortunately resulted in many missing claims in the UK bortezomib response-rebate scheme.[264] Also other studies[35,36,185,186,243] reported issues related to risk-sharing arrangements, such as high implementation and transaction costs, administrative burden, lack of transparency, challenges in measuring treatment effect, and a lack of appropriate data infrastructures. Accordingly, the first schemes in the United Kingdom included outcomes-based (response-rebate) schemes whereas in the later years most patient access schemes concerned finance-based agreements (e.g., dose-capping), which are easier to implement in practice. Recent Dutch experiences revealed insufficient real-world evidence to perform a reassessment after four years of data collection (i.e., omalizumab, infliximab, and ranibizumab); a few revisions have been converted into other risk-sharing agreements (i.e., a pay-for-performance[184] and a finance agreement[190]).

Nevertheless, at the time of the initial reimbursement decision, potential issues challenging outcomes research of bortezomib might have been in line with expectations regarding relapsed/refractory multiple myeloma treatment (i.e., small patient population, rapid advances in treatment). It also remains debatable, however, whether an outcomes-based risk-sharing agreement, such as in the United Kingdom, decreases policymaker

uncertainty regarding whether the high drug costs are worth public funding. Although such an agreement requires less data and great patient heterogeneity would not be an issue, other issues are likely to exist (e.g., validity of the outcome measure, monitoring issues, and administrative burden). Instead of requiring outcomes research in general, as in the Dutch case, policymakers could also consider requesting additional data on specific uncertain items –for example, prioritised by a Value of Information analysis– to enhance outcomes of investments of valuable resources. Policymakers could also consider requesting Bayesian updating of the existing model, which is trial-based, stochastic or model input parameter uncertainty analysis, or a synthesis of evidence from the real-world with trial follow-up and other published information.

There is currently, however, no flowchart available to policymakers that outlines the policy options available to best address the various types of uncertainty regarding value for money of a new health care technology under consideration for reimbursement. Future research might consider developing guidelines that assist policymakers in selecting the most appropriate arrangement addressing the type of uncertainty in question. Such guidelines should preferably provide a flowchart describing different options (e.g., conditional reimbursement, finance- or outcomes-based risk-sharing arrangements, and patient registry) and appropriate time frames while taking into account the type of uncertainty (e.g., medical or economic uncertainty), the type of disease (e.g., population size, acute vs. chronic), and characteristics of the drug.

To conclude, outcomes research provided valuable information on real-world patients, types of treatments, dosages, dose modifications and health care costs. Assessing (incremental) effectiveness and cost-effectiveness, however, was challenged by small patient numbers, missing data, extensive treatment variation, and great patient heterogeneity in everyday practice. Although the generated evidence improved informed decision making regarding the value of the drug in everyday practice, much uncertainty remained regarding its incremental cost-effectiveness. At reimbursement decision making, policymakers should carefully consider what type of disease-specific evidence could lead to an acceptable reduction in uncertainty regarding the question whether the (high) drug costs are worth public funding at its re-evaluation. Instead of implementing outcomes research requirements in general, policymakers should carefully consider which option (e.g., finance- or outcomes-based risk-sharing arrangement) will appropriately reduce uncertainty and ensure sufficient value for money and is worth the costs of implementation.

## Acknowledgements

The authors would like to thank all participating hospitals and involved hospital data-managers and cancer registration staff and are grateful to Daniëlle Ammerlaan, Saskia de Groot, Aat van Klooster, Annemieke Leunis, Arewika Mamakanjan, Zjwan Rashaan, Anneke Spreij, Harm van Tinteren, Bruno Vieyra, and Anneloes van Walsem for their contribution to this study.

This study was supported by the Dutch Health Care Insurance Board (CVZ).

# 11

# General discussion

## Background

Due to ever increasing health care expenditure, many countries developed policies and decision structures to contain expenditure. Although such policies can be directed at all (innovative) health care technologies in the basic benefit package, most of them have been directed at pharmaceutical products.[11] After demonstrating the efficacy, quality and safety of a pharmaceutical product to gain market approval from the market licensing authority (e.g., the European Medicines Agency, and the United States Food and Drug Administration), many countries require additional HTA evidence for reimbursement decision making. This additional requirement has been labelled as the 'fourth hurdle' to obtain market access.[12] There are differences in drug reimbursement systems across countries, and there has been a lack of evidence on the efficiency and sustainability of the diverse systems. Analytically oriented studies enhance understanding of drug reimbursement systems and their strengths and weaknesses,[13,14] and thus facilitate policy learning and provide lessons on how to improve system efficiency and sustainability.

This thesis focusses on decision making in drug reimbursement. The overall aim is to describe, analyse and evaluate systems that make decisions on reimbursement of drugs. The thesis is structured in three parts addressing six main research questions. As explorative introduction, part A investigates the social preferences for the goals of a health system. Part B focusses on drug reimbursement procedures, processes and criteria. Part C evaluates the coverage with evidence development policy tool to handle uncertainty of evidence in reimbursement decision making. This final section discusses the main findings, presents suggestions on how to improve decision making in drug reimbursement, and explores challenges for future research.

## What are the objectives of a health system?

Chapter two explored the theoretical foundation of an international framework for cross-country health system performance comparison and elicited social relative preferences of five independent health system objectives. Health related objectives were valued most importantly (equitable distribution of health [0.34] and average level of health [0.29]), followed by financial fairness (0.29), and process related objectives (utility derived from the process [0.07], and its distribution [0.06]).

At the national level, however, most health care systems share three main objectives for policymaking. The overarching objective is to improve or maintain health, but for policymaking within the constraints of limited resources and social preferences with re-

spect to equity and accessibility. The three policy objectives can be seen as the poles of a triangle (see Figure 11.1), where the aim of policymaking is to balance these objectives according to a socially acceptable equilibrium. Policymakers from different countries may differ in how they trade-off between these objectives, and hence, where they are situated within the triangle.

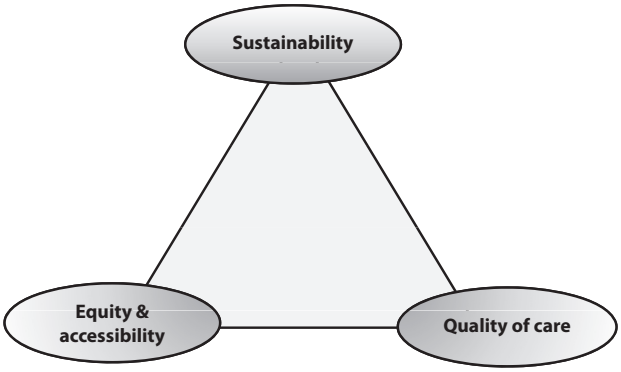


**Figure 11.1** Health care system policy objectives
Source: Le Polain, Franken et al.[15]

## What criteria are important in drug reimbursement decision making?

Chapter three and four showed that the five studied countries (Austria, Belgium, France, The Netherlands, and Sweden) use reasonably comparable criteria for decision making related to drug reimbursement. The criteria reflect the main policy objectives of the systems. Addressing the quality of care objective, therapeutic value was found to be, by far, the most important criterion in all five countries. The therapeutic value judgment is mainly based on an evaluation of the efficacy, effectiveness, safety, and adverse effects of a drug. In all five countries, added therapeutic value was reflected in a higher reimbursement basis.

Addressing the sustainability objective, cost-effectiveness and budget impact are formal criteria at the national or regional level in all five countries (cost-effectiveness only recently for new drugs in France). The assessment of cost-effectiveness, however, varies across countries from evaluating the quality of evidence to evaluating the actual cost-effectiveness ratio. None of the five countries applies a strictly defined or transparent cost-effectiveness threshold (range).

Finally, addressing the equity and accessibility objective, the severity of the disease is an important criterion in reimbursement decision making. The more severe a disease is, the higher the chance to get the treatment reimbursed (and at a higher reimbursement rate), especially in case of a rare disease without an alternative treatment. As such, the disease severity criterion is related to the medical necessity, therapeutic, and social need for a particular treatment.

## How do drug reimbursement decision-making systems handle uncertainty of evidence?

Policymakers often face the pressure ensuring timely access to promising drugs for patients in need. Uncertainty may exist on the actual clinical benefit, the adoption and diffusion rate, value for money, and the economic impact of a new drug.[18] The five studied countries vary in how they handle uncertainty of evidence in decision making. None of the five countries, however, seems to systematically postpone the initial reimbursement decision (i.e., delay access) until sufficient evidence becomes available,[16] nor systematically links the price paid for drug to the quality of the available evidence.[23]

Austria is the only country that has no system of systematic revisions, although ad hoc evaluations can be initiated. Belgium and The Netherlands have an evaluation procedure for specific groups of drugs (class I drugs in Belgium, and medical specialist drugs -*formerly expensive inpatient drugs*- in The Netherlands); an evaluation occurs within a window of one-and-a-half to three years in Belgium or four years in The Netherlands. In contrast, France systematically revises all reimbursed drugs every five years. Sweden evaluates all enlisted drugs from the old scheme (listed before 2002) according to therapeutic classes, and additionally, decides on a case-by-case basis whether a decision requires evaluation after a time period.

Many systems introduced a variety of policies to reduce uncertainty by making reimbursement conditional on additional evidence development.[18,24-26] Although such policies aim to guarantee early access to all patients in need (addressing the equity and accessibility objective), this is at the risk of failure towards the quality of care and sustainability objective (i.e., a wrong decision if the drug is not sufficiently (cost) effective or not sufficiently safe).

Part C evaluated the Dutch coverage with evidence development policy regarding its effectiveness, feasibility and appropriateness. This policy was first implemented in 2006 for expensive inpatient drugs, and from 2013 onwards, this policy has been extending

to specific groups of outpatient drugs. If a drug was included in this policy, hospitals temporarily received an additional ear-marked budget. Since the Dutch policy changed towards risk-based package management (2012), drugs included are eligible for conditional funding (i.e., hospitals negotiate with the insurance companies on the price paid for the add-on Diagnoses Treatment Combination). Conditional funding (previously temporarily additional funding) is, however, linked with the obligation to gather data on appropriate drug use and cost-effectiveness in everyday clinical practice.[229] After four years of use, an evaluation determines whether or not (additional) funding continues to exist.

Chapter eight assessed whether the Dutch coverage with evidence development policy was effective in reaching its objective (guaranteeing equal access). Although physicians needed a long adjustment period to get familiarised with the new drug, they did not indicate experiencing prescription barriers. The use of the drug significantly increased after the introduction of the policy. Some unexplained regional differences, however, continued to exist. This indicates that the policy had a positive effect on prescribing behaviour and was thus, at least partially, effective in reaching its objective.

Coverage with evidence development seems to be a promising policy option to reduce policymaker uncertainty by guaranteeing early access to all patients in need while ensuring the possibility to revise the initial decision, which was based on (highly) uncertain evidence on the drug's real world value. It is, however, debatable whether it is always feasible to obtain sufficient evidence from everyday clinical practice for an evaluation relating to whether the high drug costs are worth public funding. Chapter nine showed that this may depend on the drug and the disease. The three case studies confirmed theoretical expectations and expert opinions[219,224] that the lack of randomisation in everyday practice result in incomparable patient groups and the inability to correct for confounding. However, the chapter also revealed that it still may be feasible to obtain internally valid and generisable incremental cost-effectiveness estimates by synthesising everyday practice data with trial data (in the oxaliplatin cases). Moreover, we showed that the gathered data from everyday practice resulted in valuable evidence for policymakers regarding various initial uncertainties arising from the gap between clinical trials and everyday practice.

In all three case studies it was possible to reduce uncertainty regarding the drug's use, effects and costs in everyday practice patients. Such evidence can improve quality of care and enhance efficient allocation of scarce health care resources. However, we also showed that for some drugs and indications it may be impossible to estimate sufficiently valid and precise incremental cost-effectiveness estimates (in the bortezomib case). It is

highly debatable whether a coverage with evidence development policy is always the best option to reduce policymaker uncertainty. Chapter ten showed that other options such as finance or outcomes based agreements may be more appropriate for some drugs or indications.

## To what extent do the drug reimbursement systems satisfy the conditions of legitimate decision making?

Policymakers are forced to make trade-offs between the objectives of their health care system. These trade-offs are a matter of normative choice; the aim of policymaking is to balance the system objectives according to a socially acceptable equilibrium. Therefore, drug reimbursement systems cannot be compared on the outcomes of the trade-offs, but they can only be assessed according to what extent the decision-making process satisfies the conditions of legitimate decision making in line with social values of a country. We used the theoretical ethical accountability for reasonableness framework developed by Daniels and Sabin[102] to operationalise legitimacy of decision making. According to the Daniels and Sabin framework, a fair and legitimate procedure satisfies four conditions: (i) publicity of the decision rationale, (ii) relevance of the decision criteria, (iii) revisability of the decision in light of new evidence and arguments, and iv) enforcement of these three criteria.[102]

### Publicity of the decision rationale

Although all five systems use reasonably similar reimbursement criteria, the actual role of each criterion in the decision-making process is often not transparent. Assessment reports are usually published (except in Austria). The weighing process (i.e., appraisal) which leads to the advice (or decision), is, rarely published, however. Similarly, assessment criteria are clearly defined whereas appraisal criteria are often far less explicit. Although interviewees in our study acknowledged that therapeutic value is the most important criterion, none of the countries applies a formal hierarchy of the criteria and the actual role of each criterion relatively to other criteria is not transparent. None of the countries applies an explicit cost-effectiveness threshold (range). The role of appraisal criteria particularly lacks the publicity of the decision rationale. Chapter five showed, for example, that policymakers experience substantive difficulties in operationalising the disease severity concept and, subsequently, seem to struggle in making its actual role in the decision making process explicit. Remarkably, this chapter also revealed that policymakers, above all, prefer maintaining discretionary decision power and thus prefer continuing implicitly weighing all decision criteria.

In Sweden, applicants can withdraw their reimbursement request (e.g., if leaning towards a decision to reject reimbursement), in which case no report is published. This practice guarantees confidentially, but at the cost of transparency. Similarly, price-volume agreements in France, and recently also in The Netherlands, aim to control the budgetary impact, but because of confidentiality, they decrease the transparency of the decision rationale. Remarkably, even though The Netherlands has a separate appraisal committee, reimbursement decision rationales on outpatient drugs are hardly ever discussed at appraisal committee meetings.

**Relevance of the decision criteria**

This condition requires that all involved stakeholders understand the decision problem and acknowledge that choices need to be made to meet the system objectives. Involvement of all those who are affected by the decision increases the likelihood that the rationales for the decision will be considered relevant and acceptable.[107,265] The system objectives are reflected by the main reimbursement criteria (the therapeutic value, cost-effectiveness and budgetary impact, and disease severity criterion reflect the quality of care, sustainability, and accessibility and equity objective, respectively). This does not, however, guarantee that the general public truly understands the decision rationale. All five systems satisfactorily seem to ensure stakeholder involvement, either through direct representation in the expert advisory committee (Belgium and Austria), or through consultation of relevant stakeholders in case the committee consists of scientific experts (France, The Netherlands, and Sweden). It should be noted that only Sweden has a patient representative in the expert committee; this committee also makes the final decision.

**Revisability of the decision in light of new evidence and arguments**

Especially in case of high uncertainty and/or rapid advances in treatment and, subsequently, changes in clinical guidelines, the option of revising a decision is important whenever new evidence becomes available. Although variation exists, the five systems make limited use of policies that enforce systematic collection of new evidence and/or policies to systematically evaluate previous decisions. Only France and Sweden systematically conduct group revisions. In addition, France revises the status of all drugs every five years, and Sweden revises a decision, case-by-case, in case of important initial uncertainty. Belgium and The Netherlands only evaluate decisions for specific groups of drugs. Conversely, Austria has no system of systematic revisions.

Remarkably, however, our study revealed that, although the Dutch coverage with evidence development policy addresses the revisability condition, Dutch policymakers experienced great difficulties in enforcing the consequences of the first revisions for

expensive inpatient drugs. If reimbursement is to be made conditional on evidence development, the consequences of evaluating the new evidence (e.g., delisting, altering the price) should be enforceable. In contrast, results from systematic revisions in France[93,95] and Sweden[96] are promising.

### Enforcement of these three criteria

All five countries legally instituted a designated national reimbursement agency that falls under ministerial responsibility. All systems satisfactorily implemented formal appeal procedures for stakeholders. Although the agencies are audited or certified by external committees, self-evaluation is lacking in relation to performance of the system on the processes and outcomes. The outcome of the system is mainly monitored on pharmaceutical expenditure (addressing the sustainability objective). The systems rarely conduct an evaluation regarding the quality of care, and equity and accessibility objectives.

### Limitations of the Daniels and Sabin framework

One of the main limitations of this thesis may be the use of the Daniels and Sabin framework. Although this framework has emerged as a leading framework for fair priority setting,[266] others have criticized it. The framework focuses on procedures that ensure fairness and legitimacy of a just decision; the generated outcomes are not important, however, and thus, an important gap remains at the content level.[267] Friedman[268] suggested revising the legitimacy conditions because they are not sufficiently adequate for ensuring decisions that are acceptable to everyone. Moreover, Schokkaert[269] claimed that the framework does not contribute anything to the substantive debate how to make difficult choices in priority setting. Nevertheless, empirical evidence suggests that priority setting processes that fulfil the four conditions for accountability for reasonableness are perceived as being legitimate and fair.[106-108] We believe that using the Daniels and Sabin framework has provided us in-depth insights into reimbursement decision making, and therefore, has been extremely useful as a first step towards identifying strengths of and challenges for the systems.

## What are the strengths of and challenges for drug reimbursement systems?

Our study provided important insight into the strengths of and challenges for drug reimbursement decision-making systems. Nevertheless, it should be noted that the relative small sample of countries can be seen as one of the main limitations. The five countries may not necessarily be representative of other European countries. We observed

similarities as well as important differences in structure, organisation, and procedures of the systems, however. Because of the degree of detail of our analysis, we believe that this was only feasible in a relatively small number of countries. We also believe that only little benefit would have been gained if we had included more countries. Another limitation was that we only studied the coverage with evidence development tool regarding three indications in cancer using retrospective research designs in The Netherlands. Some of our conclusions may therefore not be generalisable to all drugs in different contexts. Nevertheless, we believe that our research enhanced understanding of drug reimbursement decision-making procedures and processes. Especially the information regarding challenges for the systems facilitates policy learning on how to improve the decision-making process.

**Strengths of the systems**

All five countries have a national system that evaluates whether or not a drug should be reimbursed out of public revenue. All share similar and clear policy objectives which are reflected by the main criteria for reimbursement. The therapeutic value criterion mainly reflects the quality of care objective, the cost-effectiveness criterion mainly reflects the sustainability objective, and the severity of disease criterion mainly reflects the equity and accessibility objective. Performance of the systems is monitored in terms of pharmaceutical expenditure (addressing the sustainability objective). Information based on health technology assessment is used at several phases in the decision-making process in order to trade-off between the system objectives.

All systems are prepared to pay out of public pocket for drugs with sufficient added therapeutic value. Stakeholders are involved in the process either through consultation or direct representation in the expert advisory committee. Implementation of reimbursement decisions is facilitated by means of guidelines and drug formularies.

**Challenges for the systems**

This thesis has put forward three main challenges: (i) transparency of the decision-making process, (ii), cyclical decision making and (iii) decision making based on value for money.

*Transparency of the decision-making process*

Transparency of the decision-making process is the first condition of legitimate decision making (publicity of the decision rationale). Theoretically, assessment and appraisal are two separate phases in the reimbursement decision-making process (see Figure 1.1). Assessment is purely descriptive in terms of evaluating the quality of and quantifying the available evidence, while appraisal entails weighing up the outcomes and thus con-

tains a value judgment. Chapter four revealed, however, that assessment and appraisal are often strongly intertwined processes in actual decision-making practice. None of the countries applies a formal hierarchy of the reimbursement criteria, nor applies an explicit cost-effectiveness threshold; the actual role of each criterion in the decision-making process lacks transparency. Especially the appraisal (i.e., appraisal criteria and the weighing process) is not sufficiently transparent.

*Cyclical decision making*

A cyclic decision-making process entails four phases (see Figure 1.1). After the assessment, appraisal, and decision-making phase, follows the evaluation phase. The evaluation phase entails (i) ascertaining the consequences of the decision, and (ii) evaluating (restart the cycle of assessment, appraisal, and decision making) whether the initial decision is still appropriate in case new evidence has become available.

Regarding ascertaining the consequences, all five systems monitor the performance of the systems in terms of pharmaceutical expenditure (addressing the sustainability objective). None of the systems, however, systematically evaluates the impact of the system against the other two system objectives (i.e., the quality of care, and the equity and accessibility objective).

Regarding evaluating whether the initial decision is still appropriate, we could demonstrate that the systems only make limited use of policy tools to systematically evaluate drugs´ relative value throughout their life cycle, although variation exists. All five systems have a so-called supply driven system: the process starts with a manufacturer´s request for reimbursement and proceeds on a case-by-case basis. This may, however, lead to pragmatic incrementalism,[91] risking a low degree of consistency across decisions. Not all systems are sufficiently equipped to systematically deal with uncertainty of evidence in the decision making. Consequently, the systems seem to experience substantial difficulties to fulfil the revisability condition of legitimate decision making.

*Decision making based on value for money*

Budgetary impact and drug expenditures are monitored in all five systems, this does not, however, guarantee value for money in the decision making. It should be noted that all five countries have an 'open-ended' pharmaceutical budget (i.e., there is no absolute budget maximum or maximum relative share of health care expenditure). In contrast to budgetary impact, information on cost-effectiveness can be used to base decisions on value for money. Cost-effectiveness is a formal reimbursement criterion in all five countries (only recently for new drugs in France). However, none of the countries applies a strictly defined or transparent cost-effectiveness threshold (range). Moreover, none of

the countries explicates the importance of cost-effectiveness relatively to other criteria. This is even true for Sweden, which is the only one of the five studied countries that has a value-based pricing system. Value-based pricing implies a direct link between the gains and the costs of a pharmaceutical product.

Remarkably, cost-effectiveness does not seem to play a major role in actual decision-making practice. Chapter six showed that, in the period January 2005 to July 2011, reimbursement dossiers for only eleven drugs included a full economic evaluation (i.e., cost-effectiveness and/or cost-utility analysis[270]) in both Sweden and The Netherlands. In The Netherlands, only 12% of all outpatient drug applications (35% of List 1B applications -non-interchangeable drugs and/or drugs with an added therapeutic value-) included pharmacoeconomic evidence. Although more pharmacoeconomic evidence was available in Sweden (97%), only 33% of dossiers included a full economic evaluation. Furthermore, chapter seven revealed that Dutch policymakers experienced great difficulties in putting restrictions on reimbursement based on cost-effectiveness, even in case of extreme high costs per Quality Adjusted Life Year.

## Suggestions to improve decision making in drug reimbursement

This thesis enhanced the understanding of drug reimbursement systems by obtaining insights into reimbursement procedures, processes, criteria, and insights into the effectiveness, feasibility, and applicability of the coverage with evidence development policy tools to handle uncertainty of evidence in the decision making process. Therefore, our study contributed to the scientific evidence base for improving reimbursement systems' legitimacy, efficiency and sustainability. The strengths of, and especially the identified challenges for the systems can be seen as opportunities for furthering the aim to improve current drug reimbursement decision-making practice.

The main overarching challenge is improving the legitimacy of the decision-making process. It should be noted, however, that even if decisions are made following a legitimate process, this does not guarantee that the outcomes are always perceived to be acceptable to the general public or politicians. As described above, according to the Daniels and Sabin framework,[102] a fair and legitimate procedure satisfies four conditions. Each of the following recommendations addresses, in no specific order, one of the legitimacy conditions. The list of recommendations is not exhaustive, but restricted to recommendations that arose from the main study findings.

### Building on existing strengths

In all five systems reimbursement decisions are made at the national level. A dedicated national system fulfils the fourth condition of legitimate decision making, because it can enforce the first three conditions. Theoretically, a national decision system ensures equitable allocation of health care resources. The main system objectives are reflected by the existing reimbursement criteria. Therefore, these criteria sufficiently facilitate making trade-offs between the system objectives even within the constraints of limited resources and social preferences with respect to equity and accessibility. Thus in principle, the existing systems are equipped for their imposed responsibilities. However, if budget responsibility is allocated, for example, to regions (as in Sweden) or to hospitals (as in The Netherlands), equitable access may not be guaranteed for patients in need, and variation in access may occur within a country. It is, therefore, important to closely monitor the outcomes of a system regarding all three policy objectives.

### Increasing the transparency of the decision-making process

One of the main challenges for the systems is improving the transparency of the decision-making process (addressing the publicity condition of legitimate decision making). Especially the use of appraisal criteria and their role in the decision-making process lacks transparency. Because assessment and appraisal are often strongly intertwined processes, the systems can better disentangle assessment and appraisal. Surprisingly, even in The Netherlands, which has a separate appraisal committee, assessment and appraisal are strongly intertwined processes. Moreover, outpatient drugs are hardly ever discussed in Dutch appraisal committee meetings. If an appraisal committee is established next to an assessment committee, reimbursement applications and evaluations are likely to be discussed at both committee meetings; especially in the case of relatively high costs and limited health gains.

Conversely, chapter five revealed that policymakers prefer maintaining discretionary decision power. This implies a great challenge for policymakers if the decision-making process is to be made more transparent. It is debatable why policymakers experience such difficulties in explicating their decision base. It could be that it is easier to make 'happy' decisions (i.e., grant reimbursement) that satisfy all stakeholders compared to making 'unhappy' decisions (i.e., denying reimbursement). Klein[271] stated, for instance, *"Ministers will always be tempted to take credit for generosity, even while seeking to avoid blame for parsimony."* However, priority setting choices need to be made in order to keep the system sustainable in the long-term. It is, therefore, of utmost importance to improve the transparency of the decision-making process.

None of the five countries uses an explicit decision framework. In chapter four, we present an explicit framework that can be used as a first step to improve the legitimacy of the decision-making process. The framework addresses both the transparency and the relevance condition of legitimate decision making. It consists of five key questions including their relevant criteria, relating to (i) the medical, therapeutic, and/or societal need, (ii) preparedness to pay for treating a particular condition, (iii) preparedness to pay for a particular treatment, (iv) preparedness to pay more compared with its alternatives, and (v) actual willingness to pay from public sources. The framework brings a certain logic in the order of the questions that need to be answered in reimbursement decision making; it allows the reconstruction of the decision-making process. The framework can be seen as a tool to allow better structuring of the decision-making process; all five questions need to be answered subsequently. Therefore, it can facilitate consistency of decision making and support the justification of decisions and, by clearly reporting the answer to each question, it can ensure the transparency of the decision-making process.

**Make the reimbursement process more cyclic**

Reimbursement decisions are inherently made under uncertain conditions. It is, therefore, crucial to ensure a cyclic decision-making process and thus evaluate a drug's social value throughout its entire life cycle. An initial reimbursement decision ought to be revisable in case of new arguments or in case that new evidence becomes available (addressing the revisability condition of legitimate decision making).

This thesis showed that the five systems make limited use of systematic revisions and that not all systems are sufficiently equipped to systematically deal with uncertainty of evidence in the decision making. Nevertheless, in recent years an increasing number of risk-sharing schemes and conditional reimbursement policies have been introduced and seem to gain more attention. Results from systematic revisions in France[93,95] and Sweden[96] are promising.

There is, however, much room for improvement. If policymakers are prioritising early access to promising drugs with high uncertainty on their actual value in everyday practice, it is even more important that evaluation options exist. Moreover, evaluation policies need to be embedded in the system (addressing the enforcement condition of legitimate decision making). Furthermore, Part C showed that it is important to evaluate policies regarding their effectiveness, feasibility and appropriateness. After a new policy has been instituted, the policy needs to be evaluated as to whether it fulfils its objectives and whether or not it requires adjustments. For instance, the Dutch coverage with evidence development policy was partly triggered by signs of 'postal code prescribing' of an expensive drug for breast cancer.[192] Chapter eight revealed, however, that although

the use of the drug significantly increased after the introduction of the policy, some unexplained regional differences remained. Thus the policy was only partially effective in reaching its objective (equitable access to expensive drugs); adjustment of the policy may still be necessary.

Coverage with evidence development schemes seem to be increasingly popular as policy instrument to handle uncertainty of evidence, and thus to ensure a cyclic process. It is, however, debatable whether it is always feasible to obtain sufficient evidence from everyday clinical practice for an evaluation regarding whether the high drug costs are worth public funding. This may depend on the type of drug and type of disease; as chapter nine showed, for some drugs and indications it may be impossible to obtain sufficiently valid and precise incremental (cost) effectiveness estimates using real-world data. Therefore, it is important to carefully consider what type of policy appropriately ensures sufficient reduction of the initial uncertainty of the evidence at the time of evaluation for the particular drug in question.

Furthermore, if the reimbursement process is to be made more cyclic, the consequences of an evaluation (e.g., delisting, altering the price) ought to be transparent at the initial decision and enforceable in the evaluation phase. New policies may require different approaches using other tools, and they may induce important new responsibilities for policymakers. For example, we revealed that Dutch policymakers experienced great difficulties in enforcing the consequences of the first evaluations for expensive inpatient drugs. However, guaranteeing early access without implementing and, if required, executing exit strategies is not sustainable in the long term. In some cases, it may even be more appropriate to reject (conditional) reimbursement at the initial phase. It should be noted that it requires political courage to revise a decision and delist a drug that was previously reimbursed; this may even be more difficult in case when media attention is present and/or if there is publicity surrounding identifiable individuals who may suffer from delisting.

**Encourage reimbursement of drugs that deliver social value for money**

All five systems can improve decision making based on value for money. All require evidence on cost-effectiveness; there is, however, much room for increasing its actual role in the decision making. Chapter six and seven revealed that cost-effectiveness may not be that important in actual decision making. It is debatable if this is sustainable in the long term, especially because new innovative drugs are inclined to increase exponentially in price. For instance, between 2003 to 2013, expenditure for cancer drugs increased in The Netherlands from 270 to 733 million Euros per year.[272] Furthermore, there seems to be a rising trend of using combination(s) of expensive treatments for the same indication.

The question arises why new innovative drugs need to be so expensive,[260,261] and how the systems can encourage appropriate price setting for drugs to ensure social value for money.

Above all, if reimbursement decision making is to be based on value for money, cost-effectiveness ought to play an important role in decision-making practice. The systems could explicate the actual role of cost-effectiveness relatively to other reimbursement criteria. Besides cost-effectiveness, other criteria may complement sustainable decision making. For instance, the Dutch Commission for the Assessment of Oncological Resources (Commission BOM) suggested a minimal expected gain of two months in life expectancy.[191] This links closely to pleas for only reimbursing interventions that, at least, produce a non-negligible health gain.[154] Using both an explicit cost-effectiveness threshold range and an explicit minimal health gain facilitates transparency of the decision making and thus may help to gain public trust in the system.

Although it is often claimed that an explicit cost-effectiveness threshold range may induce strategic behaviour, such behaviour may also be observed in case of an implicit threshold. For instance, it seems that the incremental cost-effectiveness ratio of outpatient drugs that obtained reimbursement in the last few years in The Netherlands continue to rise far above the reimbursement agency's suggested threshold maximum (€80,000 per Quality Adjusted Life Year[120]). Remarkably, several of these decisions concerned outpatient drugs that were not even discussed at Dutch appraisal committee meetings. Conversely, in England it seems generally acceptable that the National Institute for Health and Care Excellence (NICE) uses cost-effectiveness as reimbursement criterion and applies an explicit cost-effectiveness threshold range in decision making.[273] This shows that decision making can be transparently based on cost-effectiveness. Although the threshold may incentivise to set a price just below the threshold, it may, on the other hand, also incentivise to keep the price below this threshold (instead of continually rising prices to implicitly seek for a maximum as seems to be the case in The Netherlands). Nevertheless, it should be noted that political unacceptability of some of NICE decisions to reject reimbursement has led to arrangements such as patient access schemes,[274] the Cancer Drugs Fund,[135] as well as to changes to NICE's decision framework (e.g., the End of Life premium[275]).

It is important, however, to determine who is responsible for setting the threshold range, and who is responsible for executing the threshold. For example, in 2012 the Dutch agency considered a negative advice for three specialist orphan drugs for Pompe and Fabry disease. Preliminary reports were leaked to the press, which resulted in a turbulent episode. After a considerable amount of public debate and political pressure,

the Dutch agency modified its advice, recommending the continuation of funding because of the high severity of the disease, the high costs per individual patient, and the relatively low budgetary impact.[188] Remarkably, the Dutch agency did not have any new scientific evidence. The minister accepted the advice and in 2013 made a confidential price agreement with the manufacturer.[190] It is, however, highly questionable whether it was feasible to agree on a price that ensures social value for money (e.g., Myozyme for classic Pompe disease was estimated to cost €300,000 to €900,000 per Quality Adjusted Life Year[188]).

Although the Dutch agency previously suggested a threshold range depending on the severity of the disease (i.e., €10,000–€80,000 per Quality Adjusted Life Year[120]), this was never confirmed nor endorsed by the minister. It is, however, a question of debate whether the Dutch reimbursement agency, only having an advisory role, should take the lead in determining the Dutch threshold. This is especially as they not only will carry all the negative publicity, but may also be subsequently overruled by the Minister. Moreover, although medical guidelines should consider cost-effectiveness and physicians should be involved in the rationing debate, it is highly questionable whether rationing decisions should be entirely delegated to the individual physician who actually treat the patient in need. In June 2014, the Dutch Cancer Society (KWF Kankerbestrijding) published a monitoring report acknowledging that there ought to be limits to the costs for a treatment, and advising the minister to set up a committee to propose a cost-effectiveness threshold range.[272] Interestingly, the social acceptability of denying treatments seems to have changed even in the last few years. For example, medical professionals were reluctant to base treatment decisions on cost-effectiveness in the discussion regarding Pompe and Fabry disease, while currently they specifically ask for transparency regarding a cost-effectiveness threshold (range) imposed by the government.

It is of utmost importance that the rationale for priority setting comes into the public debate. In 2013, the Dutch reimbursement agency organised focus group discussions with patients and the general public regarding this debate. They concluded that there was no sense of urgency for such a public debate; this was due to a lack of knowledge of the focus group participants, and because there was a general feeling of denying 'the right to health care' for which the average citizen pays a high price.[276] Based on this, the Dutch reimbursement agency decided to not (yet) begin a wider social debate on cost-effectiveness. We believe, however, that medical professionals as well as the general public need to be involved in the debate as to how much society is willing to pay for a treatment. Consequently, it is essential to appropriately inform the general public to overcome the current lack of sense of urgency and the associated gap in knowledge.

Although politicians seem to be unwilling to set an explicit (maximum) threshold, they have to take their responsibility, as solidarity in health care may be at risk. Decision making based on value for money may induce new tasks and responsibilities for policymakers and may require a different approach taking into account the interests of all involved stakeholders.

Furthermore, a cost-effectiveness threshold may also act as gatekeeper for conditional reimbursement. In the case of extremely high costs, the drug will, even in case of optimistically estimated health gains, inevitably fail to meet the maximum of any realistic threshold. Similarly, in such cases it will most likely be impossible to agree on a price reduction (e.g., by means of a financial risk sharing agreement) which ensures social value for money. Policymakers could make better use of this knowledge at the initial decision and thus reject (conditional) reimbursement at the initial decision instead of postponing the difficult decision, for example, by means of a coverage with evidence development scheme. Policymakers could also consider linking the reimbursed price to the quality of the available evidence to ensure value for money.[23] Remarkably, only in the Swedish value based pricing system, the reimbursement and price decision is one combined decision that is made by one committee. Most systems have no, or only a relatively loose link between the price and reimbursement decision. If value based pricing is the future, drugs should be priced according to their actual value and thus a closer link should be established between the price decision and the reimbursement decision.

Regarding risk sharing schemes, it is important that policymakers are aware of the pros and cons of such schemes. Although financial risk sharing agreements divide the surplus between payers and manufacturers and results from the price-volume agreements in France are promising,[93,94] such agreements do not reflect value for money. Moreover, such agreements are most often confidential in order to maintain a high list price. This is, however, at the cost of transparency. In contrast, agreements based on outcome guarantee value for money (i.e., no cure no pay agreement). However, many issues have been reported for this sort of agreement including high implementation and transaction costs, administrative burden, and challenges in accurately measuring treatment effect.[35,36,185,186] Although coverage with evidence development schemes require an evaluation after a certain time period, and thus may ensure value for money in the long term, it is important to mention that during the first years the payer takes all the financial risks. Moreover, such schemes often heavily rely on real-world observational data. As chapter ten showed, observational data may not always sufficiently reduce the uncertainty regarding a drug's value for money.

Consequently, policymakers need a variaty of tools that address the various types of uncertainty. Depending on the type of drug and type of disease, policymakers should carefully consider, case-by-case, which option is most appropriate and will sufficiently reduce the initial uncertainty and also ensure value for money in the long term and is worth its costs of implementation.

## Challenges for future research

This thesis contributed to the scientific evidence base on reimbursement procedures, processes, criteria, and provided insights into the effectiveness, feasibility, and applicability of the coverage with evidence development policy tool. There are, however, important challenges for future research.

Our study revealed that none of the five countries studied used quantitative rating methods to support decision making. Similarly, Noorani et al.[277] found that HTA agencies in health care priority setting seldomly use quantitative rating methods. Although we acknowledge that quantitative methods should not impose decisions and some discretionary decision power should be maintained, we also believe that quantitative methods may be used as guidance to support decision making and may increase the consistency and transparency of the decision-making process. Future research could aim to develop a quantitative multi criteria decision analysis (MCDA) framework for policymakers that facilitates the incorporation of assessment and appraisal considerations (including value for money) in drug reimbursement decision making. Such a framework could be used by policymakers of the respective reimbursement agencies, expert advisory committees, appraisal committees, and ministries of health to improve the transparency of decision making on the basis of value for money. MCDA aims to facilitate decision making by applying a set of methods and approaches to explicate the impact on the decision of all applied criteria and the relative importance attached to them.[278-280] The MCDA method has been previously tested in drug reimbursement decision making (the EVIDEM framework). It was found to be extremely useful by the Canadian drug advisory committee as it supported a consistent approach and a systematic consideration of a broad range of appraisal criteria.[281,282]

Furthermore, systems increasingly use different policy tools to make the process more cyclic in order to deal with uncertainty of evidence and to improve value for money in the long term. There is, however, little evidence on the actual impact of the various tools and there is no guideline available that outlines how best to make the process cyclic and which policy tools best address which type of uncertainty. Future research could

aim to investigate in detail the effectiveness, and the relative strengths and weaknesses of these different tools and assess their actual impact on transparency in the decision-making process, and evaluate to what extent they address decision making based on value for money. A further aim could be the development of guidelines that assist policymakers in (i) selecting the most appropriate tool addressing the type of uncertainty in question, and (ii) determining the evaluation procedure. Moreover, the best case would be to incorporate these guidelines within the MCDA decision framework.

It should be noted, however, that although MCDA facilitates decision making, it does not provide a solution for all the current issues. For instance, the experienced dilemma of making difficult rationing choices (i.e., denying access to treatments) remains. It is therefore important that the rationale for priority setting is open for public debate. Policymakers need to take their responsibility and the general public needs to be involved in this difficult debate. Nevertheless, we believe that it is important to continue developing (current) methods that support decision making in actual practice in order to further advance evidence based decision making. The use of a comprehensive MCDA framework may increase the consistency and the transparency of the four phases of the decision-making process, and thus will improve legitimacy of social decision making on drug reimbursement. This may result in greater confidence in the system of the general public. It may also facilitate a closer link between the way HTA research evaluates technologies and the way governments allocate health care budgets. It is therefore essential that HTA researchers and policymakers work closely together, and make use of each other's expertise.

# Summary

## Introduction

In many countries life expectancy is increasing but, simultaneously, continuously rising health care expenditure has increased the pressure on the sustainability of the systems. To contain expenditure, many countries developed policies and decision structures to manage the basic benefit package; most of them have been directed at pharmaceutical products. After demonstrating the efficacy, quality and safety of a pharmaceutical product to gain market approval from the market licensing authority, many countries require additional Health Technology Assessment (HTA) evidence for reimbursement decision making (also called the 'fourth hurdle' to obtain market access). Due to differences in health care organisation across countries there are differences in drug reimbursement systems and reimbursement policies.

This thesis focuses on decision making in drug reimbursement. The overall aim was to describe, analyse and evaluate systems that make decisions on reimbursement of drugs. The thesis consists of three parts. Part A explores the preferences for the objectives of a health system. Part B investigates drug reimbursement procedures, processes and criteria in five European countries. Part C evaluates the coverage with evidence development policy tool to handle uncertainty of evidence in reimbursement decision making.

## Part A: Social preferences for health system objectives

In **chapter two**, we identify a comprehensive international framework facilitating cross-country comparison of health system performance, and provide its theoretical foundation. The framework consists of five independent health system objectives. Besides health itself, we included the process of health care delivery in a utility framework supporting that the process of care giving is an independent objective, irrespectively whether it affects health. We used a discrete choice experiment to elicit relative preferences for the five system objectives. Health related objectives were valued most importantly (equitable distribution of health [0.34] and average level of health [0.29]), followed by financial fairness [0.24], and process related objectives (utility derived from the process [0.07], and its distribution [0.06]). Compared to previous research, our weights, elicited from behind a 'veil of ignorance', placed much greater emphasis on health and health inequality than on process outcomes.

## Part B: Drug reimbursement procedures, processes and criteria

To obtain insight into drug reimbursement decision making, **chapter three** provides a comparative analysis of five European systems (Austria, Belgium, France, The Netherlands, and Sweden). The five systems share three main objectives for policymaking: (i) equitable access, (ii) quality of care, and (iii) sustainability of the system. The aim of policymaking is to balance these objectives according to a socially acceptable equilibrium. In the five systems, a national agency evaluates, on a case-by-case basis, whether or not a drug is worth public funding. The systems use comparable reimbursement criteria, which reflect the main systems' objectives. All are prepared to pay for drugs with sufficient added value. System impact, however, is mainly assessed by drug expenditure. The minister has discretionary decision power to alter the reimbursement advice in three of the five countries (Belgium, France and The Netherlands). The five systems make limited use of policies to systematically evaluate a drug's relative value for money throughout its life cycle.

Although the five systems make efforts to increase transparency in the decision-making process, none of the systems uses formal hierarchical reimbursement criteria nor applies an explicit cost-effectiveness threshold value. All five could improve the transparency of the decision-making process; especially appraisal lacks transparency. Based on these findings, we developed an explicit framework that can be used as a first step to improve the legitimacy of the decision-making process. Our framework, presented in **chapter four,** consists of five key questions including their relevant criteria, relating to (i) the medical, therapeutic, and/or societal need, (ii) preparedness to pay for treating a particular condition,( iii) preparedness to pay for a particular treatment, (iv) preparedness to pay more compared with its alternatives, and (v) actual willingness to pay from public sources. The framework can be seen as a tool to better structure the decision-making process; all five questions need to be answered subsequently. Systematic use of this framework in actual reimbursement decision making can facilitate consistency of decision making and support the justification of decisions and, by clearly reporting the answer to each question, it can ensure the transparency of the decision-making process.

Our study continues with exploring the importance of two reimbursement criteria (the disease severity and cost-effectiveness criterion) in actual decision making. **Chapter five** shows that the severity of the disease is an important consideration in Belgian, Dutch, French, and Swedish decision making; especially in case of high severity. All four countries could, however, improve the transparency of its actual importance relatively to the other criteria in the decision making. The operationalisation differs across the four countries. The Netherlands operationalised disease severity using the propor-

tional shortfall approach. Sweden uses categories to give an indication of the severity. However, in both countries disease severity only plays an implicit role in the decision whether to reimburse a drug. In Belgium and France, it additionally plays an explicit role in determining the willingness to pay out of public sources because both countries use disease severity in setting reimbursement levels. Our study revealed that none of the interviewees considered quantitative information on the severity of the disease, additional to qualitative information, to be of decisive importance.

Furthermore, we found the relative importance of cost-effectiveness far more modest than would generally be expected, especially in The Netherlands. **Chapter six** provides a comparative analysis of the role of pharmacoeconomic evidence in actual Dutch and Swedish reimbursement decision making. More economic evaluations were available in Sweden than in The Netherlands (97 vs. 31, respectively), mainly due to the numerous exemptions from pharmacoeconomic evidence in The Netherlands (65%). In The Netherlands, only 35% of the 118 applications on List 1B (i.e., claiming added therapeutic value) were found to include pharmacoeconomic evidence in the period January 2005 to July 2011. In all cases where drugs were rejected for reimbursement, the pharmacoeconomic evidence had been judged insufficiently robust. However, pharmacoeconomic evidence also had been judged insufficiently robust in 21% of the cases where drugs obtained reimbursement. In Sweden, drugs that were rejected for reimbursement had been judged either not cost effective (74%) or not supported by sufficiently credible data (26%). Nearly all drugs that obtained reimbursement in Sweden had been judged cost effective (92%). However, 53% of these judgements were based on a price comparison and 10% on a cost-minimisation analysis; only 33% were based on a full economic evaluation. Moreover, dossiers for only 11 drugs included a full economic evaluation in both countries; of these, the reimbursement decisions differed for four drugs.

**Chapter seven** specifically focuses on the actual impact of cost-effectiveness in Dutch decision making. Although HTA already informed Dutch policymaking in the early 1980s, evidence of health economic evaluations is only systematically used in drug reimbursement decision making. Even in drug reimbursement, however, the availability of evidence of health economic evaluations remains rather low. The chapter reveals that it is highly questionable whether health economic evaluations currently play a role in actual Dutch reimbursement decision making. Although the requirements exist in policy procedures, recent cases showed that Dutch policymakers experience great difficulties in putting restrictions on reimbursement based on evidence from health economic evaluations, even in case of extreme high costs per Quality Adjusted Life Year.

## Part C: Handling uncertainty in drug reimbursement decision making

Reimbursement decisions are inherently made under uncertain conditions. The Dutch coverage with evidence development policy addresses these uncertainties by making the decision conditional on additional evidence development. This policy was first implemented in 2006 for expensive inpatient drugs, and from 2013 onwards, this policy has been extending to specific groups of outpatient drugs.

**Chapter eight** assesses whether the Dutch policy was effective in reaching its objective (guaranteeing equal access of expensive drugs). Although interviews revealed awareness of the high treatment costs, physicians did not express experiencing prescription barriers. Nevertheless, our study showed that physicians needed a long adjustment period to get familiarised with the new drug bortezomib. The use of bortezomib significantly increased after the introduction of the policy (prescription rates increased from treating 2% of patients in 2004 to 17% in 2009), but remained below the rate estimated by the professional association of haematologists (27%). We found regional differences for everyday practice use (e.g., ranging from 13–27% in 2009) as well as clinical trial participation (e.g., ranging from 1–12% in 2006). Although our results indicate that the Dutch policy was thus, at least partially, effective in reaching its objective, the remaining regional differences may indicate the existence of residual inequality in access.

Consequently, coverage with evidence development seems to be a promising policy option to reduce policymaker uncertainty by guaranteeing early access to all patients in need while ensuring the possibility to revise the initial decision. **Chapter nine** investigates the feasibility to obtain sufficient evidence from everyday clinical practice for an evaluation regarding whether the high drug costs are worth public funding. The methods used in the three case studies were feasible to develop evidence on some aspects of drug use including types of treatments used, dosages, dose modifications, and health care costs. However, aspects such as baseline patient characteristics, reasons to start or stop a treatment, and treatment effects were less feasible because of important missing values. Despite difficulties to correct for confounding by indication, it was possible to estimate incremental cost-effectiveness by synthesising evidence in the two oxaliplatin case studies. This was, however, not possible in the bortezomib case study. The optimal approach may differ between drugs and their indications.

**Chapter ten** continues with evaluating whether a coverage with evidence development policy is always the best option to reduce policymaker uncertainty regarding the question whether the costs are worth public funding. Our study shows that the gathered data from everyday practice can provide valuable new evidence for policymakers regarding

various initial uncertainties arising from the gap between clinical trials and everyday practice. Even though patients in everyday practice received fewer treatment cycles (4 vs. 6) as well as lower dosages (13%) compared with trial patients, time to progression (6.8 vs. 6.2 months) and response rates (complete response 8% vs. 6%; very good, partial, and minimal response 55% vs. 41%) seemed reasonably similar. Moreover, the gathered evidence provided detailed data on treatment costs in everyday clinical practice. Patients treated with bortezomib were, however, not comparable to patients not treated with bortezomib despite attempts to correct for confounding. It was, therefore, impossible to develop a feasible model to obtain a valid incremental cost-effectiveness estimate. It is essential that policymakers carefully consider if a coverage with evidence development policy will sufficiently decrease initial uncertainty or whether other options (e.g., finance- and/or outcomes-based risk-sharing arrangements) may be more appropriate to ensure sufficient value for money of expensive drugs.

## Discussion

**Chapter eleven** discusses the main research findings and presents suggestions on how to improve the decision making in drug reimbursement. We evaluate the systems regarding the extent to which their decision-making process satisfies the conditions of legitimate decision making according to the Daniels and Sabin accountability for reasonableness framework. Regarding the publicity of the decision rationale condition, we found that the systems lack in sufficiently satisfying transparency of the decision-making process; especially appraisal lacks transparency. Regarding the condition of relevance of the decision criteria, we found that the reimbursement criteria reflect the system objectives, and the five systems satisfactorily seem to ensure stakeholder involvement. Regarding the revisability condition, we found that the systems make limited use of policies to systematically evaluate a drug's social value throughout its life cycle. Finally, regarding the enforcement condition, we found that all systems implemented a designated national agency responsible for reimbursement decision making.

Based on this evaluation, we identified strengths of and challenges for the systems; especially the challenges can be seen as opportunities for furthering the aim to improve current drug reimbursement decision-making practice. We then put forward four recommendations for the reimbursement systems. First, the systems can build upon their existing strengths. A shared strength is that all systems have a national system that evaluates whether or not a drug should be reimbursed out of public sources; all share similar and clear policy objectives. Second, all systems can improve transparency of their decision-making process. Hereto, we developed a framework that can be used as a tool

to better structure the decision-making process by transparently answering five key questions. Third, the systems can make the reimbursement process more cyclic in order to deal with uncertainty of evidence, for instance by implementing a coverage with evidence development policy, or other outcomes-based or financial-based risk-sharing schemes. It is important to carefully consider what type of policy will appropriately ensure a sufficient reduction of the initial uncertainty. This may depend of the type of drug and type of disease. It is, however, essential to enforce the consequences of an evaluation. Finally, we recommend encouraging reimbursement of drugs that deliver social value for money. In order to maintain the sustainability of the systems in the long term, all can improve decision making based on social value for money. There is much room for increasing the role of cost-effectiveness in actual decision-making practice.

This thesis contributes towards a better understanding of drug reimbursement systems. It provides empirical insights into reimbursement procedures, processes, and criteria, and additionally provides insights into the effectiveness, feasibility, and applicability of the coverage with evidence development policy tool to handle uncertainty of the evidence in the decision making.

# Samenvatting

## Introductie

De gemiddelde levensverwachting neemt toe in veel landen. Dit gaat gepaard met stijgende kosten, waardoor de financiële houdbaarheid van verschillende gezondheidszorgsystemen gevaar loopt. Veel landen hebben daarom de afgelopen jaren kosten beheersende beleidsmaatregelen genomen. De meeste van deze maatregelen betreffen de vergoeding van geneesmiddelen. Om op de markt gebracht te mogen worden, moet eerst de effectiviteit, kwaliteit en veiligheid van een geneesmiddel worden bewezen. Veel landen vereisen bovendien aanvullend bewijs van een medisch evaluatie onderzoek (ook wel *Health Technology Assessment* [HTA] genoemd) voor de besluitvorming over de vergoeding van het geneesmiddel. Dit laatste wordt vaak de vierde horde voor markttoegang genoemd. Aangezien ieder land de gezondheidszorg op verschillende wijze kan inrichten, bestaan er ook verschillen in de manier waarop wordt besloten of geneesmiddelen wel of niet vergoed worden uit algemene middelen.

Dit proefschrift richt zich op de besluitvorming op het gebied van geneesmiddelenvergoeding. Het overkoepelende doel was om geneesmiddelenvergoedingssystemen te beschrijven, te analyseren en te evalueren. Het proefschrift bestaat uit drie delen. Deel A beschrijft de doelen van een gezondheidszorg systeem en onderzoekt de maatschappelijk preferenties voor deze doelen. Deel B geeft een beschrijving van vijf Europese geneesmiddelenvergoedingssystemen. Dit deel richt zich op de vergoedingsprocedures en processen en geeft inzicht in het gebruik van vergoedingscriteria in deze landen. Deel C evalueert de beleidsmaatregel 'coverage with evidence development'. Dit instrument wordt binnen de geneesmiddelenvergoeding gebruikt om de gevolgen van onzekerheid in de besluitvorming te verminderen.

## Deel A: Maatschappelijke preferenties voor de doelen van de gezondheidszorg

**Hoofdstuk twee** beschrijft een uitgebreid internationaal raamwerk dat het mogelijk maakt om uitkomsten van gezondheidszorgsystemen met elkaar te vergelijken. Het hoofdstuk geeft tevens een theoretische onderbouwing voor dit kader. Het raamwerk bestaat uit vijf onafhankelijke doelen. Wij includeerden naast gezondheid zelf tevens het proces van zorg verlenen in ons utiliteitsraamwerk. Het zorgproces is namelijk een op zichzelf staand doel, onafhankelijk van of het de gezondheid beïnvloedt. Vervolgens werden de maatschappelijke preferenties gemeten voor deze vijf doelen met behulp van een discrete keuze-experiment. Gezondheidsgerelateerde doelen werden het meest belangrijk bevonden (rechtvaardige verdeling van gezondheid [0,34] en het gemiddelde

niveau van gezondheid [0,29]), gevolgd door financiële eerlijkheid [0,24], en proces gerelateerde doelen (nut afgeleid van het proces [0,07], en de verdeling ervan [0,06]). In vergelijking met eerder onderzoek vonden respondenten in ons onderzoek gezondheid en de ongelijkheid in gezondheid veel belangrijker. Dit verschil kwam mogelijk doordat onze respondenten de vragen beantwoordden vanachter een 'veil of ignorance' (sluier der onwetendheid). Proces gerelateerde uitkomsten kregen een lagere waardering.

## Deel B: Geneesmiddelenvergoedingsprocedures, processen en criteria

Dit deel van het proefschrift richt zich op de besluitvorming op het gebied van geneesmiddelenvergoeding in vijf Europese landen (België, Frankrijk, Oostenrijk, Nederland en Zweden). **Hoofdstuk drie** vergelijkt het geneesmiddelenvergoedingssysteem in deze vijf landen. De vijf systemen hebben gemeenschappelijke doelen: (i) eerlijke en gelijke toegang tot zorg, (ii) kwaliteit van zorg, en (iii) houdbaarheid van het systeem. Het doel van het beleid is het op een sociaal aanvaardbare wijze in evenwicht brengen van deze drie doelen. Ieder systeem heeft een onafhankelijk instituut dat evalueert of geneesmiddelen moeten worden gefinancierd vanuit het basispakket. De vijf systemen gebruiken hiervoor vergelijkbare vergoedingscriteria. Over het algemeen is men bereid geneesmiddelen met voldoende meerwaarde te vergoeden. De impact van het systeem wordt voornamelijk gemonitord aan de hand van het geneesmiddelen budget. Een opvallend verschil is dat in drie landen de minister discretionaire beslissingsbevoegdheid heeft (België, Frankrijk en Nederland), terwijl in twee landen dit niet het geval is (Oostenrijk en Zweden). De vijf systemen maken beperkt gebruik van beleidsmogelijkheden om geneesmiddelen doorlopend te evalueren.

De vijf systemen trachten met toenemende mate transparant te zijn in hun besluitvormingsproces, maar geen van de systemen past een formele hiërarchie toe van de vergoedingscriteria of maakt gebruikt van een expliciete kosteneffectiviteit drempelwaarde. Ons onderzoek toont aan dat de transparantie van het besluitvormingsproces verbeterd kan worden in alle vijf systemen; vooral de *appraisal* fase is onvoldoende transparant. In **hoofdstuk vier** presenteren wij een kader dat gebruikt kan worden om de legitimiteit van het besluitvormingsproces te vergroten. Dit kader bestaat uit vijf belangrijke vragen inclusief de daarbij behorende vergoedingscriteria. Deze vragen hebben betrekking op de (i) medische, therapeutische, en/of maatschappelijke noodzaak, (ii) bereidheid te betalen voor het behandelen van een bepaalde aandoening, (iii) bereidheid te betalen voor een bepaalde behandeling, (iv) bereidheid om meer te betalen in vergelijking tot alternatieve behandelingen, en de (v) daadwerkelijke bereidheid om deze behandeling te betalen uit publiekelijk gefinancierde middelen. Het door ons ontwikkelde kader

biedt een hulpmiddel om het besluitvormingsproces beter te structureren. Hiertoe dienen alle vijf vragen achtereenvolgens beantwoord te worden. Het systematisch beantwoorden van de vijf vragen verbetert de consistentie van de beslissingen en hun onderbouwing. Bovendien wordt door het duidelijk rapporteren van het antwoord op iedere vraag de transparantie van het besluitvormingsproces verbeterd.

Vervolgens richt het onderzoek zich op het gebruik van twee specifieke vergoedingscriteria (ziektelast en kosteneffectiviteit). **Hoofdstuk vijf** toont aan dat ziektelast een belangrijke overweging is in de besluitvorming in België, Frankrijk, Nederland en Zweden. Dit is vooral het geval indien de ziektelast hoog is. Ondanks dat zijn de landen weinig transparant in hoe belangrijk ziektelast is ten opzichte van de andere vergoedingscriteria. Het begrip ziektelast wordt in de vier landen op verschillende wijzen geoperationaliseerd. Het Nederlandse systeem gebruikt de 'proportional shortfall' methode om de ziektelast weer te geven, terwijl in Zweden de ernst van de ziekte wordt uitgedrukt in een aantal categorieën. Echter, in beide landen speelt ziektelast enkel een impliciete rol in de daadwerkelijke besluitvorming met betrekking tot de vergoeding van een geneesmiddel. In België en Frankrijk speelt ziektelast tevens een expliciete rol bij het bepalen van de maatschappelijke betalingsbereidheid. Beide landen gebruiken de ernst van de ziekte bij het bepalen van het vergoedingsniveau. Ons onderzoek toonde aan dat geen enkele respondent kwantitatieve informatie over de ernst van de ziekte (naast kwalitatieve informatie) van doorslaggevende waarde vond in de besluitvorming.

Tevens hebben wij de rol van kosteneffectiviteit in het besluitvormingsproces in Nederland en Zweden nader onderzocht (**hoofdstuk zes**). De rol van dit criterium lijkt veel bescheidener dan vaak wordt gesteld. Vooral in Nederland lijkt de daadwerkelijke rol vrij beperkt te zijn. In Zweden was veel vaker informatie van economische evaluaties beschikbaar dan in Nederland (97 vs 31, respectievelijk). Dit verschil leek voornamelijk te kunnen worden verklaard door het grote aantal gegeven vrijstellingen voor het aanleveren van doelmatigheidsonderzoek in Nederland (65%). In de periode van januari 2005 tot juli 2012 was slechts doelmatigheidsbewijs aanwezig in 35% van de 118 Nederlandse vergoedingsdossiers voor bijlage 1B (dat wil zeggen dossiers voor geneesmiddelen met aangetoonde therapeutische meerwaarde). Het doelmatigheidsonderzoek werd als onvoldoende onderbouwd beoordeeld voor alle geneesmiddelen die een negatief vergoedingsadvies kregen. Echter, tevens 21% van het doelmatigheidsonderzoek voor geneesmiddelen die een positief vergoedingsadvies kregen, werd beoordeeld als zijnde onvoldoende onderbouwd. Geneesmiddelen die in Zweden een negatief vergoedingsadvies kregen, werden niet kosteneffectief bevonden (74%) of kregen het oordeel dat er onvoldoende bewijs was met betrekking tot de kosteneffectiviteit (26%). Geneesmiddelen met een positief vergoedingsbesluit werden bijna allemaal als

kosteneffectief beschouwd (92%). Het is van belang om op te merken dat 53% van deze beoordelingen enkel gebaseerd was op een prijsvergelijking en 10% op een kostenminimalisatie studie. Slechts 33% van de positieve beoordelingen was gebaseerd op een volledig uitgevoerde economische evaluatie studie. Opvallend was dat er slechts voor 11 geneesmiddelen in zowel Nederland als Zweden een volledige economische evaluatie studie was uitgevoerd. Het vergoedingsbesluit verschillende voor 4 van de 11 geneesmiddelen.

**Hoofdstuk zeven** richt zich vervolgens specifiek op de rol van doelmatigheid in de besluitvorming in Nederland. HTA informatie was al in de 80-er jaren beschikbaar voor beleidsmakers. HTA informatie wordt echter alleen systematisch vereist bij besluitvorming betreffende geneesmiddelenvergoeding; maar ook hier lijkt de rol van doelmatigheid tamelijk gering. Ondanks dat het aanleveren van doelmatigheidsbewijs noodzakelijk is voor bepaalde groepen geneesmiddelen, is het de vraag of doelmatigheid wel daadwerkelijk een rol speelt in de dagelijkse besluitvorming. Uit recente casussen bleek dat Nederlandse beleidsmakers grote moeilijkheden ondervonden om beperkingen te stellen aan de vergoeding van bepaalde geneesmiddelen op basis van doelmatigheid. Dit was zelfs het geval bij extreem hoge kosten per voor kwaliteit gecorrigeerd levensjaar.

## Deel C: Omgaan met onzekerheid in de besluitvorming

Het besluit om een geneesmiddel wel of niet te vergoeden wordt inherent gemaakt met een mate van onzekerheid in het bewijs over de waarde van het geneesmiddel. In Nederland werd in 2006 een beleidsmaatregel (*coverage with evidence development*) ingevoerd voor dure intramurale geneesmiddelen. Deze maatregel werd ingesteld om gelijke toegang tot geneesmiddelen te garanderen en richt zich op de onzekerheid in het bewijs tijdens de initiële besluitvorming. Voor een positief vergoedingsadvies geldt de voorwaarde dat er aanvullend bewijs over het geneesmiddel moet worden verzameld door middel van uitkomstenonderzoek in de dagelijkse praktijk. Vanaf 2013 is de toepassing van deze maatregel verder uitgebreid naar specifieke groepen extramurale geneesmiddelen. **Hoofdstuk acht** onderzoekt of de ingestelde beleidsmaatregel het gewenste effect bereikte in Nederland (het garanderen van gelijke toegang tot dure geneesmiddelen) voor één van de geneesmiddelen (bortezomib) die onder de maatregel viel. Uit interviews bleek dat medisch specialisten zich bewust waren van de hoge kosten van het geneesmiddel maar dat zij geen barrières ondervonden om het middel in de dagelijkse praktijk te gebruiken. De studie toonde tevens aan dat artsen een lange periode nodig hadden voordat zij het middel regelmatig voorschreven. Ondanks dat het gebruik van het middel toe nam na het invoeren van conditionele aanvullende

vergoeding aan ziekenhuizen (van 2% in 2004 tot 17% in 2009), bleef het gebruik onder het niveau geschat door de beroepsvereniging van hematologen (27%). Verder toonde het onderzoek regionale verschillen aan: zowel in het gebruik in de dagelijkse praktijk (variërend van 13% tot 27% in 2009) als ook in deelname aan klinische studies (variërend van 1% tot 12% in 2006). Het onderzoek toonde enerzijds aan dat de beleidsmaatregel een positief effect had op de toegang tot het geneesmiddel, maar anderzijds bleek dat er nog steeds regionale verschillen waren die duiden op een ongelijke toegang tot het geneesmiddel.

De beleidsmaatregel 'coverage with evidence development' lijkt een veelbelovende beleidsoptie aangezien het zowel een snelle toegang tot nieuwe geneesmiddelen waarborgt als ook gelijktijdig de mogelijkheid biedt om het vergoedingsbesluit te herzien. In **hoofdstuk negen** onderzoeken wij of het haalbaar is om voldoende bewijs te verzamelen, door middel van uitkomstenonderzoek in de dagelijkse klinische praktijk, voor een beoordeling of geneesmiddelen moeten worden gefinancierd vanuit het basispakket. De methoden die wij hebben gebruikt in drie *case studies* maakten het mogelijk om voldoende bewijs te verzamelen over welke patiënten worden behandeld, hoe geneesmiddelen worden gebruikt in de dagelijkse praktijk (dosis, dosisaanpassingen), en welke kosten dit met zich meebrengt. Echter, deze methoden bleken minder goed om bewijs te verzamelen over patiënt karakteristieken, redenen om te starten of stoppen met een behandeling, en de effecten van een behandeling. Het grootste probleem was dat er veel belangrijke informatie miste. In twee *case studies* (oxaliplatin bij twee indicaties) was het mogelijk om een incrementele kosteneffectiviteitsratio te schatten ondanks dat er problemen waren met het corrigeren voor verstorende variabelen (*confounding by indication*). In de derde case studie (bortezomib) bleek dit echter niet mogelijk. De optimale aanpak van uitkomstenonderzoek kan verschillend zijn per geneesmiddel en per indicatie.

Vervolgens wordt in **hoofdstuk tien** besproken of het toepassen van deze beleidsmaatregel de beste oplossing biedt voor het omgaan met onzekerheid in het bewijs. Onze studie laat zien dat uitkomstenonderzoek waardevolle informatie opleverde voor beleidsmakers ten aanzien van diverse onzekerheden als gevolg van verschillen tussen behandelingen gegeven in klinische *trial* onderzoeken en behandelingen gegeven in de dagelijkse praktijk. Patiënten die werden behandeld in de dagelijkse praktijk ontvingen minder behandelingscycli (gemiddeld 4 versus 6) en werden behandeld met lagere doseringen (13% lager) in vergelijking tot patiënten die werden behandeld in *trial* verband. De respons op de behandeling leek vergelijkbaar (tijd tot progressie 6.8 versus 6.2 maanden; complete respons 8% versus 6%; zeer goede, gedeeltelijke en minimale respons 55% versus 41%). Uitkomstenonderzoek leverde tevens gedetailleerde

gegevens over de kosten van de behandeling in de dagelijkse praktijk. Patiënten die met bortezomib waren behandeld waren echter niet vergelijkbaar met patiënten die niet met bortezomib waren behandeld, ondanks pogingen te corrigeren voor verstorende variabelen. Het bleek daarom onmogelijk om een haalbaar model te ontwikkelen dat een valide incrementele kosteneffectiviteitsratio zou berekenen. Het is essentieel dat beleidsmakers zorgvuldig overwegen of uitkomstenonderzoek voldoende informatie oplevert om de initiële onzekerheid te verminderen of dat andere beleidsmaatregelen (bijvoorbeeld een prijsovereenkomst of een op uitkomsten gebaseerde overeenkomst) wellicht meer geschikt zijn om een betere kosten baten verhouding voor dure geneesmiddelen te garanderen.

## Discussie

**Hoofdstuk elf** geeft een overzicht van de belangrijkste onderzoeksresultaten en geeft suggesties ter verbetering van het besluitvormingsproces. Het geneesmiddelenvergoedingssysteem hebben wij geëvalueerd naar de mate waarin het besluitvormingsproces voldoet aan de legitimiteitsvoorwaarden analoog aan het Daniels en Sabin 'accountability for reasonableness' raamwerk. Ten aanzien van de voorwaarde 'publiciteit van de beslissing rationale', toonde ons onderzoek aan dat de vijf bestudeerde geneesmiddelenvergoedingssystemen onvoldoende transparant zijn in hun besluitvormingsproces; vooral de *appraisal* fase is onvoldoende transparant. Ten aanzien van de voorwaarde 'relevantie van de beslissingscriteria', zagen wij dat de doelstellingen van het systeem gereflecteerd werden in de toegepaste vergoedingscriteria. Bovendien leek ieder systeem de betrokkenheid van stakeholders voldoende te garanderen. Ten aanzien van de revisie voorwaarde, toonde ons onderzoek aan dat de systemen maar beperkt gebruik maken van de mogelijkheid om de maatschappelijke waarde van een geneesmiddel doorlopend te evalueren. Tenslotte, met betrekking tot de handhavingsvoorwaarde, vonden we dat alle systemen een nationale instantie hadden aangesteld, verantwoordelijk voor de besluitvorming betreffende geneesmiddelenvergoeding.

Op basis van het onderzoek in deze thesis was het mogelijk om sterke punten van en uitdagingen voor geneesmiddelenvergoedingssystemen te benoemen. Op basis hiervan geven wij vier aanbevelingen om het systeem te verbeteren. Ten eerste is het belangrijk dat de systemen voortbouwen op hun bestaande sterke punten. Alle systemen beschikken over een nationaal orgaan dat adviseert/beoordeelt of een geneesmiddel moet worden vergoed uit publiekelijk gefinancierde middelen; de systemen hebben vergelijkbare en duidelijke doelstellingen. Ten tweede is het voor alle systemen aan te bevelen om de transparantie van het besluitvormingsproces te verbeteren. Wij ontwikkelden een

raamwerk om het besluitvormingsproces beter te structureren en de transparantie van het proces te bevorderen. Ten derde is het belangrijk dat de systemen hun besluitvormingsproces meer cyclisch maken om zodoende om te gaan met onzekerheid in het initiële bewijs. Dit kan bijvoorbeeld door een voorwaardelijk vergoedingstraject af te spreken onder de conditie dat er aanvullend bewijs wordt verzameld, of door het delen van risico's door middel van een op uitkomsten gebaseerde of een financieel gebaseerde overeenkomst. Het is belangrijk om zorgvuldig af te wegen welke overeenkomst het beste past bij welk type onzekerheid; dit kan per type geneesmiddel en per indicatie verschillend zijn. Daarnaast is het essentieel om het vergoedingsbeleid zo nodig aan te passen op basis van de uitkomst van de herbeoordeling. Als laatste promoten wij het vergoeden van geneesmiddelen die voldoende maatschappelijke waarde bieden. Het is voor alle systemen mogelijk om vergoedingsbeslissingen in grotere mate te baseren op 'value-for-money'; kosteneffectiviteit kan een veel grotere rol spelen in de dagelijkse besluitvorming. Dit komt ten goede aan de duurzaamheid van het systeem.

Dit proefschrift draagt bij aan een beter inzicht in de besluitvorming op het gebied van geneesmiddelenvergoeding. Het geeft empirische inzichten in vergoedingsprocedures, processen en criteria. Bovendien geeft het proefschrift inzicht in de effectiviteit, haalbaarheid en toepasbaarheid van een beleidsmaatregel ('coverage with evidence development') om de gevolgen van onzekerheid in de besluitvorming te verminderen.

# Dankwoord

Promoveren, wie had dat ooit gedacht? Na ruim 10 jaar als verpleegkundige te hebben gewerkt, voornamelijk op de afdeling hartbewaking en intensive care in Arnhem, was het al wel een tijd duidelijk: ik wil wat anders, ik wil meer uitdaging in mijn werk! Die uitdaging zocht ik eerst in het buitenland. In Nieuw Zeeland waar ik, wederom op de afdeling hartbewaking, als verpleegkundige heb gewerkt en natuurlijk niet te vergeten, de fantastische tijd die Jason en ik samen rondtrokken door Nieuw Zeeland, Australië, Pacific eilanden en zuidoost Azië. Toen ik na 3 jaar te zijn weggeweest precies 1 week terug was in Nederland, begon ik in 2005, nog enigszins gedesoriënteerd aan het schakeljaar van de BMG.

Tijdens de HEPL master kwam het idee al eens in mij op, misschien wil ik wel bij het iMTA werken. Toen er een vacature voor junior onderzoeker beschikbaar kwam, hoefde ik dan ook niet lang na te denken. Ook toen wist ik nog niet of ik wilde gaan promoveren, het is immers nogal een grote overstap van verpleegkundige naar wetenschapper. Ondanks dat ik dit ook eerlijk vertelde in mijn sollicitatiegesprek werd ik wel aangenomen. Carin en Marc, jullie hadden dit goed in geschat, jullie boden mij de uitdaging die ik nodig had, met als resultaat dit proefschrift, en dus ja, promoveren!

Graag wil ik iedereen bedanken die op welke manier dan ook heeft bijgedragen aan de totstandkoming van dit proefschrift. Iemand die mij goed kent, weet dat ik het dankwoord hier eigenlijk bij zou willen laten. Maar, de afgelopen zes jaar is duidelijk geworden dat het dankwoord het meest gelezen hoofdstuk van een proefschrif is, en een hoofdstuk met één zin is geen hoofdstuk. Bovendien heb ik natuurlijk zelf kunnen ervaren dat je een proefschrift niet alleen schrijft en daarom wil ik dan ook graag een aantal mensen persoonlijk bedanken.

Allereerst wil ik mijn promotor Carin Uyl-de Groot en mijn co-promotor Marc Koopmanschap bedanken. Jullie gaven mij veel meer dan alleen een baan. Carin, bedankt voor al het vertrouwen dat je mij vanaf het begin hebt gegeven, en voor de ruimte die je mij biedt om verder te groeien. Jouw onderzoek is belangrijk voor patiënten, je onderzoekshart ligt op de goede plek. De plek waar de patiënten dit heel hard nodig hebben, namelijk het bevorderen van kwaliteit van zorg en een eerlijke toegang tot zorg. Ondanks je drukke agenda, is het toch altijd mogelijk om een gaatje te vinden. Zelfs als dit bijna onmogelijk lijkt, creëert 'onmisbare' Cobi toch een gaatje, of biedt Whatsapp alsnog gegarandeerd succes! Marc, we hebben heel wat uurtjes samen zitten bomen, onderzoek moet nuttig zijn en meer inhouden dan alleen wetenschap bedrijven. We hebben samen in het Escher project veel verder kunnen kijken dan de standaard 'kruisjesschema's' en daardoor veel over het beleid en beleidsmakers geleerd. Ons onderzoek hebben we uitgebreid kunnen presenteren, zowel bij Nederlandse beleidsmakers als

ook op verschillende congressen. Jammer dat het ons tot nu toe nog niet gelukt is om verdere financiering te vinden om ons onderzoek te vervolgen. Misschien moeten we toch niet te vaak hardop zeggen dat economische evaluaties alleen maar dienen om HTA onderzoekers aan het werk te houden?

Tevens wil ik mijn tweede promotor Ton de Boer graag bedanken. Ook al was je wat meer op afstand betrokken bij mijn proefschrift, jouw feedback was een goede aanvulling en daardoor kon je ons onderzoek naar een hoger niveau brengen, mede jouw ervaring als lid van de wetenschappelijke adviesraad van het zorginstituut was inhoudelijk belangrijk.

De promotiecommissie wil ik graag bedanken voor het lezen en beoordelen van dit proefschrift en voor het opponeren bij de verdediging.

De hoofdstukken in dit proefschrift zijn geschreven in samenwerking met collega onderzoekers. Ik ben alle coauteurs dankbaar voor hun inbreng, kennis en inzichten. Graag wil ik een aantal in het bijzonder bedanken: Jennifer Gaultney, Chantal van Gils, Ken Redekop, Hedwig Blommestein en Irina Cleemput. Jennifer en Chantal, wij begonnen niet precies tegelijkertijd aan de pilot uitkomstenonderzoek studies, maar wel alle drie even gedreven. Hoera, dit is het laatste proefschrift, kunnen we dan nu eindelijk het feestje vieren dat de pilot studies echt afgelopen zijn? Jennifer, wij hebben samen heel hard aan die onmogelijke bortezomib studie gewerkt, en onze powerpoint slide met die enorme behandelvariatie hebben we al heel vaak op veel verschillende plekken terug gezien. Ik weet zeker dat ook het laatste artikel van bortezomib gepubliceerd gaat worden, ik geloof er nog steeds in! Chantal, ook nadat je bij iMTA wegging bleven we samen werken, onze opgedane pilot kennis konden we goed gebruiken om de melanoom studies beter op te zetten en uit te voeren. Jammer dat onze samenwerking nu eindigt, maar ik wens je veel plezier met je nieuwe baan in Brussel. Ken, ook jij was betrokken bij de pilots, ik heb heel veel geleerd van je kritische reflecties. Hedwig, je begon als student assistent bij de bortezomib pilot, inmiddels ben je al jaren een fantastische collega en nog steeds bezig met multipel myeloom. Blijf zoals je bent! Irina, we hebben samen hard gewerkt aan de vijf landen vergelijking, wat ontzettend leuk dat jij nu opponeert bij mijn verdediging.

Ik kijk terug op ruim zes fijne jaren bij GE-iMTA, collega's dank hiervoor, jullie hebben hier aan bijgedragen. Heel bijzonder vond ik onze deelname aan de ROPARUN, ik vond het fantastisch, ik zou het zo weer doen (inclusief de TravelJohns). Henk Sonneveld, jij was maar voor 1 jaar mijn kamergenoot, maar dit was wel ons beider eerste jaar bij de iBMG, waardoor we veel konden delen. Ik vind het leuk dat we elkaar nog af en toe zien.

Rian Rijnsburger, de jaren erna was jij mijn kamergenote, wij hebben veel meer gedeeld dan alleen de echte lekkere koffie. Ik weet niet wie van ons het meest verslaafd is, maar jij lijkt wel degene te zijn die onthoudingsverschijnselen krijgt als er geen echte koffie is. Ik wil me hierbij nog eenmaal verontschuldigen voor al die hoge stapels waar jij altijd tegen aan moest kijken, jouw stapeltjes waren altijd netter en veel meer georganiseerd dan die van mij. Wat jammer dat je de iBMG gaat verlaten, ik ga je gezelligheid missen. Ik wens je heel veel plezier in je nieuwe baan!

En dan natuurlijk, niet het minst belangrijk, mijn paranimfen: Anja Mallison en Anouk O'Prinsen. Anja, ik moest gaan rekenen om te bedenken hoe lang wij elkaar al niet kennen. Ik ben maar gestopt bij 30 jaar, anders lijken we zo oud. In die jaren hebben wij zo ontzettend veel gedeeld, het is jammer dat we elkaar mede door de afstand niet vaker zien, maar eigenlijk maakt het niet uit of jij nu in Duitsland, Amerika, of Engeland woont, of dat ik in Nieuw Zeeland of Nederland woon of ergens anders op de wereld vertoef. Gelukkig is er altijd een goedkope telefoonaanbieder te vinden! Wat fijn dat jij mijn paranimf bent. Anouk, wij hebben elkaar leren kennen in het schakelprogramma en samen de HEPL master gedaan. Tijdens onze opleiding, maar ook nu nog steeds bediscussiëren wij regelmatig hoe dat nou toch zit in die klinische studies of doelmatigheidsstudies. Jij hebt de vorderingen in mijn proefschrift de afgelopen jaren steeds bijgehouden, waarschijnlijk ben je een van de weinige die de meeste hoofdstukken van mijn proefschrift heeft gelezen.

Tot slot vrienden en familie, een aantal van jullie heeft heel hard meegeholpen met de pilot testen van hoofdstuk twee. Verder hebben jullie natuurlijk op allerlei verschillende manieren indirect bijgedragen aan dit proefschrift. Nathalie, speciaal voor jou, het laatste hoofdstuk van de Libelle is hierbij klaar! Jason, bedankt voor alle ruimte die ik altijd van je krijg. Tess, jij brengt zoveel vrolijkheid en gezelligheid in ons leven!

# PhD Portfolio

PhD student: Margreet Franken

Institute: Institute of Health Policy and Management, Erasmus University Rotterdam, The Netherlands

PhD period: 2008–2014

## PhD Training

Pharmaceutical Policy Analysis, Summer School, Utrecht University, Utrecht, The Netherlands (2008)

Presentations skills, Erasmus University Rotterdam, The Netherlands (2008, 2009, 2010)

Regression Analysis, Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, The Netherlands (2009)

Survival Analysis, Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, The Netherlands (2009)

Training in Problem-based Learning (PGO), Erasmus University, Rotterdam, The Netherlands (2011)

Discrete choice modelling, Erasmus University, Rotterdam, The Netherlands (2011)

Drug Discovery & Development Cycle course, Top Institute Pharma, Zeist, The Netherlands (2011)

Facilitators for Drug Development, Top Institute Pharma Training program, The Netherlands (2009)

## Teaching

Socio-medical sciences, tutor in bachelor and premaster program Health Sciences, Institute of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands (2010–2013)

Quality and efficiency in healthcare, tutor and lecturer in bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands (2012–2014)

Supervision and co-supervision of several master theses, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands (2009–2014)

## International conferences

International Health Economist Association (iHEA): Beijing, China (2009)

European Conference on Health Economics, (ECHE): Helsinki, Finland (2010)

Health Technology Assessment International (HTAi): Dublin, Ireland (2010); Rio de Janeiro, Brazil (2011)

Pharmaceutical Pricing and Reimbursement Information (PPRI): Vienna, Austria (2011)

International Society for Pharmacoeconomics and Outcomes Research (ISPOR): Athens, Greece (2008); Paris, France (2009); Prague, Czech Republic (2010); Berlin, Germany (2012); Dublin, Ireland (2013); Amsterdam, The Netherlands (2014)

## Presentations at international conferences

### Issue panels

Real world experience with access with evidence development (AED): Learning from partial successes, Health Technology Assessment International (HTAi), Dublin, Ireland (2010)

Decision making in health care based on economic evaluation: reality or just wishful thinking? Experiences from four European countries, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Amsterdam, The Netherlands (2014)

### Podium presentations

Health System Goals Valuation: A discrete choice experiment behind a veil of ignorance, International Health Economist Association (iHEA), Beijing, China (2009)

Treatment variation complicates real-world pharmacoeconomics: Daily clinical practice of bortezomib in relapsed or refractory multiple myeloma, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Paris, France (2009)

Dynamics in daily practice challenge real-world pharmacoeconomics: Clinical practice of bortezomib in relapsed or refractory multiple myeloma, Health Technology Assessment International (HTAi), Dublin, Ireland (2010)

Convergence or divergence in European drug reimbursement systems? Can the Dutch, Swedish, Belgian, French and German system learn from each other in deciding on value for money? European Conference on Health Economics, (ECHE), Helsinki, Finland (2010)

Deciding on value for money: a comparison of the Dutch, Belgian, Swedish and French drug reimbursement systems, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Prague, Czech Republic (2010)

Sustainability of drug reimbursement systems: a comparison of the Austrian, Belgian, Dutch, French and Swedish system, Health Technology Assessment International (HTAi), Rio de Janeiro, Brazil (2011)

Performance of drug reimbursement systems: A comparison of the Austrian, Belgian, Dutch, French and Swedish systems, Pharmaceutical Pricing and Reimbursement Information (PPRI), Vienna, Austria (2011)

A detailed comparison of Dutch and Swedish drug reimbursement decisions: what evidence is available, which criteria are used, and is the decision-making process transparent? International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Berlin, Germany (2012)

**Workshops**

'One size does not fit all.' The Netherlands' first experiences with performing outcomes research on behalf of health care policy making: challenging the tension between the optimal and the feasible, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Prague, Czech Republic (2010)

As real as it gets: challenges in setting up patient registries for the collection of real-world data on behalf of policy making, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Amsterdam, The Netherlands (2014)

### Poster presentations

Decision making in drug reimbursement: A first glance at the role of pharmacoeconomics in the Dutch system, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Paris, France (2009)

The role of pharmacoeconomics in Dutch drug reimbursement: A toothless tiger? International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Madrid, Spain (2011) [Poster award finalist]

Policymaker, please carefully consider your needs: Does outcomes research of bortezomib for advanced multiple myeloma reduce uncertainty? International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Berlin, Germany (2012)

A comparative study of the role of disease severity in drug reimbursement decision-making in Belgium, France, The Netherlands and Sweden, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Dublin, Ireland (2013)

## Other meetings, workshops and contributions

### Presentations

Presentation *"Pilots outcomes research"* at the pharmaceutical advisory committee (CFH), Diemen, The Netherlands (2010)

Presentation *"Challenges of real-world data collection"* at The Netherlands Organisation for Health Research and Development (ZonMw), The Hague, The Netherlands (2010)

Presentation *"Performance of drug reimbursement systems: A comparison of the Austrian, Belgian, Dutch, French and Swedish system"* at Dutch Ministry of Health, The Hague, and the Dutch Health Care Insurance Board (CVZ), Diemen, The Netherlands (2011)

Poster presentation *"Legitimacy of decision-making in drug reimbursement: A comparison of the Austrian, Belgian, Dutch, French and Swedish system"* at Top Institute Pharma Springmeeting (2011)

### National conferences and meetings

Pharma National Conference, Rotterdam, The Netherlands (2008)

Lowlands Health Economists' Study Group (lolaHESG): Berg en Terblijt, The Netherlands (2009); Egmond aan Zee, The Netherlands (2010); Soesterberg, The Netherlands [paper discussed by Sylvia Vijgen] (2011); Voorne, The Netherlands [paper discussed by Johan Polder, and discussant of paper Reina de Kinderen et al.] (2014)

Winter meeting World Health Organisation and Utrecht University, Utrecht, The Netherlands (2010)

Symposium *"Back to the future",* Dutch Association for Technology Assessment in Health care (NVTAG), Rotterdam, The Netherlands (2011)

**Workshops and courses**

The appraisal process: work in progress, two-day course by the Dutch Association for Technology Assessment in Health care (NVTAG) and Dutch Health Care Insurance Board (CVZ), The Netherlands (2009)

Meten van maatschappelijk draagvlak pakketadviezen, one-day workshop by Netherlands Institute for health services research (NIVEL) and Leiden Univerisity Medical Center (LUMC), and VU University Medical Center (VUMC), The Netherlands (2009)

Methodological workshop citizen and patient participation, Belgian Health Care Knowledge Centre, Brussels, Belgium (2012)

**Other**

Bachelor program Pharmaceutical innovation, lecturer *"Drug reimbursement"*, Utrecht University, Utrecht, The Netherlands (2012)

Research seminars at the Institute for Health Policy and Management (2008–2014)

Market access payers advisory board meetings (2011, 2012, 2013, 2014)

## Scientific Award

Best new investigator podium presentation for the presentation *"Treatment variation complicates real-world pharmacoeconomics: Daily clinical practice of bortezomib in relapsed or refractory multiple myeloma"* at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) annual European Conference, Paris, France (2009)

# List of publications

## List of scientific publications

Siok Swan Tan, Chantal Van Gils, Margreet Franken, Leona Hakkaart-van Roijen, Carin Uyl-de Groot. The Unit Costs of Inpatient Hospital Days, Outpatient Visits, and Daycare Treatments in the Fields of Oncology and Hematology. *Value in Health*. 2010;13(6):712-719.

Irina Cleemput, Margreet Franken, Maïté Le Polain, Marc Koopmanschap. European drug reimbursement systems' legitimacy: Five-country comparison and policy tool. *International Journal of Technology Assessment in Health Care*. 2012;28(4):358-366.

Margreet Franken, Maïté Le Polain, Irina Cleemput, Marc Koopmanschap. Similarities and differences between five European drug reimbursement systems. *International Journal of Technology Assessment in Health Care*. 2012;28(4):349-357.

Frank Sandmann, Margreet Franken, Adri Steenhoek, Marc Koopmanschap. Do reassessments reduce the uncertainty of decision making? Reviewing reimbursement reports and economic evaluations of three expensive drugs over time. *Health Policy*. 2013;112(3):285-296.

Jennifer Gaultney, Margreet Franken, Siok Swan Tan, William Redekop, Peter Hujgens, Pieter Sonneveld, Carin Uyl-de Groot. Real-world health care costs of relapsed/refractory multiple myeloma during the era of novel cancer agents. *Journal of Clinical Pharmacy and Therapeutics*. 2013;38(1):41-47.

Margreet Franken, Chantal van Gils, Jennifer Gaultney, Gepke Delwel, Wim Goettsch, Peter Huijgens, Adri Steenhoek, Cees Punt, Miriam Koopman, William Redekop, Carin Uyl-de Groot. Practical feasibility of outcomes research in oncology: lessons learned in assessing drug use and cost-effectiveness in The Netherlands. *European Journal of Cancer*. 2013;49(1):8-16.

Margreet Franken, Fredrik Nilsson, Frank Sandmann, Anthonius de Boer, Marc Koopmanschap. Unravelling drug reimbursement outcomes: A comparative study of the role of pharmacoeconomic evidence in Dutch and Swedish reimbursement decision making. *Pharmacoeconomics*. 2013;31(9):781-797.

Margreet Franken, Xander Koolman. Health System Goals: A discrete choice experiment to obtain societal valuations. *Health Policy*. 2013;112(1-2):28-34.

Hedwig Blommestein, Margreet Franken, Sylvia Verelst, Michel van Agthoven, Peter Huijgens, Carin Uyl-de Groot. Access to expensive cancer drugs in Dutch daily practice: Should we be concerned? *The Netherlands Journal of Medicine*. 2014;72(4):235-241.

Margreet Franken, Jennifer Gaultney, Hedwig Blommestein, Peter Huijgens, Pieter Sonneveld, William Redekop, Carin Uyl-de Groot. Policymaker, please consider your needs carefully: Does outcomes research in relapsed or refractory multiple myeloma reduce policymaker uncertainty regarding value for money of bortezomib? *Value in Health*. 2014;17(2):245-253.

Margreet Franken, Marc Koopmanschap, Adri Steenhoek. Health economic evaluations in reimbursement decision making in The Netherlands: Time to take it seriously? *German Journal for Evidence and Quality in Health Care*. 2014. [In Press]

Jennifer Gaultney, Margreet Franken, Carin A. Uyl-de Groot, William Redekop, Peter Huijgens, Bronno van der Holt, Henk Lokhorst, Pieter Sonneveld. Experience with outcomes research into the real-world effectiveness of novel therapies in Dutch daily practice from the context of conditional reimbursement. [Revision submitted]

Margreet Franken, Elly Stolk, Anthonius de Boer, Marc Koopmanschap. A comparative study of the role of disease severity in drug reimbursement decision making in four European countries. [Submitted]

## List of other publications

Adri Steenhoek, Marc Koopmanschap, Margreet Franken, Frans Rutten. Nieuwe geneesmiddelen: niet goed, geld terug? *Nederlands Tijdschrift voor Geneeskunde*. 2010;154:A2042.

Margreet Franken, Jennifer Gaultney, Hedwig Blommestein, Siok Swan Tan, Chantal van Gils, Peter Huijgens, Pieter Sonneveld, William Redekop, Carin Uyl- de Groot. Pilot outcomes research: Effects and costs of bortezomib in relapsed or refractory multiple myeloma. Report to the Dutch Health Care Insurance Board (CVZ). 2010.

Margreet Franken, Chantal van Gils, Jennifer Gaultney, Peter Huijgens, Adri Steenhoek, William Redekop, Carin Uyl- de Groot. Policy regulations for expensive medicines: Methodological issues based on pilot outcomes research studies. Report to the Dutch Health Care Insurance Board (CVZ). 2010.

Maïté le Polain, Margreet Franken, Marc Koopmanschap, Irina Cleemput. Drug reimbursement systems: international comparison and policy recommendations. Brussels: Belgian Health Care Knowledge Centre (KCE). KCE Reports 147C. D/2010/10.273/90. 2010.

Marc Koopmanschap, Hans Severens, Margreet Franken. Internationale aspecten bij besluitvorming inzake geneesmiddelenvergoeding. *Economische Statistische Berichten: Dossier Gepaste Zorg*. 2012;97(4644S):36-39.

# About the author

Margreet Franken is a researcher at the institute for Medical Technology Assessment (iMTA), Erasmus University Rotterdam, The Netherlands. She obtained a Bachelor degree in Nursing (1992), after which she specialised in coronary (1995) and intensive care nursing (1998). From 1992 until 2008 she worked as a quality care manager, clinical nurse educator, and specialised coronary and intensive care nurse in hospitals in Arnhem, Rotterdam and Christchurch (New Zealand). In 2008, Margreet obtained a masters degree cum laude in Health Economics, Policy and Law from the Erasmus University in Rotterdam. Since this time she has been working as a researcher at the institute for Medical Technology Assessment of the institute of Health Policy and Management at the Erasmus University in Rotterdam. Her research interests are in the field of health economics and include drug reimbursement policies and outcomes research. She is currently involved in disease registries for melanoma, colon and lung cancer.

# References

1. Van der Horst A, van Erp F, de Jong J. *Trends in health and health care* [in Dutch: Trends in gezondheid en zorg]. Den Haag: Centraal Planbureau. Policy Brief 11; 2011.
2. OECD. Health at a glance 2013: OECD indicators. 2013.
3. Folland S, Goodman AC, Stano M. *The economics of health and health care*. 5th ed. New Jersey, United States of America: Pearson Prentice Hall Upper Saddle River; 2007.
4. StatLine. *Zorgrekeningen; uitgaven en financiering vanaf 1972*. Available from http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLnl&PA=71988ned&LA=nl. Updated May, 2014.
5. De Meijer C, Wouterse B, Polder J, Koopmanschap M. The effect of population aging on health expenditure growth: A critical review. *European Journal of Ageing*. 2013;10(4):353-361.
6. Draborg E, Gyrd-Hansen D, Bo Poulsen P, Horder M. International comparison of the definition and the practical application of health technology assessment. *International Journal of Technology Assessment in Health Care.* 2005;21(01):89-95.
7. Banta HD, Bonsel G, Casparie AF, et al. *Medische technology assessment en gezondheidsbeleid*. Alphen aan den Rijn: Samson Stafleu; 1989.
8. Oliver A, Mossialos E, Robinson R. Health technology assessment and its influence on health-care priority setting. *International Journal of Technology Assessment in Health Care*. 2004;20(01):1-10.
9. Banta D. The development of health technology assessment. *Health Policy*. 2003;63(2):121-132.
10. Jonsson E. Development of health technology assessment in Europe. *International Journal of Technology Assessment in Health Care*. 2002;18(02):171-183.
11. Towse A. The efficient use of pharmaceuticals: Does Europe have any lessons for a medicare drug benefit? *Health Affairs*. 2003;22(3):42-45.
12. Hutton J, McGrath C, Frybourg JM, Tremblay M, Bramley-Harker E, Henshall C. Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). *International Journal of Technology Assessment in Health Care*. 2006;22(1):10-18.
13. Vuorenkoski L, Toiviainen H, Hemminki E. Decision-making in priority setting for medicines: A review of empirical studies. *Health Policy*. 2008;86(1):1-9.
14. Martikainen JE. Should the decision-making process of reimbursing medicines be improved? *International Journal of Technology Assessment in Health Care*. 2012;28(04):347-348.
15. Le Polain M, Franken MG, Koopmanschap MA, Cleemput I. *Drug reimbursement systems: International comparison and policy recommendations.* Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE). 2010. KCE Reports 147C. D/2010/10.273/90
16. Steenhoek A, Koopmanschap MA, Franken MG, Rutten FFH. New drugs: No cure, no pay? [in Dutch: Nieuwe geneesmiddelen: Niet goed, geld terug?] *Nederlands Tijdschrift voor Geneeskunde*. 2010;152:A2024.

17. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: When is there sufficient evidence? *Value in Health*. 2005;8(4):433-446.

18. Stafinski T, McCabe CJ, Menon D. Funding the unfundable: Mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics*. 2010;28(2):113-142.

19. Borst E. Changes in the attachments of the pharmaceutical benefit scheme [in Dutch: Wijziging bijlagen regeling farmaceutische hulp 1996]. *Official gazette of the Dutch State*. 2000;April, 6:10.

20. Raftery J. NICE and the challenge of cancer drugs. *BMJ*. 2009;338:271-272.

21. Van Luijn JC, Gribnau FW, Leufkens HG. Availability of comparative trials for the assessment of new medicines in the European union at the moment of market authorization. *British Journal of Clinical Pharmacology*. 2007;63(2):159-162.

22. Goldberg NH, Schneeweiss S, Kowal MK, Gagne JJ. Availability of comparative efficacy data at the time of drug approval in the United States. *The Journal of the American Medical Association*. 2011;305(17):1786-1789.

23. Claxton K. OFT, VBP: QED? *Health Economics*. 2007;16(6):545-558.

24. Carbonneil C, Quentin F, Lee-Robin SH. A common policy framework for evidence generation on promising health technologies. *International Journal of Technology Assessment in Health Care*. 2009;25(Suppl 2):56-67.

25. Garrison LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR real-world data task force report. *Value in Health*. 2007;10(5):326-335.

26. Hutton J, Trueman P, Henshall C. Coverage with evidence development: An examination of conceptual and policy issues. *International Journal of Technology Assessment in Health Care*. 2007;23(4):425-432.

27. Briggs A, Ritchie K, Fenwick E, Chalkidou K, Littlejohns P. Access with evidence development in the UK: Past experience, current initiatives and future potential. *Pharmacoeconomics*. 2010;28(2):163-170.

28. Klemp M, Frønsdal KB, Facey K. What principles should govern the use of managed entry agreements? *International Journal of Technology Assessment in Health Care*. 2011;27(1):77-83.

29. Tunis SR, Pearson SD. Coverage options for promising technologies: Medicare's 'coverage with evidence development'. *Health Affairs*. 2006;25(5):1218-1230.

30. Mohr PE, Tunis SR. Access with evidence development: The US experience. *Pharmacoeconomics*. 2010;28(2):153-162.

31. Trueman P, Grainger DL, Downs KE. Coverage with evidence development: Applications and issues. *International Journal of Technology Assessment in Health Care*. 2010;26(01):79-85.

32. Krumholz HM. Outcomes research: Generating evidence for best practice and policies. *Circulation*. 2008;118(3):309.

33. Clancy CM, Eisenberg JM. Outcomes research: Measuring the end results of health care. *Science*. 1998;282(5387):245.

34. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy*. 2010;96(3):179-190.

35. Adamski J, Godman B, Ofierska-Sujkowska G, et al. Risk sharing arrangements for pharmaceuticals: Potential considerations and recommendations for European payers. *BMC health services research*. 2010;10(1):153.

36. Carlson JJ, Garrison LP,Jr, Sullivan SD. Paying for outcomes: Innovative coverage and reimbursement schemes for pharmaceuticals. *Journal of Managed Care Pharmacy*. 2009;15(8):683-687.

37. World Health Organization. *The world health report 2000 - health systems: Improving performance*. Geneva: World Health Organization; 2000.

38. Ratcliffe J, Buxton M. Patients' preferences regarding the process and outcomes of life-saving technology. *International Journal of Technology Assessment in Health Care*. 1999;15(2):340-351.

39. Tyler TR. Process utility and help seeking: What do people want from experts? *Journal of Economic Psychology*. 2006;27(3):360-376.

40. Van Empel IWH, Dancet EAF, Koolman XHE, et al. Physicians underestimate the importance of patient-centredness to patients: A discrete choice experiment in fertility care. *Human Reproduction*. 2011;26(3):584-593.

41. Parsons T. *The structure of social action*. New York: The Free Press; (1937) 1968.

42. Deci EL, Ryan RM. *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Publishing Co; 1985.

43. Deci EL, Ryan RM. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*. 2000;11(4):227-268.

44. Smith PC, Mossialos E, Papanicolas I, Leatherman S. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press; 2009.

45. Shibuya K, Boerma J. Measuring progress towards reducing health inequalities. *Bulletin of the World Health Organization*. 2005;83(3):162.

46. McKie J, Richardson J. *Neglected equity issues in cost effectiveness analysis: Part 1: Severity of Pre-Treatment Condition, Realisation of Potential for Health, Concentration and Dispersion of Health Benefits, and Age-Related Social Preferences*. Victoria: Monash University. Research Paper 7; 2005.

47. Norman R, Hall J, Street D, Viney R. Efficiency and equity: A stated preference approach. *Health Economics*. 2013;22(5):568-581.

48. Frey BS, Benz M, Stutzer A. Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics*. 2004;160(3):377-401.

49. Murray CJL, Knaul F, Musgrove P, Xu K, Kawabata K. *Defining and measuring fairness in financial contribution to the health system*. GPE Discussion Paper Series No 24. 2000.

50. Xu K, Klavus J, Evans DB, Hanvoravongchai P, Zeramdini R, Murray CJL. Understanding household catastrophic health expenditures: A multi-country analysis. In: Murray CJL, Evans DB, eds. *Health system performance assessment: Debates, methods and empiricism*. Geneva: World Health Organization; 2003:566-573.

51. Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD health care quality indicators project. *International Journal for Quality in Health Care*. 2006;18(suppl 1):5.

52. OECD. *The OECD health project: Towards high-performing health systems*. Paris: Organisation for Economic Cooperation and Development; 2004.

53. Murray CJL, Lauer J, Tandon A, Frenk J. *Overall health system achievement for 191 countries*. GPE Discussion Paper Series, No.28; 2000.

54. Gakidou EE, Murray CJL, Evans DB. Quality and equity: Preferences for health system outcomes. In: Murray C.J.L., Evans DB, eds. *Health system performance assessment: Debates, methods and empiricism*. Geneva: World Health Organization; 2003:667-765.

55. Almeida C, Braveman P, Gold MR, et al. Methodological concerns and recommendations on policy consequences of the world health report 2000. *The Lancet*. 2001;357(9269):1692-1697.

56. Williams A. Science or marketing at WHO? A commentary on 'World health 2000'. *Health Economics*. 2001;10(2):93-100.

57. Smith P. Developing composite indicators for assessing overall system efficiency. In: Smith P, ed. *Measuring up: Improving health system performance in OECD countries*. Paris: Organisation for Economic Cooperation and Development; 2002:275-295.

58. Gakidou E, Murray CJL, Frenk J. *Measuring preferences on health system performance assessment. Geneva: World Health Organization*. GPE Discussion Paper Series, No. 20; 2000

59. Richardson J, Wildman J, Robertson IK. A critique of the World Health Organisation's evaluation of health system performance. *Health Economics*. 2003;12(5):355-366.

60. Oswaldo Cruz Foundation. *Report of the workshop "Health systems performance - the World Health Report 2000"*. Available at http://www.who.int/health-systems-performance/docs/articles/oswaldocruz_report.pdf. Accessed December/11, 2011.

61. Lauer JA, Evans DB, Murray CJL. Measuring health system attainment: The impact of variability in the importance of social goals. In: Murray CJL, Evans DB, eds. *Health systems performance assessment: Debates, methods and empiricism*. Geneva: World Health Organization; 2003:677-681.

62. Lauer J, Lovell CAK, Murray C, Evans D. World health system performance revisited: The impact of varying the relative importance of health system goals. *BMC health services research*. 2004;4(1):19.

63. Lancaster KJ. A new approach to consumer theory. *The Journal of Political Economy*. 1966;74(2):132-157.

64. Louviere JJ, Hensher DA, Swait JD. *Stated choice methods: Analysis and applications*. Cambridge: Cambridge University Press; 2000.

65. Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Economics*. 2003;12(7):545-551.

66. Rawls J. *A theory of justice*. Cambridge: Belknap Press of Harvard University Press; 1971.

67. Keeney RL, Raiffa H. *Decisions with multiple objectives: Preferences and value tradeoffs*. 2nd ed. Cambridge: Cambridge University Press; 1993.

68. Kjær T. A review of the discrete choice experiment-with emphasis on its application in health care. *Health Economics Papers* 1; 2005.

69. Ortiz JP, Valentine NB, Gakidou EE, et al. Inequality in responsiveness: Population surveys from 16 OECD countries. In: Murray CJL, Evans DB, eds. *Health systems performance assessment: Debates, methods and empiricism.* Geneva: World Health Organization; 2003:653-663.

70. Otten F, Bos W, Vrooman C, Hoff S. *Poverty notice 2006* [in Dutch: Armoedebericht 2006]. 2006.

71. Street DJ, Burgess L, Louviere JJ. Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*. 2005;22(4):459-470.

72. Williams A, Cookson R. Equity in health. In: Culyer AJ, Newhouse JP, eds. *North Holland handbook of health economics.* Amsterdam: Elsevier; 2000:1863-1909.

73. Salkeld G, Ryan M, Short L. The veil of experience: Do consumers prefer what they know best? *Health Economics*. 2000;9(3):267-270.

74. Ryan M, Amaya-Amaya M. Threats to and hopes for estimating benefits. *Health Economics*. 2005;14(6):609-619.

75. Harris A, Buxton M, O'Brien B, Drummond FRM. Using economic evidence in reimbursement decisions for health technologies: Experience of 4 countries. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2001;1(1):7-12.

76. Drummond M, Jönsson B, Rutten F, Stargardt T. Reimbursement of pharmaceuticals: Reference pricing versus health technology assessment. *The European Journal of Health Economics*. 2011;12(3):263-271.

77. Sorenson C. The role of HTA in coverage and pricing decisions: A cross-country comparison. *Euro Observer*. 2009;11(1):1-11.

78. Stevens A, Milne R. Health technology assessment in England and Wales. *International Journal of Technology Assessment in Health Care*. 2004;20(01):11-24.

79. Niezen M, de Bont A, Stolk E, Eyck A, Niessen L, Stoevelaar H. Conditional reimbursement within the Dutch drug policy. *Health Policy*. 2007;84(1):39-50.

80. Cohen J, Faden L, Predaris S, Young B. Patient access to pharmaceuticals: An international comparison. *The European Journal of Health Economics*. 2007;8(3):253-266.

81. Cohen J, Stolk E, Niezen M. The increasingly complex fourth hurdle for pharmaceuticals. *Pharmacoeconomics*. 2007;25(9):727-734.

82. Levy AR, Mitton C, Johnston KM, Harrigan B, Briggs AH. International comparison of comparative effectiveness research in five jurisdictions: Insights for the US. *Pharmacoeconomics*. 2010;28(10):813-830.

83. Wirtz V, Cribb A, Barber N. Reimbursement decisions in health policy: Extending our understanding of the elements of decision-making. *Health Policy*. 2005;73(3):330-338.

84. Garattini L, Cornago D, De Compadri P. Pricing and reimbursement of in-patent drugs in seven European countries: A comparative analysis. *Health Policy*. 2007;82(3):330-339.

85. Vuorenkoski L, Toiviainen H, Hemminki E. Drug reimbursement in Finland: A case of explicit prioritizing in special categories. *Health Policy*. 2003;66(2):169-177.

86. Anell A. Priority setting for pharmaceuticals. the use of health economic evidence by reimbursement and clinical guidance committees. *The European Journal of Health Economics*. 2004;5(1):28-35.

87. Le Pen C, Priol G, Lilliu H. What criteria for pharmaceuticals reimbursement? *The European Journal of Health Economics*. 2003;4(1):30-36.

88. Stolk EA, Poley MJ. Criteria for determining a basic health services package. recent developments in The Netherlands. *The European Journal of Health Economics*. 2005;6(1):2-7.

89. Hutton J, McGrath C, Frybourg JM, Tremblay M, Bramley-Harker E, Henshall C. Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). *International Journal of Technology Assessment in Health Care*. 2006;22(1):10-18.

90. OECD. *OECD health data 2010*. Available at http://www.irdes.fr/EcoSante/DownLoad/OECD-HealthData_FrequentlyRequestedData.xls. Updated 2010.

91. Klein R. Priorities and rationing: Pragmatism or principles? *BMJ*. 1995;311(7008):761.

92. Kristensen FB, Lampe K, Chase DL, et al. Practical tools and methods for health technology assessment in Europe: Structures, methodologies, and tools developed by the European network for health technology assessment, EUnetHTA. *International Journal of Technology Assessment in Health Care*. 2009;25(S2):1-8.

93. Grandfils N. *Drug price setting and regulation in France*. Working Papers DT 16; 2008. Available at http://www.irdes.fr/EspaceAnglais/Publications/WorkingPapers/DT16DrugPriceSettingRegulationFrance.pdf

94. Van Ganse E, Chamba G, Bruet G, Becquart V, Stam, C., Marty,C., Lopes S. *Pharmaceutical Pricing and Reimbursement Information: France Pharma profile*. 2008.

95. Sermet C, Andrieu V, Godman B, Van Ganse E, Haycox A, Reynier JP. Ongoing pharmaceutical reforms in France: Implications for key stakeholder groups. *Applied Health Economics and Health Policy*. 2010;8(1):7-24.

96. TLV. *Hypertension: A review of medicines for lowering blood pressure*. Solna: TLV. Available at http://www.tlv.se%2Fupload%2Fgenomgangen%2Freview-blood-pressure.pdf&ei=3dcVVLDYEfSg7AaXt4GABQ&usg=AFQjCNGxpHCVkqSVnjT3RY4pZsAHwcwl_A&bvm=bv.75097201,d.ZGU; 2008.

97. Dreyer NA, Garner S. Registries for robust evidence. *The Journal of the American Medical Association*. 2009;302(7):790.

98. Garrison Jr LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real world data for coverage and payment decisions: The ISPOR real world data task force report. *Value in Health*. 2007;10(5):326-335.

99. Thomson S, Foubister T, Figueras J, Kutzin T, Permanand G, Bryndová L. *Addressing financial sustainability in health systems*. Copenhagen: WHO Regional Office for Europe. Available at: http://www.euro.who.int/__data/assets/pdf_file/0005/64949/E93058.pdf?ua=1; 2009.

100. High Level Group on Health Services and Medical Care. *Background document setting out health system objectives of particular interest to different European committees*. 2004;HLG/2004/5. Available at: http://ec.europa.eu/health/ph_overview/co_operation/healthcare/docs/high-level_2004_007_en.pdf.

101. U.S. Department of Health and Human Services. *Strategic plan fiscal year 2010-2015*.

102. Daniels N, Sabin J. Limits to health care: Fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philosophy & Public Affairs*. 1997;26(4):303-350.

103. Franken MG, Le Polain M, Cleemput I, Koopmanschap MA. Similarities and differences between five European drug reimbursement systems. *International Journal of Technology Assessment in Health Care*. 2012;28(4):349-357.

104. Cleemput I, Van Wilder P, Huybrechts M, Vrijens F. Belgian methodological guidelines for pharmacoeconomic evaluations: Toward standardization of drug reimbursement requests. *Value in Health*. 2009;12(4):441-449.

105. Obyn C, Cleemput I. The capital cost and productivity of MRI in a Belgian setting. *JBR-BTR*. 2010;93(2):92-96.

106. Denis A, Mergaert L, Fostier C, Cleemput I, Simoens S. Orphan diseases and orphan medicines: A Belgian and European study. *Journal de pharmacie de Belgique*. 2009;(4)(4):131-137.

107. Jansson S. Implementing accountability for reasonableness–the case of pharmaceutical reimbursement in Sweden. *Health Economics, Policy and Law*. 2007;2(02):153-171.

108. Pasternack I, Anttila H, Mäkelä M, et al. Testing the HTA core model: Experiences from two pilot projects. *International Journal of Technology Assessment in Health Care*. 2009;25(S2):21-27.

109. Daniels N, Teagarden JR, Sabin JE. An ethical template for pharmacy benefits. *Health Affairs*. 2003;22(1):125-137.

110. Annemans L, Arickx F, Belle O, et al. *A call to make valuable innovative medicines available in the European Union: Recommendations for a coordinated action to stimulate, measure, and valorise*

*pharmaceutical innovation*. Available at: https://www.inami.fgov.be/information/all/studies/study-20100923-24/pdf/background_report_en.pdf; 2010.

111. Hasman A, Hope T, Osterdal LP. Health care need: Three interpretations. *Journal of Applied Philosophy*. 2006;23(2):145-156.

112. Cookson R, Dolan P. Principles of justice in health care rationing. *Journal of Medical Ethics*. 2000;26(5):323-329.

113. Culyer AJ. The principle objective of the NHS should be to maximise aggregate health. In: New B, ed. *Rationing: Talk and action in health care*. London: King's Fund and British Medical Journal Publishing Group; 1997.

114. Senn A. Capability and well-being. In: Nussbaum M, Senn A, eds. *The quality of life*. Oxford: Oxford University Press; 1993.

115. Williams A. Intergenerational equity: An exploration of the 'fair innings' argument. *Health Economics*. 1997;6(2):117-132.

116. Nord E. Severity of illness versus expected benefit in societal evaluation of healthcare interventions. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2001;1(1):85-92.

117. Stolk EA, Van Donselaar G, Brouwer WBF, Busschbach JJV. Reconciliation of economic concerns and health policy: Illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics*. 2004;22(17):1097-1107.

118. McKie J, Richardson J. The rule of rescue. *Social Science & Medicine*. 2003;56(12):2407-2419.

119. Van de Wetering E, Stolk E, Van Exel N, Brouwer W. Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *The European Journal of Health Economics*. 2013;14(1):107-115.

120. Zwaap J, Mastenbroek CG, van der Holt B, van der Heijden LA. *Management of the basic benefit package in practice 2* [in Dutch: Pakketbeheer in praktijk 2]. 2009;29033174.

121. McCabe C, Claxton K, Tsuchiya A. Orphan drugs and the NHS: Should we value rarity? *BMJ*. 2005;331(7523):1016-1019.

122. Cleemput I, Neyt M, Thiry N, De Laet C, Leys M. *Threshold values for cost-effectiveness in health care Health Technology Assessment (HTA)*. Brussels: Belgian Health Care Knowledge Centre (KCE); 2008. KCE reports 100C (D/2008/10.273/96)

123. Schotte E. *Een onderzoek naar mogelijke factoren die een impact kunnen hebben op de terugbetalingsbeslissing van geneesmiddelen*. Gent: University of Gent. Available at: http://lib.ugent.be/fulltxt/RUG01/001/392/012/RUG01-001392012_2010_0001_AC.pdf; 2009.

124. LFN. *The pharmaceutical benefits board: Areas of responsibility and tasks*. Available at: http://www.tlv.se/Upload/English/ENG-lfn-responsibility-tasks.pdf; 2002.

125. Denis A, Mergaert L, Fostier C, Cleemput I, Simoens S. A comparative study of European rare disease and orphan drug markets. *Health Policy*. 2010;97(2):173-179.

126. Syrett K. Health technology appraisal and the courts: Accountability for reasonableness and the judicial model of procedural justice. *Health Economics, Policy and Law*. 2006:1-20.

127. Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: A comparative analysis of five countries and implications for coverage decisions. *Health Policy*. 2012;108(2-3):167-177.

128. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics*. 1999;8(1):25-39.

129. Dolan P, Edlin R, Tsuchiya A. *The relative societal value of health gains to different beneficiaries*. 2008; HEDS Discussion Paper 08/12.

130. Cleemput I, Franken M, Le Polain M, Koopmanschap M. European drug reimbursement systems' legitimacy: Five-country comparison and policy tool. *International Journal of Technology Assessment in Health Care*. 2012;28(4):358-366.

131. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: An empirical inquiry into social value. *Health Policy*. 2005;74(3):343-355.

132. Velasco-Garrido M, Schreyögg J, Stargardt T, Busse R. *Identification of health baskets in nine EU countries*. Working Papers 2006;60(2-3):63-88.

133. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: A discrete-choice experiment with assessment of external validity. *Pharmacoeconomics*. 2013;31(4):345-355.

134. Koopmanschap MA, Stolk EA, Koolman AHE. Dear policymaker: Have you made up your mind? A discrete choice experiment. *International Journal of Technology Assessment in Health Care*. 2010;26(2):198-204.

135. Linley WG, Hughes DA. Societal views on NICE, cancer drugs fund and value-based pricing criteria for prioritising medicines: A cross-sectional survey of 4118 adults in Great Britain. *Health Economics*. 2013;22(8):948-964.

136. Bobinac A, Van Exel N, Rutten FF, Brouwer WB. Willingness to pay for a quality-adjusted life-year: The individual perspective. *Value in Health*. 2010;13(8):1046-1055.

137. Lancsar E, Wildman J, Donaldson C, Ryan M, Baker R. Deriving distributional weights for QALYs through discrete choice experiments. *Journal of Health Economics*. 2011;30(2):466-478.

138. Green C. Investigating public preferences on 'severity of health'as a relevant condition for setting healthcare priorities. *Social Science & Medicine*. 2009;68(12):2247-2255.

139. Richardson JRJ, McKie J, Peacock SJ, Iezzi A. Severity as an independent determinant of the social value of a health service. *The European Journal of Health Economics*. 2011;12(2):163-174.

140. Shah KK. Severity of illness and priority setting in healthcare: A review of the literature. *Health Policy*. 2009;93(2):77-84.

141. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: An empirical study of public preferences. *The European Journal of Health Economicsomics*. 2014;15:389-399.

142. Olsen JA. Priority preferences: "end of life" does not matter, but total life does. *Value in Health*. 2013;16(6):1063-1066.

143. Cookson R, McCabe C, Tsuchiya A. Public healthcare resource allocation and the rule of rescue. *Journal of Medical Ethics*. 2008;34(7):540-544.

144. Nord E. Concerns for the worse off: Fair innings versus severity. *Social Science & Medicine*. 2005;60(2):257-263.

145. Nord E. Severity of illness versus expected benefit in societal evaluation of healthcare interventions. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2001;1(1):85-92.

146. Stolk EA, van Donselaar G, Brouwer WB, Busschbach JJ. Reconciliation of economic concerns and health policy. *Pharmacoeconomics*. 2004;22(17):1097-1107.

147. Franken M, Nilsson F, Sandmann F, De Boer A, Koopmanschap M. Unravelling drug reimbursement outcomes: A comparative study of the role of pharmacoeconomic evidence in Dutch and Swedish reimbursement decision-making. *PharmacoEconomics*. 2013;31(9):781-797.

148. Gerkens S, Merkur S. *Belgium: Health system review. Health systems in transition*. 2010;12(5):1-266.

149. De Pouvourville G. A French approach to cost-effectiveness analysis? *The European Journal of Health Economics*. 2010;11(6):521-523.

150. Bouvenot G. *Le rôle de la commission de la transparence de la HAS dans la pertinence de l'admission au remboursement des médicaments par l'assurance maladie*. Regards: protection sociale. 2013;44:59-66.

151. Pharmaceutical Policy and Management Unit. *Global analytical report on the content of pharmanet - singular pathway*. Brussels: National Institute for Health and Disability Insurance. 2009:1-80.

152. Dolan P, Tsuchiya A. Health priorities and public preferences: The relative importance of past health experience and future health prospects. *Journal of Health Economics*. 2005;24(4):703-714.

153. Cookson R. Can the NICE "End-of-life premium" be given a coherent ethical justification? *Journal of Health Politcs, Policy and Law*. 2013;38(6):1129-1148.

154. Kvamme MK, Gyrd-Hansen D, Olsen JA, Kristiansen IS. Increasing marginal utility of small increases in life-expectancy? Results from a population survey. *Journal of Health Economics*. 2010;29(4):541-548.

155. Martin DK, Pater JL, Singer PA. Priority-setting decisions for new cancer drugs: A qualitative case study. *The Lancet*. 2001;358(9294):1676-1681.

156. Harris A, Buxton M, O'Brien B, Drummond M. Using economic evidence in reimbursement decisions for health technologies: Experience of 4 countries. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2001;1(1):7-12.

157. Mason A, Drummond M, Ramsey S, Campbell J, Raisch D. Comparison of anticancer drug coverage decisions in the United States and United Kingdom: Does the evidence support the rhetoric? *Journal of Clinical Oncology*. 2010;28(20):3234-3238.

158. Rosenberg-Yunger ZRS, Daar AS, Thorsteinsdóttir H, Martin DK. Priority setting for orphan drugs: An international comparison. *Health Policy*. 2011;100(1):25-34.

159. Vegter S, Rozenbaum MH, Postema R, Tolley K, Postma MJ. Review of regulatory recommendations for orphan drug submissions in The Netherlands and Scotland: Focus on the underlying pharmacoeconomic evaluations. *Clinical Therapeutics*. 2010;32(9):1651-1661.

160. Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: A comparative analysis of five countries and implications for coverage decisions. *Health Policy*. 2012;108(2-3):167-177.

161. Bending M, Hutton J, McGrath C. A comparison of pharmaceutical reimbursement agencies' processes and methods in France and Scotland. *International Journal of Technology Assessment in Health Care*. 2012;28(02):187-194.

162. Dakin HA, Devlin NJ, Odeyemi IA. "Yes", "no" or "yes, but"? multinomial modelling of NICE decision-making. *Health Policy*. 2006;77(3):352-367.

163. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics*. 2004;13(5):437-452.

164. Harris AH, Hill SR, Chin G, Li JJ, Walkom E. The role of value for money in public insurance coverage decisions for drugs in Australia: A retrospective analysis 1994-2004. *Medical Decision Making*. 2008;28(5):713-722.

165. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, USA; 2005.

166. Van der Meijden C, Grahlmann C. *Assessment procedure for outpatient drugs* [in Dutch: Procedure beoordeling extramurale geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2011;2010117719.

167. Cheung PK. *Description of anti HIV drugs* [in Dutch: Signalement HIV-remmende geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2011;300.

168. Schippers EI. *Proposals for provisional and conditional inclusion of new drugs in the drug reimbursement scheme* [in Dutch: Voorstellen voor voorlopige en voorwaardelijke opname van nieuwe geneesmiddelen in het pakket]. Ministerial Letter. 2011;GMT-U-3065578.

169. Hoomans T, Severens JL, van der Roer N, Delwel GO. Methodological quality of economic evaluations of new pharmaceuticals in The Netherlands. *Pharmacoeconomics*. 2012;30(3):219-227.

170. Rutten F. Economic evaluation and health care decision-making. *Health Policy*. 1996;36(3):215-229.

171. Berg M, van der Grinten T, Klazinga N. Technology assessment, priority setting, and appropriate care in Dutch health care. *International Journal of Technology Assessment in Health Care*. 2004;20(01):35-43.

172. Hoomans T, van der Roer N, Severens JL, Delwel GO. Cost-effectiveness of new drugs impacts reimbursement decision making but room for improvement [in Dutch: Kosteneffectiviteit van

nieuwe geneesmiddelen: Van belang bij geneesmiddelenvergoeding, maar voor verbetering vatbaar]. *Nederlands Tijdschrijft voor Geneeskunde*. 2010;154(A958):1-9.

173. Rutten F, van der Linden J. Integration of economic appraisal and health care policy in a health insurance system; the Dutch case. *Social Science & Medicine*. 1994;38(12):1609-1614.

174. Van Rossum W. Decision-making and medical technology assessment: Three Dutch cases. *Knowledge, Technology & Policy*. 1991;4(1):107-124.

175. Dunning Committee. Committee on choices in health care: Choices in health care. 1992.

176. CVZ. Van zorg verzekerd: Informatie over het college voor zorgverzekeringen. Diemen: College voor Zorgverzekeringen. 2009.

177. Couwenbergh BTLE, van der Meer FM, Weghaus-Reus SE, Schelleman H, Zwaap J. *Management of the basic benefit package in practice part 3* [in Dutch: Pakketbeheer in de praktijk deel 3]. Diemen: College voor Zorgverzekeringen. 2013;2013095825.

178. Ligtenberg G, Staal PC, Goettsch W, Knies S. *Cost-effectiveness in health care* [in Dutch: Kosteneffectiviteit in de zorg]. Diemen: College voor Zorgverzekeringen. 2013;2013082521.

179. Klink A. *Medicines policy October 8, 2009* [in Dutch: Geneesmiddelenbeleid 8 oktober 2009]. Kamerstuk. 2009;GMT-2960544:1-7.

180. Dupree R, Pasman P. *Management of the basic benefit package for specialist drugs* [in Dutch: Pakketbeheer specialistische geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2013;2013077869.

181. CVZ. *Guidelines for pharmacoeconomic evaluations, updated version* [in Dutch: Richtlijnen voor farmaco-economisch onderzoek, geactualiseerde versie]. Diemen: College voor Zorgverzekeringen. 2006;25001605.

182. van der Meijden C. *Reimbursement procedure for outpatient drugs* [in Dutch: Procedure beoordeling extramurale geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2012;2010117719.

183. Rutte M, Samson D. *Coalition report* [in Dutch: Bruggen slaan: Regeerakkoord VVD-PvdA]. 2012. http://www.rijksoverheid.nl/regering/documenten-en-publicaties/rapporten/2012/10/29/regeerakkoord.html.

184. College voor Zorgverzekeringen. *Pay-for-performance agreement for omalizumab (xolair) for treatment of severe asthma* [in Dutch: Pay-for-performance overeenkomst omalizumab (xolair) voor de behandeling van ernstige astma]. Diemen: College voor Zorgverzekeringen. 2012;ZA/2012056679.

185. Neumann PJ, Chambers JD, Simon F, Meckley LM. Risk-sharing arrangements that link payment for drugs to health outcomes are proving hard to implement. *Health Affairs*. 2011;30(12):2329-2337.

186. Sudlow CLM, Counsell CE. Problems with UK government's risk sharing scheme for assessing drugs for multiple sclerosis. *BMJ*. 2003;326(7385):388-392.

187. CVZ. *Pharmacotherapeutic report ipilimumab (yervoy®) for the indication 'inoperable or metastatic melanoma in advanced stage adults* [in Dutch: Farmacotherapeutisch rapport ipilimumab (yervoy®) bij de indicatie 'inoperabel of gemetastaseerd melanoom in een gevorderd stadium bij volwassenen die al eerder behandeld zijn' (ZA/2012010215)]. Diemen: College voor Zorgverzekeringen. 2012.

188. CVZ. *Advice regarding alglucosidase alfa (myozyme ®) for the indication of alglucosidase alfa deficiency (Pompe disease)* [in Dutch: Advies alglucosidase alfa (myozyme ®) bij alglucosidase alfa deficientie (ziekte van Pompe)]. Diemen: College voor Zorgverzekeringen. 2012;ZA/2012134239.

189. Kievit J. *Orphan durgs, and Pompe and Fabry disease: Health care policy with compassion as well as intellect* [in Dutch: Weesgeneemiddelen, en de ziekten van Fabry en Pompe: Over gezondheidszorgbeleid met compassie zowel als verstand]. 2012.

190. *Ministerial point of view Pompe and Fabry* [in Dutch: Standpunt Pompe en Fabry]. 2013;155165-111008-GMT. Accessed October 28, 2013.

191. Commissie BOM. Systematic assessment of oncological treatments [in Dutch: Systematische beoordeling van oncologische middelen]. *Medische Oncologie*. 2012;2012(1):26-29.

192. Borstkankervereniging Nederland. *Report on the treatment of patients with breast cancer with metastatic HER2-positive tumours* [in Dutch: Rapport onderbehandeling van borstkankerpatiënten met uitgezaaide HER2-positieve tumoren]. Available at: www.borstkanker.nl/php/downloaditem.php?id=1248. 2005.

193. Delwel GO. *Guidance for outcomes research*. Diemen: College voor Zorgverzekeringen 2008;270.

194. Jagannath S, Barlogie B, Berenson J, et al. A phase 2 study of two doses of bortezomib in relapsed or refractory myeloma. *British Journal of Haematology*. 2004;127(2):165-172.

195. Richardson PG, Barlogie B, Berenson J, et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. *The New England Journal of Medicine*. 2003;348(26):2609-2617.

196. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *The New England Journal of Medicine*. 2005;352(24):2487-98.

197. Kumar SK, Rajkumar SV, Dispenzieri A, et al. Improved survival in multiple myeloma and the impact of novel therapies. *Blood*. 2008;111(5):2516-2520.

198. Verelst SGR, Blommestein HM, Karim-Kos HE, Huijgens PC, Sonneveld P. Trends in incidence and survival of multiple myeloma in The Netherlands in the last two decades: Results from a national population based study [Abstract ASH annual meeting]. *Blood*. 2011;118(21):118:5071.

199. Sonneveld P, Zweegman S, Vellenga E, et al. Guidelines for treatment of plasma cell dyscrasias in 2010 [in Dutch: Richtlijn behandeling plasmacelaandoeningen anno 2010]. *Nederlands Tijdschrift voor Hematologie*. 2010;7(3):84-95.

200. Groot MT, Huijgens PC, Uyl-de Groot CA. Introduction of expensive pharmaceuticals in haemato-oncology in The Netherlands and throughout Europe. *European Journal of Hospital Pharmacy Practice*. 2006;12:30-36.

201. Wilking N, Jönsson B, Högberg D. *Comparator report on patient access to cancer drugs in Europe*. Stockholm: Karolinska Institutet. 2009.

202. Creswell JW. *Research design: Qualitative, quantitative, and mixed methods approaches*. California, United States of America: SAGE Publications, Incorporated; 2003.

203. Netherlands Cancer Registry. Available at: www.iknl.nl.

204. Franken MG, Gaultney JG, Blommestein HM, et al. *Pilot outcomes research: Effects and costs of bortezomib in relapsed or refractory multiple myeloma*. 2010;2012075715.

205. College voor Zorgverzekeringen. *Evaluation bortezomib (Velcade®)* [in Dutch: Beoordeling bortezomib (Velcade®)]. Diemen: College voor Zorgverzekeringen. 2007;07/30.

206. HOVON. *HOVON 65 MM / GMMG HD4 clinical picture: MM (multiple myeloom) ISRCTN64455289*. Available at http://www.hovon.nl/trials/trials-by-type/mm.html?action=showstudie&studie_id=5&categorie_id=3;.

207. HOVON. *HOVON 86 MM clinical picture: MM (multiple myeloom) EudraCT number:2007-002533-37*. Available at http://www.hovon.nl/trials/trials-by-type/mm.html?action=showstudie&studie_id=60&categorie_id=3;.

208. GIP. *GIP database*. www.gipdatabank.nl.

209. Rogers Everett M. *Diffusion of innovations*. New York: Simon & Schuster Ltd. 1995.

210. Berwick DM. Disseminating innovations in health care. *The Journal of the American Medical Association*. 2003;289(15):1969-1975.

211. Wagelaar M, Lubberding S, Hugtenburg J, Wagner C. *Accessibility of expensive drugs: A study into drug use in breast cancer and multiple myeloma* [in Dutch: Toegankelijkheid van dure geneesmiddelen: Een onderzoek naar geneesmiddelgebruik bij mammacarcinoom en multipel myeloom]. Available at: http://www.onderzoekpatientveiligheid.nl/duregeneesmiddelenrapport.pdf; 2010.

212. Gatsonis C. The promise and realities of comparative effectiveness research. *Statistics in Medicine*. 2010;29(19):1977-1981.

213. Sullivan P, Goldmann D. The promise of comparative effectiveness research. *The Journal of the American Medical Association*. 2011;305(4):400.

214. Lipscomb J, Donaldson MS, Hiatt RA. Cancer outcomes research and the arenas of application. *JNCI Monographs*. 2004;2004(33):1-7.

215. Rubin DB. On the limitations of comparative effectiveness research. *Statistics in Medicine*. 2010;29(19):1991-1995.

216. Tunis SR, Benner J, McClellan M. Comparative effectiveness research: Policy context, methods development and research infrastructure. *Statistics in Medicine*. 2010;29(19):1963-1976.

217. Dreyer NA, Tunis SR, Berger M, Ollendorf D, Mattox P, Gliklich R. Why observational studies should be among the tools used in comparative effectiveness research. *Health Affairs*. 2010;29(10):1818.

218. Chalkidou K, Tunis S, Lopert R, et al. Comparative effectiveness research and evidence based health policy: Experience from four countries. *Milbank Quarterly*. 2009;87(2):339-367.

219. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: Approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: The international society for pharmacoeconomics and outcomes research good research practices for retrospective database analysis. ISPOR Task Force Report—Part II. *Value in Health*. 2009;12(8):1053-1061.

220. Berger M, Mamdani M, Atkins D, Johnson M. Good research practices for comparative effectiveness research: Defining, reporting and interpreting non-randomized studies of treatment effects using secondary data sources. ISPOR Task Force report 2009—Part I. *Value in Health*. 2009;12(8):1044–1052.

221. Kupersmith J, Ommaya AK. The past, present, and future of comparative effectiveness research in the US department of veterans affairs. *American Journal of Medicine*. 2010;123(12):e3-e7.

222. Concato J, Lawler EV, Lew RA, Gaziano JM, Aslan M, Huang GD. Observational methods in comparative effectiveness research. *American Journal of Medicine*. 2010;123(12S):e16-e23.

223. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: Analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: The ISPOR good research practices for retrospective database analysis Task Force Report—Part III. *Value in Health*. 2009;12(8):1062-1073.

224. Psaty BM, Siscovick DS. Minimizing bias due to confounding by indication in comparative effectiveness research. *The Journal of the American Medical Association*. 2010;304(8):897-898.

225. Sorenson C, Naci H, Cylus J, Mossialos E. Evidence of comparative efficacy should have a formal role in European drug approvals. *BMJ*. 2011;343:d4849.

226. Krumholz HM. Real-world imperative of outcomes research. *The Journal of the American Medical Association*. 2011;306(7):754.

227. Tunis SR, Chalkidou K. Coverage with evidence development: A very good beginning, but much to be done. Commentary to Hutton et al. *International Journal of Technology Assessment in Health Care*. 2007;23(04):432-435.

228. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world: Effectiveness versus efficacy studies. *Pharmacoeconomics*. 1999;15(5):423-434.

229. Kuijpers MR, Toenders WGM. *Assessment procedure for inpatient drugs* [in Dutch: Procedure beoordeling intramurale geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2006;26022597.

230. Lokhorst HM, Schmidt-Wolf I, Sonneveld P, et al. Thalidomide in induction treatment increases the very good partial response rate before and after high-dose therapy in previously untreated multiple myeloma. *Haematologica*. 2008;93(1):124-127.

231. Van Gils CWM, Koopman M, Mol L, Redekop WK, Uyl-de Groot CA, Punt CJA. Adjuvant chemotherapy in stage III colon cancer: Guideline implementation, patterns of use and outcomes in daily practice in The Netherlands. *Acta Oncologica*. 2011(0):1-8.

232. Franken M, Gaultney J, Huijgens P, Redekop W, Uyl-de Groot C. CN7 Treatment variation complicates real-world pharmacoeconomics: Daily clinical practice of bortezomib in relapsed or refractory multiple myeloma [Abstract]. *Value in Health*. 2009;12(7):A225-A225.

233. Gaultney J, Franken M, van Gils C, Huijgens P, Redekop W, Uyl-de Groot C. PCN18 Outcomes research of bortezomib indicated for multiple myeloma in the context of the Dutch reimbursement policy for expensive medicines: Threats to the internal validity of the incremental effectiveness estimate [Abstract]. *Value in Health*. 2009;12(7):A259-A259.

234. Byrne BJ, Kishnani PS, Case L, et al. Pompe disease: Design, methodology, and early findings from the Pompe registry. *Molecular Genetics and Metabolism*. 2011.

235. Towse A, Garrison LP. Cant get no satisfaction? will pay for performance help? Toward an economic framework for understanding performance-based risk-sharing agreements for innovative medical products. *Pharmacoeconomics*. 2010;28(2):93-102.

236. Franken MG, van Gils CWM, Gaultney JG, et al. Practical feasibility of outcomes research in oncology: Lessons learned in assessing drug use and cost-effectiveness in The Netherlands. *European Journal of Cancer*. 2013;49(1):8-16.

237. Commissie Farmaceutische Hulp. Farmacotherapeutisch rapport bortezomib (Velcade®). Diemen: College voor Zorgverzekeringen. 2005;25060777.

238. Gaultney JG, Franken MG, Tan SS, et al. Real-world healthcare costs of relapsed/refractory multiple myeloma during the era of novel cancer agents. *Journal of Clinical Pharmacy and Therapeutics*. 2013;38(1):41-47.

239. Richardson PG, Sonneveld P, Schuster M, et al. Extended follow-up of a phase 3 trial in relapsed multiple myeloma: Final time-to-event results of the APEX trial. *Blood*. 2007;110(10):3557-60.

240. Pharmacoeconomic guidelines around the world. Available at: http://www.ispor.org/peguidelines/index.asp. Updated 2013.

241. Vogl DT, Stadtmauer EA, Richardson PG, et al. Impact of prior therapies on the relative efficacy of bortezomib compared with dexamethasone in patients with relapsed/refractory multiple myeloma. *British Journal of Haematology*. 2009;147(4):531-534.

242. Lokhorst HM, van der Holt B, Zweegman S, et al. A randomized phase III study on the effect of thalidomide combined with adriamycin, dexamethasone (TAD), and high-dose melphalan, followed by thalidomide maintenance in patients with multiple myeloma. *Blood*. 2010;115(6):1113-1120.

243. Garrison L, Towse A, Brigss A, de Pouvourville G, Grueger J, Mohr P, Severens J, Siviero P, Sleeper M. Performance-based risk-sharing arrangements: Good practices for design, implementation and evaluation. Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force. *Value in Health*. 2013;16(5):703-719.

244. Gaultney JG, Franken MG, Redekop WK, Huijgens PC, Uyl-de Groot CA, Sonneveld P. PRM3: Turning the tables to address the real value of real-world observational studies of novel anti-cancer agents in multiple myeloma [Abstract]. *Value in Health*. 2012;15(7):A460.

245. Connors A, Speroff T, Dawson NV, et al. The effectiveness of right heart catheterization in the initial care of critically ill patients. *The Journal of the American Medical Association*. 1996;276(11):889-897.

246. Earle CC, Tsai JS, Gelber RD, Weinstein MC, Neumann PJ, Weeks JC. Effectiveness of chemotherapy for advanced lung cancer in the elderly: Instrumental variable and propensity analysis. *Journal of Clinical Oncology*. 2001;19(4):1064-1070.

247. Röttger J, Scheller-Kreinsen D, Busse R. Patient-level hospital costs and length of stay after conventional versus minimally invasive total hip replacement: A propensity-matched analysis. *Value in Health*. 2012;15(8):999-1004.

248. Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. *American Journal of Cardiology*. 2003;92(12):1447-1451.

249. Weintraub WS, Grau-Sepulveda MV, Weiss JM, et al. Comparative effectiveness of revascularization strategies. The *New England Journal of Medicine*. 2012;366(16):1467-1476.

250. Weitzen S, Lapane KL, Toledano AY, Hume AL, Mor V. Principles for modeling propensity scores in medical research: A systematic literature review. *Pharmacoepidemiology and Drug Safety*. 2004;13(12):841-853.

251. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis A report of the ISPOR-SMDM modeling good research practices task force working Group–6. *Medical Decision Making*. 2012;32(5):722-732.

252. Ghatnekar O, Alvegård T, Conradi N, et al. Direct hospital resource utilization and costs of treating patients with multiple myeloma in southwest Sweden: A 5-year retrospective analysis. *Clinical Therapeutics*. 2008;30(9):1704-1713.

253. Armoiry X, Fagnani F, Benboubker L, et al. Management of relapsed or refractory multiple myeloma in French hospitals and estimation of associated direct costs: A multi-centre retrospective cohort study. *Journal of Clinical Pharmacy and Therapeutics*. 2011;36(1):19-26.

254. Koleva D, Cortelazzo S, Toldo C, Garattini L. Healthcare costs of multiple myeloma: An Italian study. *European Journal of Cancer Care*. 2011;20(3):330-336.

255. Mehta J, Duff SB, Gupta S. Cost effectiveness of bortezomib in the treatment of advanced multiple myeloma. *Managed Care Interface*. 2004;17(9):52-61.

256. Hornberger J, Rickert J, Dhawan R, Liwing J, Aschan J, Löthgren M. The cost-effectiveness of bortezomib in relapsed/refractory multiple myeloma: Swedish perspective. *European Journal of Haematology*. 2010;85(6):484-491.

257. Möller J, Nicklasson L, Murthy A. Cost-effectiveness of novel relapsed-refractory multiple myeloma therapies in Norway: Lenalidomide plus dexamethasone vs bortezomib. *Journal of Medical Economics*. 2011;14(6):690-697.

258. De Groot S, Blommestein H, Franken M, Uyl-de Groot C. PCN154 Are population-based registries a suitable tool for outcomes research in cancer? Experiences from four registries [Abstract]. *Value in Health*. 2012;15(7):A437-A438.

259. Brenner H, Gondos A, Pulte D. Recent major improvement in long-term survival of younger patients with multiple myeloma. *Blood*. 2008;111(5):2521-2526.

260. Light DW, Warburton R. Demythologizing the high costs of pharmaceutical research. *BioSocieties*. 2011;6(1):34-50.

261. Kantarjian H. The price of drugs for chronic myelogenous leukemia (CML) is a reflection of unsustainable prices of cancer drugs: From the perspective of a large group of CML experts. *Blood*. 2013;121:4439-4442.

262. National Institute for Health and Clinical Excellence. *Bortezomib monotherapy for relapsed multiple myeloma*. 2007;NICE Technology Appraisal Guidance 129. Accessed 27-07-2012.

263. McCabe CJ, Stafinski T, Edlin R, Menon D. Access with evidence development schemes: A framework for description and evaluation. *Pharmacoeconomics*. 2010;28(2):143-152.

264. Williamson S. A. *Report into the uptake of patient access schemes in the NHS*. 2009. Available at: http://www.bopawebsite.org/contentimages/publications/Report_into_Uptake_of_Patient_Access_Schemes_8_11_09.pdf.

265. Daniels N, Sabin JE. What are fairness and consistency in a national pharmacy benefit? *Medical Care*. 2001;39(4):312-314.

266. Martin DK, Giacomini M, Singer PA. Fairness, accountability for reasonableness, and the views of priority setting decision-makers. *Health Policy*. 2002;61(3):279-290.

267. Belgisch raadgevend comité voor bio-ethiek. *Advice regarding fiancial issues of expensive drugs* [in Dutch: Advies over de problematiek van financiering van dure geneesmiddelen]. 2014;Advise number 58.

268. Friedman A. Beyond accountability for reasonableness. *Bioethics*. 2008;22(2):101-112.

269. Schokkaert E. *Tension between innovation and solidarity* [in Dutch: De spanning tussen innovatie en solidariteit]. 2012:1. Available at: http://icuro.be/documents/content/ICURO-Symposium19-10-2012-VM-2ErikSCHOKKAERT[Compatibiliteitsmodus].pdf. Accessed May 30, 2014.

270. Drummond MF, Sculpher MJ, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, USA; 2005.

271. Klein R. *The new politics of the NHS: From creation to reinvention*. 6th ed. Oxford: Radcliffe Publishing; 2010.

272. Tax SEM, van der Hoeven JJM. *Accessibility to expensive cancer drugs: Now and in the future* [in Dutch: Toegankelijkheid van dure kankergeneesmiddelen: Nu en in de toekomst]. Available at:

http://www.kwf.nl/sitecollectiondocuments/sck%20rapport%20toegankelijkheid%20van%20dure%20kankergeneesmiddelen.pdf; 2014.

273. Raftery J. Health economic evaluation in England. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*. 2014. [In Press]

274. Raftery J. Multiple sclerosis risk sharing scheme: A costly failure. *BMJ*. 2010;340:c1672.

275. NICE. *Appraising life-extending, end of life treatments: Supplementary advice to the appraisal committees*. 2009. Available at http://www.nice.org.uk/guidance/gid-tag387/resources/appraising-life-extending-end-of-life-treatments-paper2.

276. College voor Zorgverzekeringen. *Entering the discussion about cost-effectiveness* [in Dutch: In gesprek over kosteneffectiviteit]. Diemen: College voor Zorgverzekeringen. 2013.

277. Noorani HZ, Husereau DR, Boudreau R, et al. Priority setting for health technology assessments: A systematic review of current practical approaches. *International Journal of Technology Assessment in Health Care*. 2007;23(3):310-315.

278. Devlin NJ, Sussex J. *Incorporating multiple criteria in HTA: Methods and processes*. London: Office of Health Economics. 2011.

279. Baltussen R, Youngkong S, Paolucci F, Niessen L. Multi-criteria decision analysis to prioritize health interventions: Capitalizing on first experiences. *Health Policy*. 2010;96(3):262-264.

280. Thokala P. *Multiple criteria decision analysis for health technology assessment*. Sheffield: School of Health and Related Research, University of Sheffield. 2011.

281. Tony M, Wagner M, Khoury H, et al. Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): Field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Services Research*. 2011;11(1):329.

282. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): Applying the EVIDEM framework to medicines appraisal. *Medical Decision Making*. 2011.

Part of: TI PHARMA

project escher