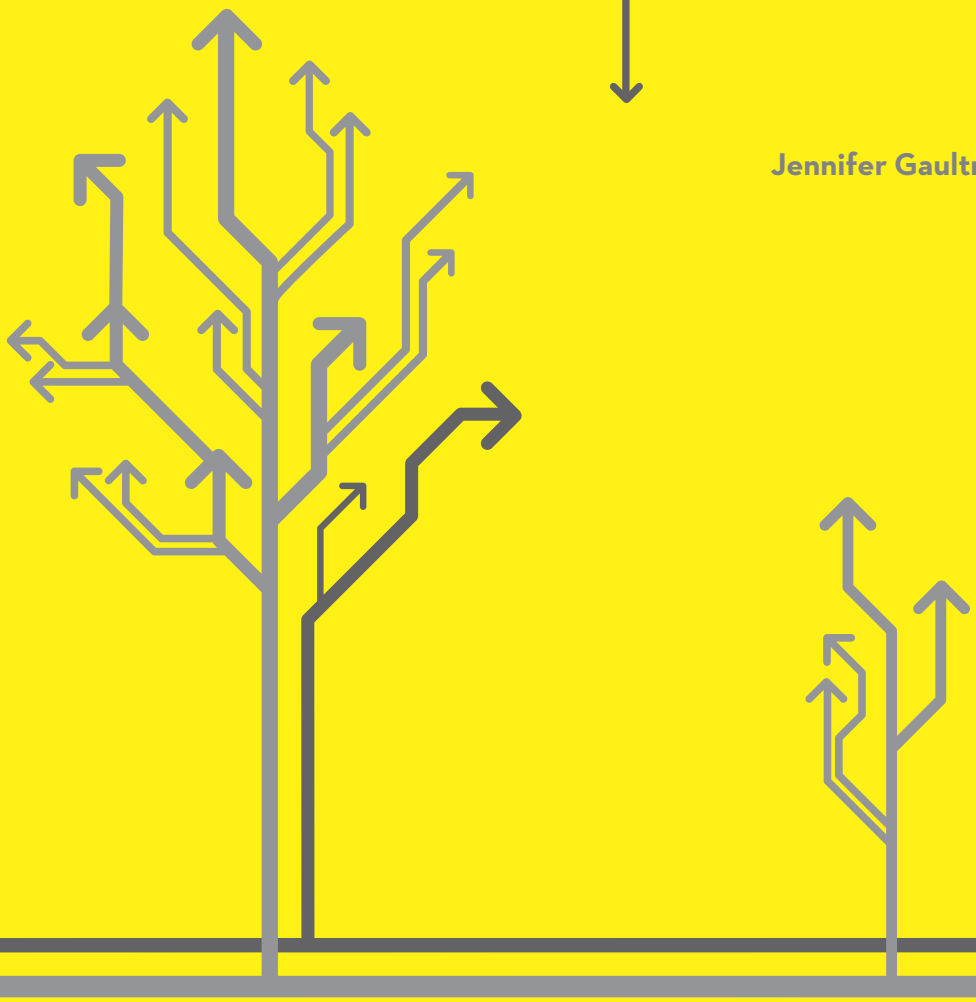
# Economic evaluations

of targeted therapy and risk-stratified treatment approaches

# in multiple myeloma

Jennifer Gaultney



# ECONOMIC EVALUATIONS OF TARGETED THERAPY AND RISK-STRATIFIED TREATMENT APPROACHES IN MULTIPLE MYELOMA

**Jennifer Gaultney**


**Funding**
The studies performed in this thesis were supported by the Dutch Healthcare Insurance Board (CVZ), the Center for Translational Molecular Medicine (CTMM), project BioCHIP (grant 03O-102), an unrestricted grant from PamGene B.V., and the Netherlands Organization for Health Research and Development (ZonMW).

Publication of this thesis was financially supported by PamGene B.V. and SkylineDx B.V.

Cover: © germina – Fotolia.com
Lay-out and printing: Optima Grafische Communicatie B.V., Rotterdam

ISBN/EAN: 978-94-6169-464-5

2013 © J.G. Gaultney
All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without prior written permission of the author or, when appropriate, of the publishers of the publications.

**ECONOMIC EVALUATIONS OF TARGETED THERAPY AND RISK-STRATIFIED TREATMENT APPROACHES IN MULTIPLE MYELOMA**

**Economische evaluaties van doelgerichte therapie en risico-gestratificeerde behandeling bij de ziekte van Kahler**

**Thesis**

to obtain the degree of Doctor from the Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. H.A.P. Pols
and in accordance with the decision of the Doctorate Board.

The public defense shall be held on
Friday January 17, 2014 at 11.30 hours

by

**Jennifer Gail Gaultney**
born in Atlanta, Georgia, USA

**DOCTORAL COMMITTEE**

| | |
|---|---|
| **Promoters:** | Prof.dr. C.A. Uyl-de Groot |
| | Prof.dr. P. Sonneveld |
| **Other members:** | Prof.dr. M.P.M.H. Rutten-van Mölken |
| | Prof.dr. P.C. Huijgens |
| | Prof.dr. M.J. IJzerman |
| **Copromoter:** | Dr. W.K. Redekop |

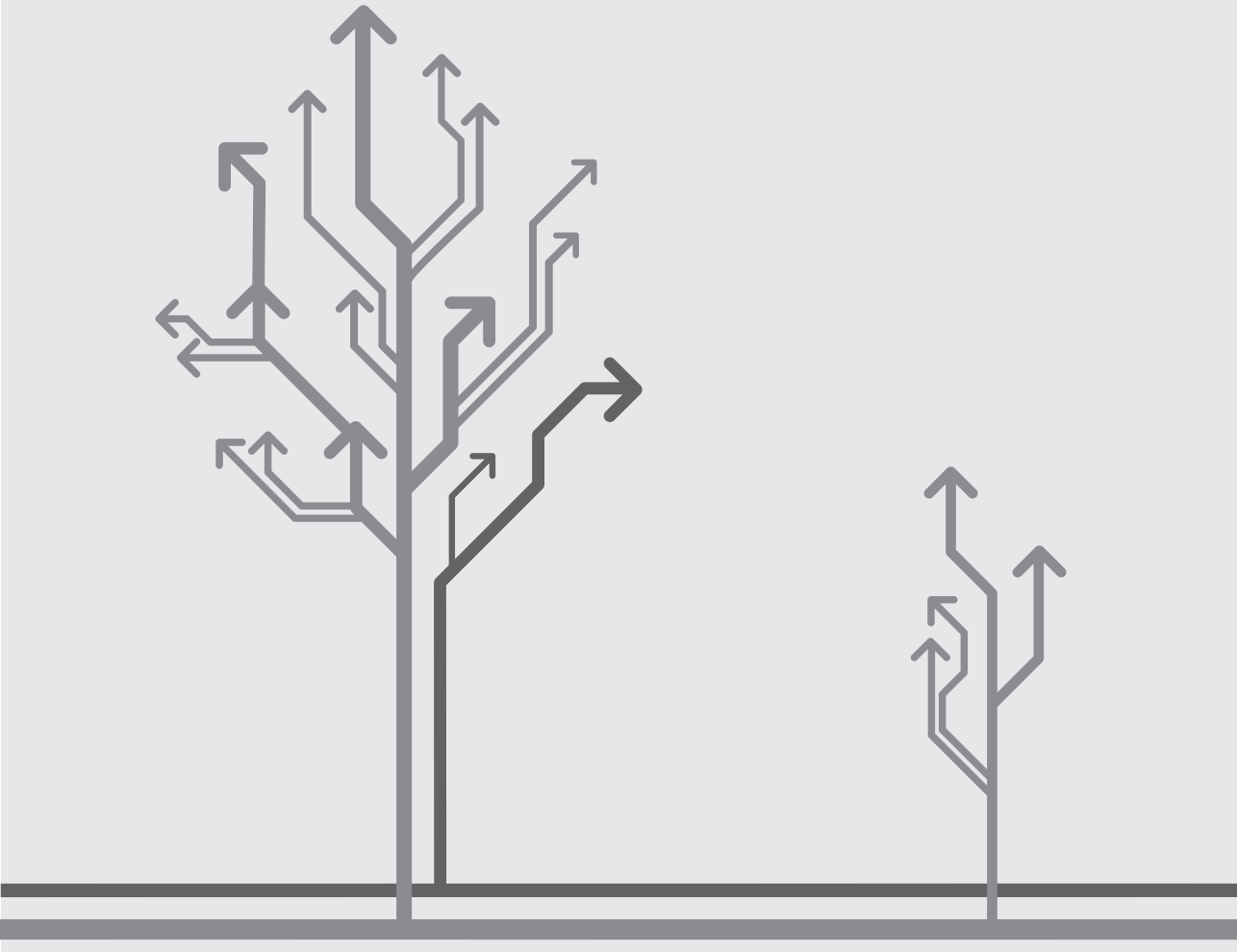
*For Mom.*

**CONTENTS**

**Part 3 – Economic evaluations of risk-stratified treatment approaches**

# Chapter 1.

## Introduction



## MULTIPLE MYELOMA

### Epidemiology, clinical characteristics, and prognosis

Multiple myeloma (MM) is a malignant plasma cell disorder accounting for 1% of all cancer diagnoses worldwide and 13% of all hematologic malignancies [1]. Worldwide, the incidence of MM is 0.4 to 5 per 100,000 people per year [2]. Incidence rates are higher among males than females, people of African descent, and increase rapidly until age 84 and then decline. In the Netherlands, the annual incidence rate is 5 per 100,000 with a median age of 70 at diagnosis [3].

The clinical characteristics of MM vary from asymptomatic patients to those with malignant disease. Nonetheless, the disease has a very characteristic presentation such as lytic bone disease, renal insufficiency, anemia, hypercalcemia, the presence of M-proteins found in the serum and/or urine and immunodeficiency [4]. Disorders of the central and peripheral nervous system are also common [5]. The development of hemorrhagic diathesis or thrombosis is also a risk for MM patients with bleeding present in 15-30% of patients and a 3% risk of thrombosis [6].

MM is considered a severe disease in terms of quality-of-life impact and life expectancy and remains incurable. The survival of multiple myeloma patients has improved substantially in the past decade and is attributable to the introduction of novel ('targeted') therapies bortezomib (Velcade®), thalidomide (Thalidomid®), and lenalidomide (Revlimid ®), as well as improvements in the use of autologous stem cell transplantation and supportive therapy [7, 8]. Newly diagnosed patients can now expect to live an average of 5-7 years, with some patients living longer than 10 years [9, 10]. In fact, many patients are dying with and not necessarily due to the underlying disease.

Prognosis has shown to be associated with treatment but an independent relationship exists with a number of clinical factors and markers for tumor burden and biology. The first staging system to predict prognosis for MM patients was devised 25 years ago by Durie and Salmon (DS) and is based on measurement of hemoglobin, paraprotein concentration and renal and bone disease [11]. Due to advances in treatment that produce complete remissions (CR) and the complexity of the application of the DS staging system, a new International Staging System (ISS) was developed. By applying regression techniques on data from 17 institutions worldwide collected for 11,171 patients [12], the ISS is based on two known markers of disease activity in MM, albumin and $\beta_2$-microglobulin, which were found to be significant predictors of survival. The ISS divides patients into 3 distinct prognostic groups on the basis of these two markers, which can be evaluated at virtually any clinical laboratory (Table 1). It is important to note that the DS and ISS

staging system were studied in patients treated prior to the advent of new active agents. Thus, additional studies validating such prognostic factors are necessary in the present era of novel therapy.

**Table 1.** International Staging System for Myeloma

| Stage | Definition | % of patients | Median survival (months) |
|---|---|---|---|
| I | Serum albumin ≥ 3.5 g/dL and serum $\beta_2$-microglobulin < 3.5 µg/mL | 29 | 62 |
| II | Neither stage I or III | 38 | 44 |
| III | Serum β2-microglobulin ≥ 5.5 µg/mL | 34 | 29 |

Recent studies have found addition clinical factors and molecular markers for acquired chromosomal aberrations to have independent prognostic significance (Figure 1). Clinical factors shown to have an indication of prognosis include elevated lactate dehydrogenase (LDH) or presence of renal failure detected by elevated creatinine levels at diagnosis [13]. Approximately 25% of MM patients define a high-risk population that does not benefit from conventional treatment to justify the morbidity and cost of the procedure and should be steered toward more investigational therapies [14]. These patients are defined by having any one of the following factors: 1) detection of either t(4;14) or t(14;16), 2) deletion of 17p13, 3) deletion of chromosome 13, 4) or aneuploidy by molecular genetics or 5) a plasma cell labeling index (PCLI) greater than 3%. Other molecular markers reported to be high-risk factors include amplifications of the chromosome 1q21 region [15]. Standard-risk disease is defined as the absence of any adverse prognostic marker, the presence of t(11;14) or hyperdiploidy [16, 17]. An ideal prognostic system would be one that associates molecular markers and other clinical parameters.

Despite such progress, challenges remain in prognostic prediction in MM [18]. For example, further validation of the prognostic value of chromosomal abnormalities in independent cohorts of patients is needed, especially those treated with novel therapies, as well as an extension of the prognostic accuracy of newer methods such as microarray technology. To tackle such challenges, it is necessary to establish the use of the currently available techniques for detecting chromosomal aberrations, particularly florescence in situ hybridization (FISH), as part of good clinical practice for MM.

### Current treatment approach

Newly diagnosed patients with symptomatic (active) disease should be treated immediately, whereas asymptomatic (smoldering) myeloma requires only clinical observation, since early treatment with conventional chemotherapy has shown no benefit [19]. The prevailing method that currently dictates the treatment approach towards newly diag-

nosed patients is a uniform treatment approach. With uniform treatment, all patients within a stage group are treated irrespective of their prognostic risk in a similar manner. In general, the treatment paradigm is characterized by two distinctive phases: 1) treatment of newly diagnosed disease, (i.e., induction or initial therapy, or first-line); and 2) treatment of relapsed or refractory disease (i.e., second-line, third-line, fourth-line, etc., and palliative care) (Figure 2). The protocol for upfront therapy depends on eligibility for hematologic stem cell transplant (HSCT). Factors associated with ineligibility for transplantation include older age, poor physical condition and coexisting conditions. Since MM is incurable, the majority of patients will relapse or become refractory to upfront therapy while others will either die before becoming eligible or be deemed unfit for second-line therapy.

***Induction therapy in transplant eligible patients***

Initial therapy consists of three to six cycles of high dose therapy (HDT) followed by HSCT. HDT typically consists of a two-drug combination consisting of dexamethasone and one or more novel agent (i.e., bortezomib, thalidomide, and lenalidomide) [20-22]. More recently, three and four-drug combinations including one or more novel agents have been introduced [1].

HSCT can be either autologous or allogeneic and is administered as either a single or double (i.e., tandem) transplantation. Allogeneic transplantation however is typically less frequently administered on account of the high-risk for death and complication and is usually given within a trial setting.

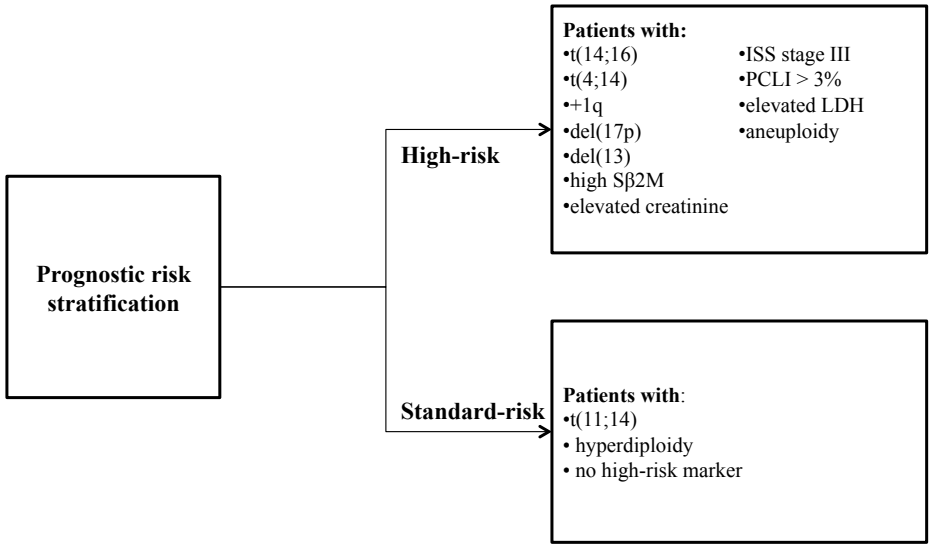



**Figure 1.** Prognostic risk stratification in multiple myeloma.

***Induction therapy in transplant ineligible patients***

Combination therapy with melphalan and prednisone plus either thalidomide [23] or bortezomib [24, 25] is now considered the standard of initial care for patients not eligible for transplantation. The combination of lenalidomide and low-dose dexamethasone has also been shown to be effective as initial therapy in this patient group [26].

***Consolidation and maintenance therapy***

To increase the likelihood of response to induction therapy, consolidation therapy may be given and usually consists of two to four cycles of combination therapies. To prolong response, maintenance therapy may be given and typically consists of continuous therapy with single agents following induction or consolidation until the time of disease progression. Consolidation and maintenance are widely accepted although there are no specific guidelines for their use on account of the conflicting evidence regarding their benefit [1].

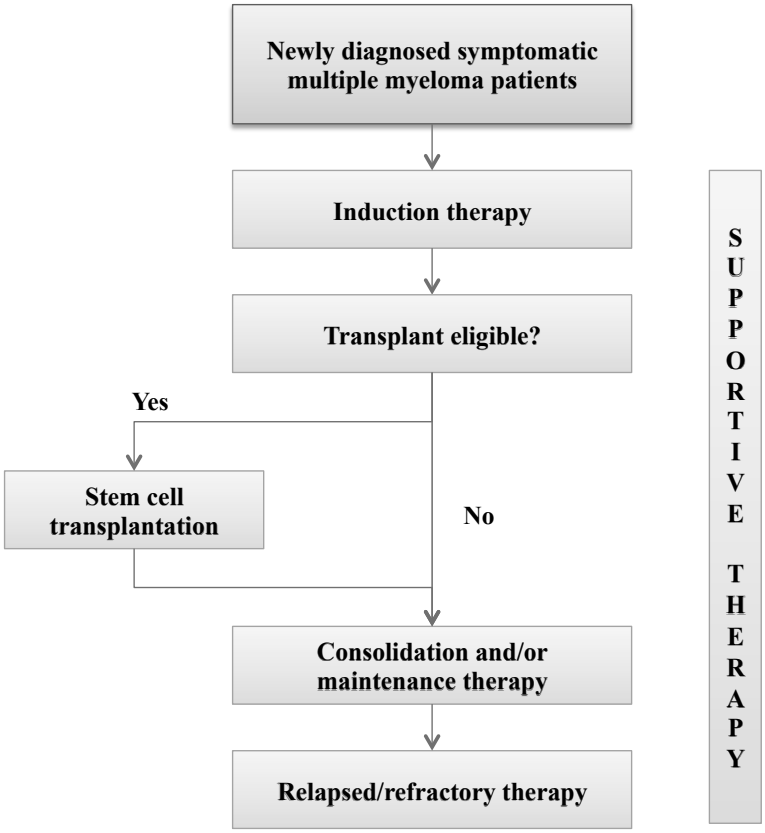



**Figure 2.** Treatment paradigm of multiple myeloma.

***Relapsed/refractory treatment***

The majority of patients eventually relapses or becomes refractory to upfront therapy. The treatment combination administered at first relapse typically depends on the patient's response to initial therapy [27]. A complete response to therapy warrants re-treatment with the same therapy combination. Patients may also be retreated if relapse occurs after 2 years of retreatment or if relapse occurs after 1 year of remission following treatment for first relapse. Otherwise, a different therapy from that administered for induction therapy is given. The treatment of choice for these patients is combination therapy consisting of dexamethasone and bortezomib [28] or lenalidomide [29, 30]. Autologous HSCT is also an option for patients who did not undergo transplantation previously or for those who received transplantation and had a prolonged duration or remission [31].

***Supportive therapy***

Supportive therapy consists mainly of the following: erythropoiesis-stimulating agents to prevent anemia, treatment of bone and neurological pain, therapy to prevent bone lesions and fractures, and prophylaxis for infection [1].

**Risk-stratified treatment**

The term 'risk-stratified treatment' refers to the matching of a treatment regimen to a patient who is demonstrated to be similar to a cohort of patients that exhibit a particular treatment response based on the presence or absence of a clinical marker [32]. As summarized by Trusheim et al. (2007), a clinical marker "can be based on gene-expression profiles, individual proteins, proteomic patterns, metabonomics, histology, imaging, physician's clinical observations and even self-reported patient surveys." In other words, it is not defined by its technology but rather by its reliable, predictive correlation to differential patient responses [32]. Trusheim et al. (2007) further states that the necessary conditions for stratified medicine include either one of the following: (i) underlying disease variability reflecting multifactorial etiology, or currently indistinguishable clinical presentations for biologically distinct conditions, (ii) multiple relevant targets for medical intervention, (iii) differential ADME (absorption, distribution, metabolism and excretion), or (iv) adaptiveness of the disease leading to treatment resistance. In addition, both of the following should be present: (v) multiple treatment options with heterogeneous responses for the disease and (vi) a logistically and medically acceptable clinical marker [32]. The disease multiple myeloma satisfies these conditions as there are (ii) multiple relevant targets for medical intervention and (v) multiple treatment options with heterogeneous responses that can be distinguished by means of a number of (vi) logistically and medically acceptable clinical markers.

***Risk-stratified treatment approaches in multiple myeloma***

On account of the recent advances in understanding the prognostic significance of biomarkers in MM and the development of molecular diagnostic techniques, risk-stratified treatment is gaining more and more support. In a risk-stratified treatment approach, patients are stratified to a more personalized therapy option based on their status for risk markers demonstrated to be associated with improved health outcomes when given a specific therapy. As a result, patients who would otherwise gain little or no health benefits from a particular therapy are stratified to either an alternative treatment option demonstrated to be beneficial for patients with their risk profile or experimental therapy when no such alternative exists. By improving care for particular risk groups, such a strategy offers the potential to improve the overall quality and effectiveness of care.

A risk-stratified treatment approach for multiple myeloma stratifies patients into either of two groups: standard and high-risk. Standard-risk patients have a median overall survival of 6 to 7 years while high-risk patients are defined as having a median expected survival of less than 2 to 3 years despite autologous HSCT [33]. In standard-risk patients, the strategy is disease control with sequential use of drugs to preserve quality-of-life along with prolongation of survival. In high-risk patients, a more aggressive approach with combination therapies is used to achieve and maintain a complete response to therapy since recent data demonstrate this to be an important requirement for long-term survival in this patient population [34].

Arising from the heterogeneity in patient prognosis, there are a number of risk-stratified treatment approaches that have been proposed. The treatment of high-risk patients could be altered from that of standard-risk patients by means of the following: (i) addition of bortezomib to non-bortezomib combinations, (ii) a more intensive dosage schedule for bortezomib (i.e., higher starting dose or additional cycles), (iii) no thalidomide maintenance, or (iv) the addition of allogeneic HSCT. Early studies of bortezomib have found evidence for (i) when stratifying the treatment effect according to cytogenetic markers [24, 25] , which were later confirmed in independent studies [35-39]. . In strategy (iii), the treatment of standard-risk patients would be altered by the addition of thalidomide maintenance, which was shown to improve PFS with a late OS benefit in patients with no adverse cytogenetic marker, while high-risk patients would not receive thalidomide maintenance since it showed no significant PFS and worse OS [40]. For options (ii) and (iv), no data is currently available. Whether these risk-stratified approaches in multiple myeloma lead to clinically important improvements in patient prognosis still requires validation in prospective studies.

**Economic burden of multiple myeloma**

Like most indications in oncology, the costs of care for multiple myeloma are increasing, presumably due to advances in therapy and standard use of HSCT. The ageing population is also a catalyst for increasing its overall budget impact in the future.

The cost of novel agents and HSCT are substantially higher compared to the conventional chemotherapies that were used to treat multiple myeloma prior to the turn of the century. Based on the most recent published estimates of the average wholesale prices in the USA, the cost of one 4-week cycle of thalidomide and lenalidomide is US$7,362 (100-mg capsules, quantity 28) or $88,3344 per year (12 cycles) and $11,447 (25 mg capsules, quantity 21) or $137,364 per year (12 cycles), respectively [41]. Furthermore, the cost of one 21-day cycle of bortezomib was estimated to be $6,450 (3.5-mg vial, quantity four) or $70,950 per patient (11 cycles). These costs do vary per regimen depending on the dose per cycle and number of cycles administered. Prices also vary further across countries as a result of variation in pricing agreements and the availability of generics. In the Netherlands, for example, thalidomide is generically available resulting in a lower cost compared to the US, with the cost of one 4-week cycle equal to €196 (100-mg capsules, quantity 28) or €2,352 per year (12 cycles) [42]. Lenalidomide and bortezomib are not available as generics resulting in similarly high prices for these therapies in the Netherlands compared to the US. The costs of one 4-week cycle of lenalidomide correspond to €5,390 (25 mg capsules, quantity 21) or €64,678 per year (12 cycles), while for bortezomib one 21-day cycle costs €3,816 (3.5-mg vial, quantity four) or €41,976 per patient (11 cycles) [42]. The costs of HSCT are also high, with the median cost per patient for autologous and allogeneic HSCT in 1997 US dollars being $55,500 and $105,300, respectively[43], compared to €45,670 and €101,919 in the Netherlands, respectively [44]. Further adding to the number of therapy combinations and possibly increasing costs, it is expected that second- and third-generation proteasome inhibitors and immunomodulatory agents will be available in the near future [45]. Taking this into consideration, payers of health care are questioning whether the increased economic burden of treatment options for multiple myeloma is justified by the health gain they produce.

In the case of novel therapies, a substantially higher reimbursement price has indeed created concerns with value for money and access to care. Some payers have introduced reimbursement schemes to manage uncertainty about value and patient access to expensive therapies. In multiple myeloma, such schemes have been introduced for the drug bortezomib. In the UK, a risk-sharing scheme was introduced after NICE found that bortezomib was not considered to offer value for money in relapsed/refractory patients[46]. The agreement, which was put forward by the manufacturer, states that the NHS will reimburse the manufacturer for the costs of the drug only for patients who

respond after 4 cycles. In the meantime, the clinical experience with bortezomib in daily practice will generate outcomes research data to be used to further assess the 'real-world' costs and health benefits of the drug in the UK. In the Netherlands, a reimbursement scheme requiring outcomes research within four years of access in daily practice was also introduced to assess the therapeutic value, appropriate use (e.g., patient characteristics, types of treatments, dosages, and dose modifications), costs, and cost-effectiveness of expensive inpatient drugs[1][47], with the drug bortezomib qualifying for such a scheme. Based on the results of outcomes research, the decision for reimbursement is revisited after four[2] years. Although both the UK and the Netherlands have introduced reimbursement schemes requiring outcomes research to clarify the uncertainty around the costs and health benefits of bortezomib, it has never been addressed whether observational data collected during the drug's diffusion phase is useful in this regard. To address this issue, the Dutch healthcare insurance board (i.e., College voor Zorgverzekering or CVZ) requested that a pilot outcomes research study be performed for two drugs, one of which included bortezomib in the indication of relapsed/refractory MM.

Besides reimbursement schemes, implementation of risk-stratified treatment approaches in daily practice could improve not only health outcomes but also the efficiency of care. By making upfront treatment decisions based on prognosis, patients with a favorable prognosis who stand to gain little from a particular treatment protocol would be spared unnecessary and often expensive therapy. On the other hand, risk groups that benefit equally from two or more therapy options that differ primarily in costs can be stratified toward the least costly option. There remains a need to address not only the potential therapeutic value of risk-stratified treatment options in multiple myeloma but also their potential cost-effectiveness. By assessing their potential therapeutic and economic impact, useful evidence can be generated that motivates stakeholders to support future studies evaluating their clinical utility.

## THE ECONOMIC EVALUATION OF HEALTH CARE

Health technology assessment (HTA) is a useful tool for addressing the "economic impact and efficiency of health technology in terms of its short- and long-term health benefits and resource use" [48]. HTA is multidisciplinary by nature in that it focuses on the medi-

1. Expensive was defined by having a projected budget impact equal to or higher than 0.5% of the total hospital budget for inpatient drugs at the macro level.

2. The initial timeframe was three years but has since been changed to four years.

cal, organizational, economic, ethical and social consequences of health technology in order to meet the needs of a variety of policymakers seeking information throughout the lifecycle of a technology.

An economic evaluation is a commonly applied method in HTA to address questions concerning the economic impact of health technologies. Economic evaluations are generally defined as the "comparative analysis of alternative courses of action in terms of their costs and consequences [49]." The steps taken in an economic evaluation include the identification, measurement, valuation and comparison of all costs and consequences of the alternatives under consideration. Depending on how the consequences are measured, economic evaluations can be classified as cost-effectiveness analysis (CEA), cost-utility analysis (CUA), or cost-benefit analysis, with practical application mainly limited to CEA and CUA. The results generated by such an analysis provide evidence of the relative value of one alternative compared to another, given the analytical perspective used in the analysis. In recent decades, economic evaluations have more frequently been applied to support decision making in the field of health care as a result of the increasing number of treatment options available for patients and the rising health care expenditure. In many settings, an economic evaluation of a fully developed health technology is required for reimbursement decisions of new therapies and can also be used for inclusion in the treatment guidelines [50]. The common question being asked is simply whether the additional therapeutic value of a health care technology or program is in balance with its costs. For reimbursement decision-making specifically in the Netherlands, CUAs are the preferred form of economic evaluation [51].

**Early-stage versus classical economic evaluations**

The timing at which an economic evaluation is performed can generally be classified as either early-stage, which corresponds to the development phases preceding the market approval phase, or as the classical from which corresponds to the reimbursement phase and beyond (Figure 3). As one would expect, the stakeholders involved in decision making differ according to whether an early-stage or classical HTA is conducted. For an early-stage HTA, the stakeholders are often investors and developers/manufacturers, while for a classical HTA decision making is often conducted by regulators, payers, and others involved in the diffusion phase of the technology such as care providers, hospital managers and drafters of treatment guidelines. The choice of data source used as the basis of the evidence for costs and consequences also depends on the development phase of the health care technology. Sources of evidence include clinical trials, retrospective and prospective observational studies, or modeling studies which are based on a variety of sources. Most often it is the case that clinical trial and/or observational data

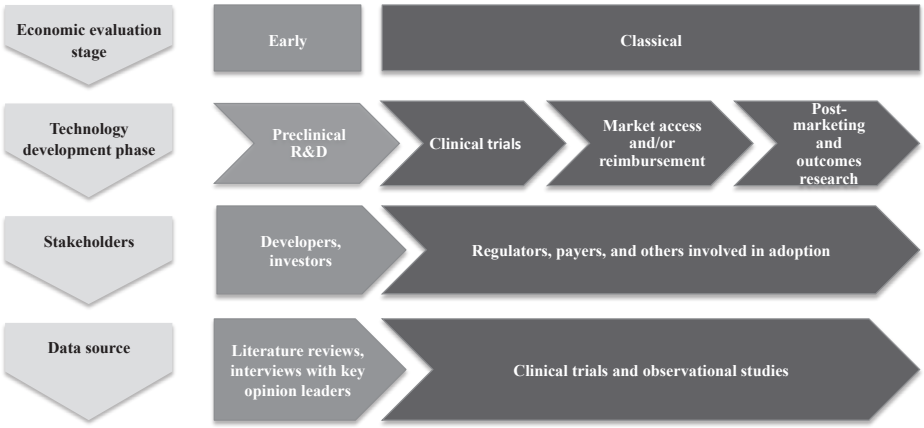



**Figure 3.** Early-stage versus the classical form of economic evaluations.

is used for classical economic evaluations, while modeling studies based on a variety of sources are more commonly used in early-stage economic evaluations [52].

Although economic evaluations can be conducted at any point in time, there are always tradeoffs inherent in decision-making. For example, the earlier a technology is assessed, the more likely its diffusion can be curtailed if it is unsafe or ineffective. At an early stage, it may also provide information through a market assessment or it can be used in go/no go decision regarding further development, future trial design, or price determination [53]. On the other hand, due to the nature of the data available during the early development phase and since the methods are intended to support policy level decision-making, the findings of an early assessment may not be definitive or perhaps misleading for the stakeholders involved [52]. For example, the technology may not yet be perfected, its users may not yet be proficient, the production costs may not yet have stabilized, the technology may not have been applied in enough circumstances to recognize its potential benefits, and its long-term effectiveness and safety may be completely uncertain. It could be argued, however, that these issues are not inherent to early-stage economic evaluations, per se, as they could be present in the classical form as well. Namely, the use of outcomes research, as a classical form, in assessing the value of a new technology in daily practice also presents methodological challenges that impact the usefulness of the results in decision-making [54]. The extent to which there are similar challenges to performing early-stage and classical deserves further attention.

### Economic evaluations in multiple myeloma

A number of partial economic evaluations, namely cost studies, in MM have been conducted using patient-level data from the hospital perspective in Sweden [55] and, more

recently, France [56] and Italy [57]. The costs reported for Swedish patients were mainly based on conventional therapies, and the results from Italy did not provide information on the use of novel agents. A more recent cost study conducted in France found that total costs were substantial mainly due to novel agents, although this conclusion was based mainly on the increasing costs due to bortezomib. It remains unclear to what extent the increased costs of treating MM are caused by the costs of novel therapies alone or other treatment-related costs. Knowledge on the major cost drivers in multiple myeloma could pinpoint resource use that represents good value for money or, conversely, instances of waste that could be targeted by cost reduction initiatives.

Previous reviews of full economic evaluations in multiple myeloma have found that economic evidence is scarce and largely lacking for novel therapies [58, 59]. Continued expansion in the availability of costly novel therapies will enhance the role of economic evaluations in reimbursement decisions and amendments to the treatment guidelines. To assess the validity of the results obtained from economic evaluations, the quality of the methodology applied to produce the results should be taken into account by users of their results. A critical review of the trends and quality of the methodology applied in economic evaluations in multiple myeloma, particularly for novel agents, has not been conducted. Therefore, the robustness of the results reported by existing studies remains questionable to users of economic evaluations. Furthermore, it is useful to identify the unmet needs and shortcomings of existing evaluations in multiple myeloma so as to inform researchers performing future evaluations of methods for quality improvement.

**Economic evaluation of stratified treatment approaches**

Stratified treatment approaches have the potential to improve treatment efficacy but also improve safety and quality-of-life as well as reduce costs. Demonstration of the clinical and economic utility of clinical markers is often needed for reimbursement decisions or to convince care providers of their usefulness in care. However, it can be difficult to estimate the utility of stratified treatment approaches, particularly in the early phases of product development or in cases where conducting a randomized controlled trial (RCT) is unethical. There are additional barriers to their development such as the development costs and time needed to validate clinical biomarkers, lower potential revenues on account of the inherent exclusion of patients, and the time and money needed to educate providers about the therapeutic implications of the clinical biomarker [60, 61]. Economic evaluations provide a framework by which all of the relevant benefits of stratified medicine can be distinguished and valued. Since economic evaluations can be performed at an early-stage development phase, it could serve useful in clarifying both the therapeutic and economic value of clinical markers in the absence of a RCT. However, there is a paucity of economic evaluations in the area of stratified medicine, with only

11 references found in a systematic review [62]. The application of economic evaluations in addressing the potential value of stratified medicine approaches deserves more attention, particularly in the area of hematologic malignancies such as multiple myeloma where multiple treatment options with heterogeneous responses are available.

## THESIS AIMS

The overall aim of this thesis is to evaluate the health benefits and costs of current and hypothetical treatment options in MM. This thesis begins with examining the evidence available from existing economic evaluations in MM and the quality thereof. To assess the methodological challenges to performing outcomes research for novel therapies in MM, this thesis then examines the feasibility of evaluating the appropriate use, effectiveness and cost-effectiveness of bortezomib in relapsed/refractory patients treated in Dutch daily practice. Subsequently, the potential therapeutic and economic value for hypothetical risk-stratified treatment approaches in MM is evaluated by means of an early-stage economic evaluation. Based on the findings of two studies, a secondary aim is to assess the feasibility and usefulness of performing early-stage economic evaluations of hypothetical risk-stratified treatment approaches in hematological malignancies.

### Research questions

1. What is the economic evidence generated by and the quality of the methodology applied in all existing economic evaluations of the treatment of multiple myeloma?
2. What is known about the efficacy, costs and impact on health-related quality-of-life particularly for novel therapies in multiple myeloma?
3. What is the real-world effectiveness and appropriate use of bortezomib for multiple myeloma in Dutch daily practice?
4. What are the real-world costs of treating multiple myeloma with bortezomib versus other therapies in Dutch daily practice?
5. What is the real-world cost-effectiveness of the novel therapy bortezomib in advanced multiple myeloma compared to the standard of care in Dutch daily practice?
6. Is there potential therapeutic and economic value at the early phase of technology development for risk-stratified treatment approaches in multiple myeloma using new and existing molecular markers?
7. Can the method of cost-effectiveness analysis be applied at an even earlier phase of technology development to demonstrate the potential value for risk-stratified treatment approaches in chronic myeloid leukemia using hypothetical molecular markers?

## THESIS OUTLINE

The research questions addressed in this thesis are categorized into three sections with each chapter addressing a separate research question. Part 1 addresses the state of knowledge about the efficiency of care in multiple myeloma and the quality of all previous work in this area (Chapter 2) and with a particular focus on novel agents (Chapter 3). Part 2 addresses the feasibility of outcomes research in evaluating the appropriate use and effectiveness (Chapter 4), costs (Chapter 5), and relative cost-effectiveness (Chapter 6) of bortezomib in multiple myeloma patients treated in Dutch daily practice. Finally, the potential value of stratified treatment in multiple myeloma is assessed in Part 3. Chapter 7 presents an early-stage economic evaluation to assess the potential cost-effectiveness of a risk-stratified approach in multiple myeloma using new and existing molecular markers to predict survival. Chapter 8 presents the application of an early-stage economic evaluation in an even earlier phase of technology development to assess the potential cost-effectiveness of a risk-stratified treatment approach for chronic myeloid leukemia, also a hematological disorder, using a hypothetical molecular marker to predict upfront treatment response. The final chapter of this thesis (Chapter 9) provides a detailed discussion of the implications of these findings for policy makers, care providers, and researchers. Some observations are then discussed regarding the extent to which there are similar challenges to performing early-stage economic evaluations and outcomes research in MM alongside providing recommendations for initiatives to meet the objectives of both.

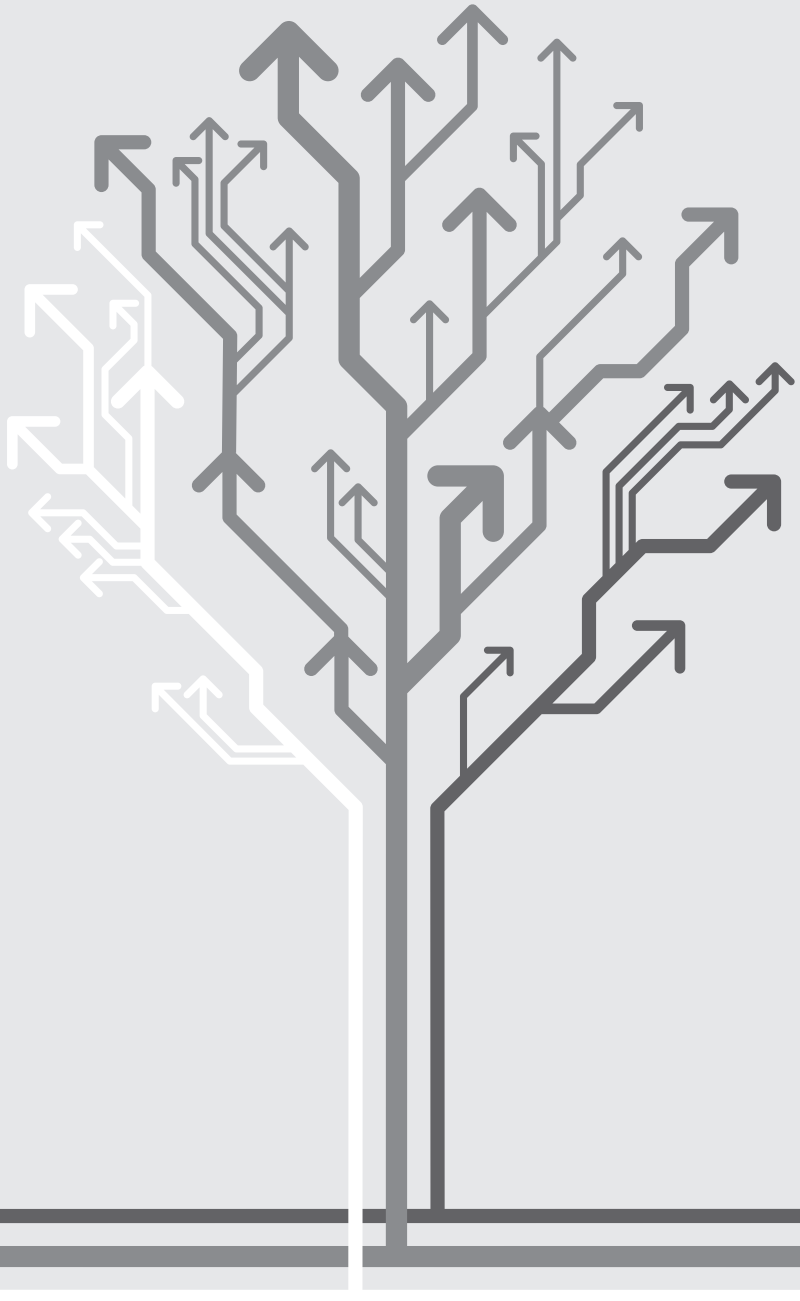
## REFERENCES

1. Palumbo A, Anderson K. Multiple myeloma. N Engl J Med. 2011;364(11):1046-60.
2. Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB. Cancer in Five Continents, Volume VIII. Lyon: International Agency for Research on Cancer (IARC); 2002.
3. Cancer figures [Cijfers over kanker] [database on the Internet]. Available from: http://cijfersoverkanker.nl/.
4. Kyle RA, Gertz MA, Witzig TE, Lust JA, Lacy MQ, Dispenzieri A, et al. Review of 1027 patients with newly diagnosed multiple myeloma. Mayo Clin Proc. 2003;78(1):21-33.
5. Dispenzieri A, Kyle RA. Neurological aspects of multiple myeloma and related disorders. Best Pract Res Clin Haematol. 2005;18(4):673-88.
6. Caers J, Vande broek I, De Raeve H, Michaux L, Trullemans F, Schots R, et al. Multiple myeloma--an update on diagnosis and treatment. Eur J Haematol. 2008;81(5):329-43.
7. Brenner H, Gondos A, Pulte D. Recent major improvement in long-term survival of younger patients with multiple myeloma. Blood. 2008;111(5):2521-6.
8. Kumar SK, Rajkumar SV, Dispenzieri A, Lacy MQ, Hayman SR, Buadi FK, et al. Improved survival in multiple myeloma and the impact of novel therapies. Blood. 2008;111(5):2516-20.
9. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA Cancer J Clin. 2010;60(5):277-300.
10. Raab MS, Podar K, Breitkreutz I, Richardson PG, Anderson KC. Multiple myeloma. Lancet.

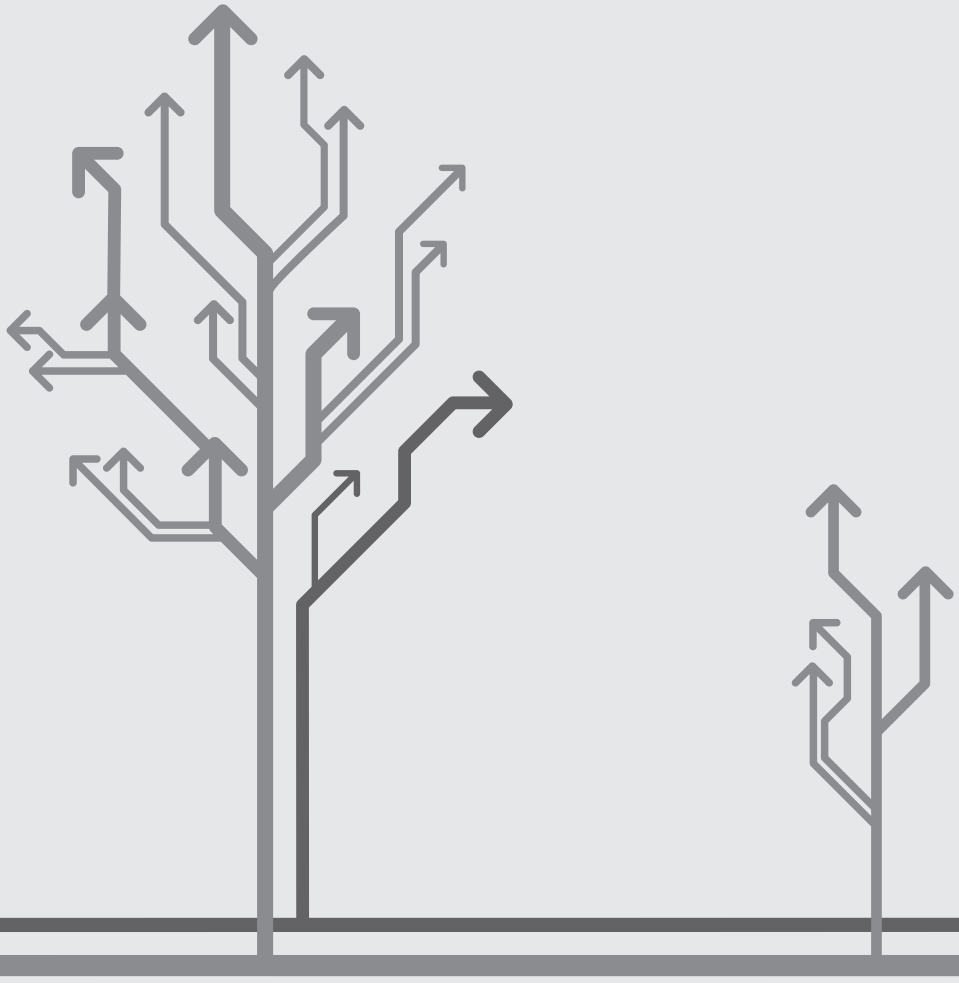
2009;374(9686):324-39.
11. Durie BG, Salmon SE. A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. Cancer. 1975;36(3):842-54.
12. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J, et al. International staging system for multiple myeloma. J Clin Oncol. 2005;23(15):3412-20.
13. Munshi NC, Anderson KC, Bergsagel PL, Shaughnessy J, Palumbo A, Durie B, et al. Consensus recommendations for risk stratification in multiple myeloma: report of the International Myeloma Workshop Consensus Panel 2. Blood. 2011;117(18):4696-700.
14. Stewart AK, Bergsagel PL, Greipp PR, Dispenzieri A, Gertz MA, Hayman SR, et al. A practical guide to defining high-risk myeloma for clinical trials, patient counseling and choice of therapy. Leukemia. 2007;21(3):529-34.
15. Hanamura I, Stewart JP, Huang Y, Zhan F, Santra M, Sawyer JR, et al. Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. Blood. 2006;108(5):1724-32.
16. Avet-Loiseau H, Attal M, Moreau P, Charbonnel C, Garban F, Hulin C, et al. Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. Blood. 2007;109(8):3489-95.
17. Brioli A, Kaiser MF, Pawlyn C, Wu P, Gregory WM, Owen R, et al. Biologically defined risk groups can be used to define the impact of thalidomide maintenance therapy in newly diagnosed multiple myeloma. Leuk Lymphoma. 2013;doi:10.3109/10428194.2012.760736.
18. Avet-Loiseau H. Role of genetics in prognostication in myeloma. Best Pract Res Clin Haematol. 2007;20(4):625-35.
19. Kyle RA, Remstein ED, Therneau TM, Dispenzieri A, Kurtin PJ, Hodnefield JM, et al. Clinical course and prognosis of smoldering (asymptomatic) multiple myeloma. N Engl J Med. 2007;356(25):2582-90.
20. Harousseau JL, Attal M, Avet-Loiseau H, Marit G, Caillot D, Mohty M, et al. Bortezomib plus dexamethasone is superior to vincristine plus doxorubicin plus dexamethasone as induction treatment prior to autologous stem-cell transplantation in newly diagnosed multiple myeloma: results of the IFM 2005-01 phase III trial. J Clin Oncol. 2010;28(30):4621-9.
21. Rajkumar SV, Jacobus S, Callander NS, Fonseca R, Vesole DH, Williams ME, et al. Lenalidomide plus high-dose dexamethasone versus lenalidomide plus low-dose dexamethasone as initial therapy for newly diagnosed multiple myeloma: an open-label randomised controlled trial. Lancet Oncol. 2010;11(1):29-37.
22. Rajkumar SV, Rosinol L, Hussein M, Catalano J, Jedrzejczak W, Lucy L, et al. Multicenter, randomized, double-blind, placebo-controlled study of thalidomide plus dexamethasone compared with dexamethasone as initial therapy for newly diagnosed multiple myeloma. J Clin Oncol. 2008;26(13):2171-7.
23. Waage A, Palumbo AP, Fayers P, Beksac M, Hulin C, Mary J, et al. MP versus MPT for previously untreated elderly patients with multiple myeloma: A meta-analysis of 1,682 individual patient data from six randomized clinical trials. ASCO Meeting Abstracts. 2010;28(15_suppl):8130.
24. Mateos MV, Richardson PG, Schlag R, Khuageva NK, Dimopoulos MA, Shpilberg O, et al. Bortezomib plus melphalan and prednisone compared with melphalan and prednisone in previously untreated multiple myeloma: updated follow-up and impact of subsequent therapy in the phase III VISTA trial. J Clin Oncol. 2010;28(13):2259-66.
25. San Miguel JF, Schlag R, Khuageva NK, Dimopoulos MA, Shpilberg O, Kropff M, et al. Bortezo-

mib plus melphalan and prednisone for initial treatment of multiple myeloma. N Engl J Med. 2008;359(9):906-17.
26. Zonder JA, Crowley J, Hussein MA, Bolejack V, Moore Jr DF, Whittenberger BF, et al. Lenalidomide and high-dose dexamethasone compared with dexamethasone as initial therapy for multiple myeloma: A randomized Southwest Oncology Group trial (S0232). Blood. 2010;116(26):5838-41.
27. Kastritis E, Palumbo A, Dimopoulos MA. Treatment of relapsed/refractory multiple myeloma. Semin Hematol. 2009;46(2):143-57.
28. Richardson PG, Sonneveld P, Schuster MW, Irwin D, Stadtmauer EA, Facon T, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med. 2005;352(24):2487-98.
29. Dimopoulos M, Spencer A, Attal M, Prince HM, Harousseau JL, Dmoszynska A, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. N Engl J Med. 2007;357(21):2123-32.
30. Weber DM, Chen C, Niesvizky R, Wang M, Belch A, Stadtmauer EA, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. N Engl J Med. 2007;357(21):2133-42.
31. Alvares CL, Davies FE, Horton C, Patel G, Powles R, Morgan GJ. The role of second autografts in the management of myeloma at first relapse. Haematologica. 2006;91(1):141-2.
32. Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. Nat Rev Drug Discov. 2007;6(4):287-93.
33. Rajkumar SV. Treatment of multiple myeloma. Nat Rev Clin Oncol. 2011;8(8):479-91.
34. Barlogie B, Crowley J. Could CR mean cure? Blood. 2011;118(3):483.
35. Avet-Loiseau H, Leleu X, Roussel M, Moreau P, Guerin-Charbonnel C, Caillot D, et al. Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). J Clin Oncol. 2010;28(30):4630-4.
36. Cavo M, Tacchetti P, Patriarca F, Petrucci MT, Pantani L, Galli M, et al. Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: A randomised phase 3 study. The Lancet. 2010;376(9758):2075-85.
37. Neben K, Lokhorst HM, Jauch A, Bertsch U, Hielscher T, van der Holt B, et al. Administration of bortezomib before and after autologous stem cell transplantation improves outcome in multiple myeloma patients with deletion 17p. Blood. 2012;119(4):940-8.
38. Sonneveld P, Schmidt-Wolf IG, van der Holt B, El Jarari L, Bertsch U, Salwender H, et al. Bortezomib Induction and Maintenance Treatment in Patients With Newly Diagnosed Multiple Myeloma: Results of the Randomized Phase III HOVON-65/ GMMG-HD4 Trial. J Clin Oncol. 2012.
39. Cavo M, Tacchetti P, Patriarca F, Petrucci MT, Pantani L, Galli M, et al. Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: a randomised phase 3 study. Lancet. 2010;376(9758):2075-85.
40. Morgan GJ, Gregory WM, Davies FE, Bell SE, Szubert AJ, Brown JM, et al. The role of maintenance thalidomide therapy in multiple myeloma: MRC Myeloma IX results and meta-analysis. Blood. 2012;119(1):7-15.
41. Bonkowski J, Vermeulen LC, Kolesar JM. The clinical utility of lenalidomide in multiple myeloma and myelodysplastic syndromes. J Oncol Pharm Pract. 2009;16(4):223-32.
42. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Farmacotherapeutisch Kompas. [cited 2013 January 30]; Available from: http://www.fk.cvz.nl/.
43. Lee SJ, Klar N, Weeks JC, Antin JH. Predicting costs of stem-cell transplantation. J Clin Oncol. 2000;18(1):64-71.
44. Blommestein HM, Verelst SG, Huijgens PC, Blijlevens NM, Cornelissen JJ, Uyl-de Groot CA. Real-world

costs of autologous and allogeneic stem cell transplantations for haematological diseases: a multicentre study. Ann Hematol. 2012.
45. Mitsiades CS, Hideshima T, Chauhan D, McMillin DW, Klippel S, Laubach JP, et al. Emerging treatments for multiple myeloma: beyond immunomodulatory drugs and bortezomib. Semin Hematol. 2009;46(2):166-75.
46. Williamson S. Patient access schemes for high-cost cancer medicines. Lancet Oncol. 2010;11(2):111-2.
47. Kuijpers MR, Toenders WGM. Assessment procedure for inpatient drugs [Procedure beoordeling intramurale geneesmiddelen]. Diemen: Dutch Health Care Insurance Board; 2006 [cited January 2013]; Available from: http://www.cvz.nl/binaries/live/cvzinternet/hst_content/nl/documenten/rapporten/2006/rpt0604+intramurale+geneesmiddelen.pdf.
48. Velasco-Garrido M, Busse R. Health technology assesssment: an introduction to objectives, role of evidence, and structure in Europe. Copenhagen: WHO regional Office for Europe 2005 [cited 2013 February ]; Available from: http://www.euro.who.int/__data/assets/pdf_file/0018/90432/E87866.pdf.
49. Drummond MF, Schulpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. Third ed. Oxford: Oxford University Press; 2005.
50. van Oostenbruggen MF, Jansen RB, Mur K, Kooijman H. Penny and pound wise: pharmacoeconomics from a governmental perspective. Pharmacoeconomics. 2005;23(3):219-26.
51. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Guidelines for pharmacoeconomic research, updated version. Diemen, Netherlands: CVZ; 2006 [cited 2013 January 30]; Available from: http://www.cvz.nl/binaries/live/cvzinternet/hst_content/nl/documenten/rubriek+zorgpakket/cfh/richtlijnen+farmaco-economisch+onderzoek+-+engels.pdf.
52. Ijzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. Appl Health Econ Health Policy. 2011;9(5):331-47.
53. Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: a methodological and empirical review. Int J Technol Assess Health Care. 2008;24(4):465-72.
54. Lu CY. Observational studies: a review of study designs, challenges and strategies to reduce confounding. Int J Clin Pract. 2009;63(5):691-7.
55. Ghatnekar O, Alvegard T, Conradi N, Lenhoff S, Mellqvist UH, Persson U, et al. Direct hospital resource utilization and costs of treating patients with multiple myeloma in Southwest Sweden: a 5-year retrospective analysis. Clin Ther. 2008;30(9):1704-13.
56. Armoiry X, Fagnani F, Benboubker L, Facon T, Fermand JP, Hulin C, et al. Management of relapsed or refractory multiple myeloma in French hospitals and estimation of associated direct costs: a multicentre retrospective cohort study. J Clin Pharm Ther. 2011;36(1):19-26.
57. Koleva D, Cortelazzo S, Toldo C, Garattini L. Healthcare costs of multiple myeloma: an Italian study. Eur J Cancer Care (Engl). 2011;20(3):330-6.
58. Moeremans K, Annemans L. An update: health economics of managing multiple myeloma. Eur J Cancer. 2006;42(11):1684-91.
59. Wisloff F, Gulbrandsen N, Nord E. Therapeutic options in the treatment of multiple myeloma: pharmacoeconomic and quality-of-life considerations. Pharmacoeconomics. 1999;16(4):329-41.
60. Deverka PA, Doksum T, Carlson RJ. Integrating molecular medicine into the US health-care system: opportunities, barriers, and policy challenges. Clin Pharmacol Ther. 2007;82(4):427-34.
61. Gurwitz D, Zika E, Hopkins MM, Gaisser S, Ibarreta D. Pharmacogenetics in Europe: barriers and opportunities. Public Health Genomics. 2009;12(3):134-41.
62. Phillips KA, Van Bebber SL. A systematic review of cost-effectiveness analyses of pharmacogenomic interventions. Pharmacogenomics. 2004;5(8):1139-49.



# Part 1 – Review of economic evaluations in multiple myeloma

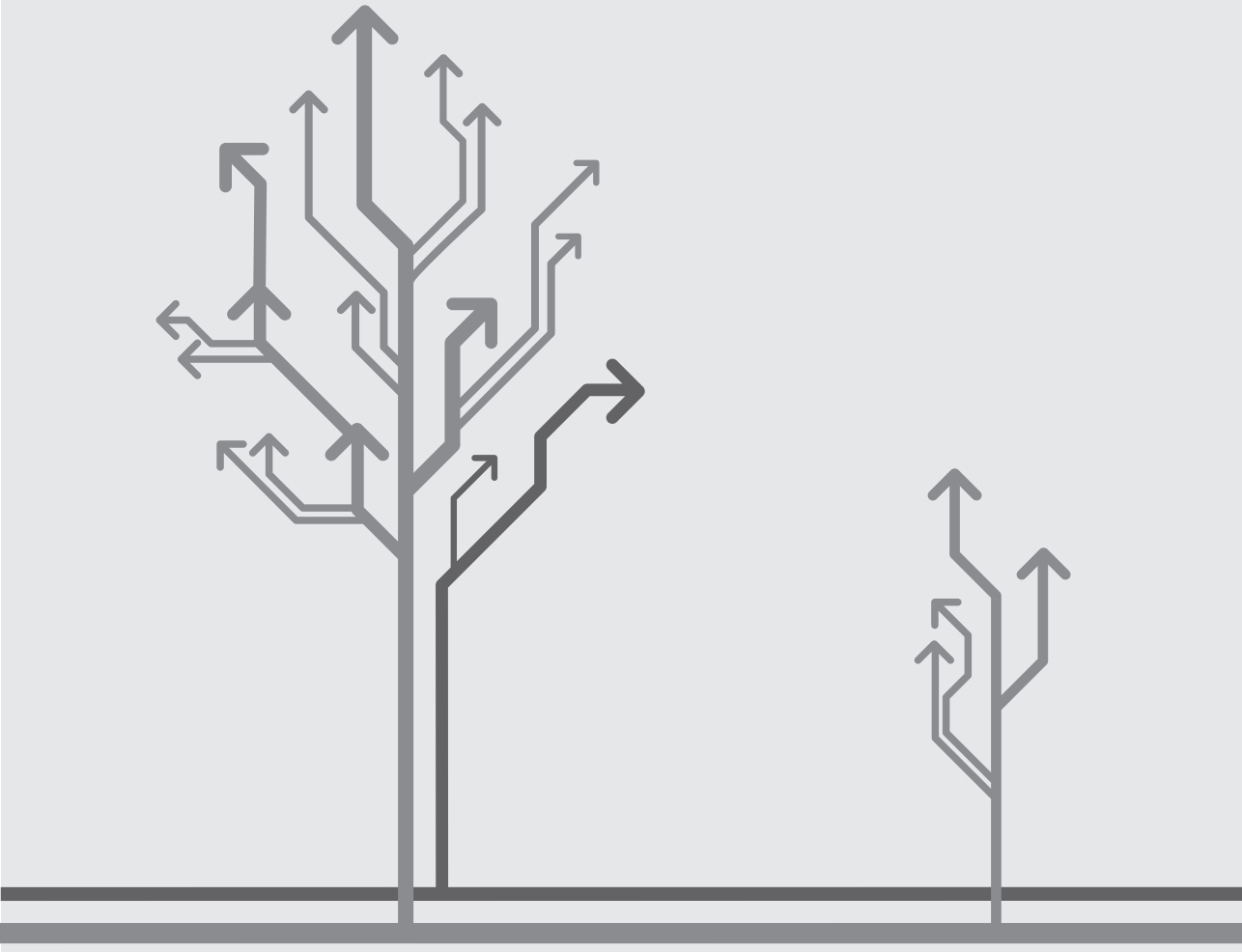


# Chapter 2.

## Critical review of economic evaluations in multiple myeloma: An overview of the economic evidence and quality of the methodology

Jennifer Gaultney, William K. Redekop, Pieter Sonneveld, Carin A. Uyl-de Groot

*Published in: European Journal of Cancer. 2011 Jul;47(10):1458-67.*



## ABSTRACT

Continued expansion in the availability of costly alternative therapies in multiple myeloma will enhance the role of economic evaluations in reimbursement decisions and amendments to the treatment guidelines. The quality of economic evaluations should be taken into account by clinicians involved in decision-making. A systematic review and critique of the methodology was performed to assess the trends and quality in economic evaluations in multiple myeloma to date. A literature search was conducted to identify full economic evaluations in multiple myeloma as of December 2009. Details of the economic evaluation methods applied were extracted. Each study underwent a quality assessment based on the Drummond checklist for high-quality economic evaluations in healthcare. Eighteen published economic evaluations were identified. Stem cell transplantation in combination with intensive chemotherapy has been demonstrated to be cost-effective, while interferon alpha is generally ineffective at additional costs. Evaluations have become less frequent in the last decade, especially for newer therapies despite their important contribution to improvements in outcomes. The quality of the methodology applied and its documentation can be improved in many aspects. As users of the results of economic evaluations, clinicians involved in guiding decision-making should be critical of the quality of economic evaluations in multiple myeloma. To ensure access to and identification of high quality studies, researchers conducting economic evaluations of future advances should strive toward evaluations that fulfill the Drummond criteria and are properly documented.

## INTRODUCTION

Multiple myeloma is a progressive hematologic malignancy accounting for approximately 0.8% of all cancer diagnoses and 0.9% of all cancer deaths worldwide [1]. It is considered a severe disease that remains incurable.

The treatment paradigm for multiple myeloma usually consists of initial treatment which may include stem cell transplantation (SCT) followed by maintenance therapy to prolong patient response to initial therapy and treatment for relapsed or disease refractory to initial therapy. Supportive agents to alleviate symptoms of the disease or side effects of treatment are also commonly administered. Novel therapies, such as thalidomide, bortezomib and lenalidomide, promise improved outcomes [2] at increased acquisition costs. Additional agents currently under development are expected to be added to the list of options available in the future.

Previous reviews of economic evaluations in multiple myeloma have found that economic evidence is scarce and largely lacking for novel therapies [3, 4]. Earlier advances in treatment, such as SCT, resulted in higher incremental costs but led to relatively greater incremental effectiveness. More recent advances, however, have been shown to be marginally less cost-effective compared with earlier advances, despite being considered cost-effective. Accordingly, future treatments will likely become marginally less cost-effective, thus enhancing the role of cost-effectiveness studies in reimbursement decisions for multiple myeloma.

Continued expansion in the availability of costly alternative therapies will enhance the role of economic evaluations in reimbursement decisions and amendments to the treatment guidelines. To assess the validity of the results obtained from economic evaluations, the quality of the methodology applied to produce the results should be taken into account by users of their results. Existing reviews of economic evaluations in multiple myeloma did not include a quality assessment component. Therefore, the validity of the results reported by such studies remains questionable to users of economic evaluations.

The objective of this review was to provide an overview of the trends in methodology. A critical appraisal of the methodology was also conducted to identify whether the quality was adequate for users of their results in decision-making and to identify needed improvements for future evaluations.

## METHODS

### Search

A literature search was performed as of 31 December 2009 (Medline, EMBASE, Cochrane Database of Systematic Reviews) using the search strategy described in Appendix A. The following inclusion and exclusion criteria were used: (1) treatment for multiple myeloma was the main topic of the article; (2) only research articles were considered and abstracts and reviews were excluded after screening the references for relevant publications; and (3) only articles using the English language were considered. Full papers were obtained for publications whose titles or abstracts were considered relevant or where the title or abstract information was not sufficient to make a decision. Full papers were rejected if the article did not include both a costing and an effectiveness element to the study. Studies using data from a patient population heterogeneous for many types of diseases in addition to multiple myeloma were excluded if both cost and effect estimates specific to multiple myeloma patients were not reported.

### Data extraction

Details of the economic evaluation methods applied were extracted, including the type of evaluation, comparators, source of data for effectiveness, time horizon, payer perspective, inclusion of indirect costs, effectiveness outcomes, and incremental costs and effectiveness. Source of funding for the study was also extracted.

### Quality assessment

The criteria used for the quality assessment were based on Drummond's checklist (Table 1), which is a standard quality assessment checklist specifically designed to critically assess economic evaluations [5]. It provides a list of ten questions to assist users of economic evaluations in separating the various elements of methodology applied in economic evaluations so that each can be scrutinized. Questions 1 through 9 and all subquestions for criterion 10 were used. Each of the questions was operationalized to enable a 'yes' or 'no' answer for each item on the checklist. If the article did not provide enough information to determine a clear answer to the question, the article was scored with a 'no' for the criterion in question. If a criterion was met only for the costs or effects, this was also scored with a 'no' but made transparent.

**Table 1.** List of criteria adapted from Drummond et al. (2005) applied in quality assessment.

| | |
|---|---|
| **Q1** | Was a well-defined question posed in answerable form? |
| **Q2** | Was a comprehensive description of the competing alternatives given? |
| **Q3** | Was the effectiveness of the programs or services established? |
| **Q4** | Were all relevant costs/consequences for each alternative identified in light of viewpoint? |
| **Q5** | Were costs and consequences measured in appropriate physical units? |
| **Q6** | Were costs and consequences valued credibly? |
| **Q7** | Were costs and consequences adjusted for differential timing (i.e., discounted)? |
| **Q8** | Was an incremental analysis of costs and consequences of alternatives performed? |
| **Q9** | Was the impact of uncertainty in the estimates of costs and consequences examined? |
| **Q10a** | Was the conclusion easily interpretable and based on objective comparison in terms of costs and effect difference? |
| **Q10b** | Were the results compared with those of others and allowances made for methodological differences? |
| **Q10c** | Did the study discuss the generalizability of the results to other settings/patient groups? |
| **Q10d** | Did the study allude to or take account of other important factors in the choice or decision under consideration? |
| **Q10e** | Did the study discuss issues of implementation and whether free resources could be redeployed to other programs? |

## RESULTS

### Search

Figure 1 depicts the selection process conducted for this review. The search identified a total of 967 potentially relevant articles. After reviewing the titles and abstracts, 33 articles were selected for full review. Finally, we reviewed 18 studies that reported full economic evaluations, which included six cost-effectiveness analyses (CEA), five cost-minimization analyses (CMA), three cost-utility analyses (CUA), three cost-benefit analyses (CBA) and one cost-consequence analysis (CCA). No CUAs have been conducted since 2004.

### General characteristics of selected studies

All identified economic evaluations were published between the years 1994 and 2009. Details regarding the comparators and source of effectiveness are provided in Table 2. Details of the methodology are provided in Table 3.

#### *Comparators and stage of treatment*

An economic evaluation was published for each stage in the treatment paradigm for multiple myeloma (Table 2). The majority of studies evaluated initial therapy, which consisted of conventional chemotherapy and/or stem cell transplant. Few have been published for targeted agents, with only one CEA conducted for bortezomib compared to best supportive care or thalidomide in the relapsed/refractory phase of treatment.

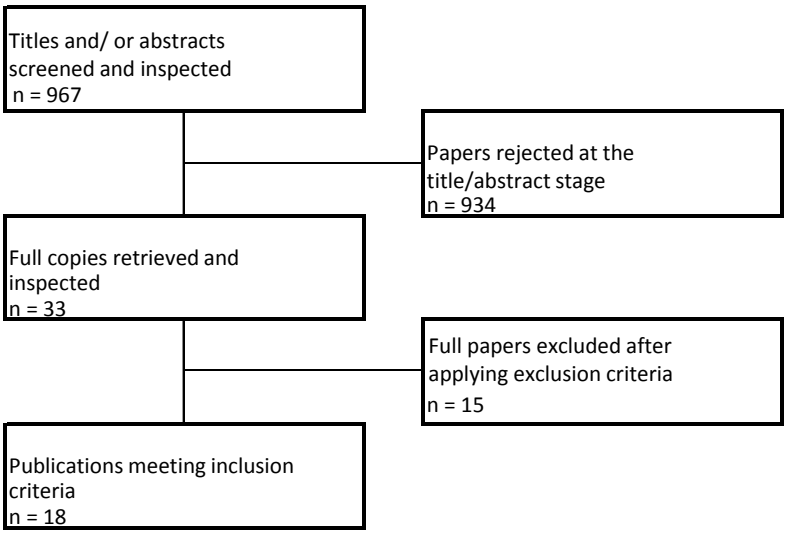



**Figure 1.** Flowchart describing the article selection process.

***Data source for effectiveness***

Effectiveness was demonstrated by a RCT (n=7), retrospective cohort study (n=4), and a meta-analysis of RCTs (n=3). The study design applied to estimate effectiveness for the intervention and alternative comparators differed for five studies. The weakest comparisons were studies that compared effectiveness estimated from a RCT design to that of estimates from a retrospective cohort study [6] and a Delphi panel supplemented with estimates from the literature [7]. One study compared effectiveness estimated from a RCT to a meta-analysis that pooled estimates from both controlled and uncontrolled studies [8]. Two studies compared estimates from a prospective design to that from a retrospective design with both groups matched for patient characteristics to simulate a retrospective case-control study [9, 10].

***Perspective***

The perspective was not stated for seven studies (Table 3). Perspectives that were stated included that of the payer, society, hospital and provider.

***Time horizon and effectiveness outcomes***

The time horizon varied between studies but was generally limited to the stage within the treatment paradigm. Evaluations of initial treatment most commonly adopted the longest time horizon and reported effectiveness outcomes in either life-years (LYs) or quality-adjusted life-years (QALYs). Studies assessing costs and benefits of SCT typically adopted a short time horizon since the objective of such analyses was to assess methods to save costs during the transplantation procedure. Consequently, effectiveness out-

comes were reported in term of costs of treatment and adverse events. For maintenance and relapsed/refractory treatment, time horizons were either one year [11] or survival from start of treatment [7, 12], with effectiveness outcomes measures being months or LYs, respectively. Economic evaluations of supportive therapies assessed the costs and effects of treatment from 1 to 3 years and reported effectiveness outcomes in terms of monetary costs [13-15].

***Incremental cost and effectiveness***

Most studies found the intervention to result in either improved or equal effectiveness compared to the alternative as well as increased costs. Some trends in incremental cost and effectiveness results can be seen. Transplantation alone or in combination with intensive therapy has generally been demonstrated to be effective though more costly [6, 9, 10, 16-18], with the exception of one study [19]. The addition of interferon alpha (IFN) to therapy has generally been shown to be limited in terms of health outcomes while at additional costs [8, 12, 20], with the exception of IFN in combination with VMPC (vincristine, melphalan, cyclophosphamide and prednisolone) [21]. Cost-savings were demonstrated for outpatient versus inpatient transplants[22], use of vincristine, adriamycin, dexamethasone (VAD) plus pegylated liposomal doxorubicin (DVd) instead of VAD plus low dose dexamethasone (VAd) [23], as well as large instead of standard volume leukapheresis collection when two transplants are required. Zoledronic acid for supportive care with bisphosphonates is as effective as pamidronate, though the incremental cost difference was not significantly different [15].

***Acknowledgement of funding***

Variation in the source of funding was observed, with studies acknowledging either no funding, funding from the manufacturer, government and non-profit or academic organizations.

**Quality of selected studies**

Figure 2 describes a summary of the quality of all studies included in this review according to the Drummond criteria. Details of the critique according to questions 1 through 9 and subquestions for question 10 for each study are available in Appendix B and C, respectively.

Nine studies (50%) included a well-defined research question by stating somewhere in the article the objective of the analysis, comparator strategies and the perspective of the analysis [6-10, 13, 17, 18, 20]. Almost all studies (89%) clearly described the alternative given [6-10, 12-16, 18-20, 22-24].

**Table 2.** Comparator and source of effectiveness from published economic evaluations in multiple myeloma.

| Study | Intervention | Alternative | Stage of treatment | Source of effectiveness estimates (intervention/alternative) | Sample size (intervention/alternative) |
|---|---|---|---|---|---|
| Donatini et al. (1994)[19] | HDM+PBSCS[a] | Chemotherapy | Initial; Relapsed/refractory | Retrospective cohort study | 16 / 11 |
| Laakso et al. (1994)[14] | Clodronate | Placebo | Supportive | RCT | 156 / 156 |
| Henon et al. (1995)[16] | HDM+Auto-BMT (group I)[b] | M2 or VAD (group II and group III)[b] | Ititial | Retrospective cohort study | 12 / 10 (group II); 15 (group III) |
| Duncan et al. (1996)[6] | ABMT | PBSCT | Transplant | RCT / Retrospective cohort study | 14 / 37 |
| Jagannath et al. (1997)[22] | Outpatient tandem autotransplants | Inpatient tandem autotransplants | Transplant | Retrospective cohort study | 91 / 160 |
| Nord et al. (1997)[20] | MP-IFN | MP | Initial | RCT | 285 / 298 |
| Trippoli et al. (1997)[12] | IFN | None | Maintenance | Meta-analysis | 393 / 423 |
| Trippoli et al. (1998)[8] | ABMT | MP or MP+IFN | Initial | Meta-analysis of controlled and uncontrolled studies | 100 / 850 (MP); 181 (MP+IFN) |
| Dranitsaris et al. (1999)[13] | Pamidronate | Placebo | Supportive | RCT | 196 / 181 |
| Ludwig et al. (2000)[11] | IFNa+VMCP | VMCP | Initial; Maintenance | Meta-analysis | 1,966 / 1,982 |
| Gulbrandsen et al. (2001)[9] | HDM+ABSCS | MP | Initial | Retrospective case-control study | 274 / 70 |
| Sampson et al. (2001)[17] | HDM+ABMT | Chemotherapy[a] | Initial | RCT | 100 / 100 |
| Kouroukis et al. (2003)[10] | VAD+ASCT | MP | Initial | Retrospective case-control study | 36 / 16 |
| Mehta et al. (2004)[7] | BMB (± previous Thal) | BSC or Thal | Relapse/refractory | RCT / Delphi panel | 202 / 6[c] |
| Van Agthoven et al. (2004)[18] | Intensive M + myeloablative cyclo + ASCT | Intensive M | Initial | RCT | 129 / 132 |
| Reed et al. (2005)[15] | Zoledronic acid | Pamidronate | Supportive | RCT | 151 / 138 |
| Porter et al. (2007)[23] | DVd | VAd | Initial | RCT | 97 / 95 |

**Table 2.** (Continued)

| Study | Intervention | Alternative | Stage of treatment | Source of effectiveness estimates (intervention/alternative) | Sample size (intervention/alternative) |
|---|---|---|---|---|---|
| Zubair et al. (2009)[24] | Large volume leukapheresis | Standard volume leukapheresis | Transplant | Retrospective cohort study | 35 / 52 |

[a]Standard chemotherapy consisting primarily of vindesine, adriamycin, prednisone, and carmustine; [b]Group I patients of DS stage I, group II patients of DS stage I, group III patients of DS stage II; [c]A total of 6 experts were surveyed to elicit effectiveness estimates based on a hypothetical cohort for which the alternative would be administered. ABMT: autologous bone marrow transplant; ABSCS: autologous blood stem cell support; ASCT: autologous stem cell transplantation; BSC: best supportive care; cyclo: cyclophosphamide; DR: Durie Salmon; DVd: vincristine, adriamycin, dexamethasone (VAD) plus pegylated liposomal doxorubicin; HDM: high-dose melphalan; IFN: interferon; M: melphalan; M2: BCMU (1,3 di[2-chloroethyl]-1-nitrosourea), eldisine, cyclophosphamide, melphalan; MP: melphalan prednisone; PBSCS: peripheral blood stem cell support; PBSCT: peripheral blood stem cell transplant; RCT: randomized controlled trial; Thal: thalidomide; VAd: VAD plus low dose dexamethasone; VMCP: vincristine, melphalan, cyclophosphamide and prednisolone.

**Table 3.** Methodological characteristics of published economic evaluations in multiple myeloma.

| Study | Type | Perspective (country) | Indirect costs | Time horizon | Discount rate | Effectiveness outcomes | Incremental Effects | Incremental Costs | Funding source |
|---|---|---|---|---|---|---|---|---|---|
| Donatini et al. (1994)[19] | CMA | Not stated (FRA) | None | 6 years | Not done | OS;<br>QoL score | Equal | US$34465 | None stated |
| Laakso et al. (1994)[14] | CBA | Not stated (FIN) | None | 2 years | Not done | AEs | Included as costs | 51 FM /day | Manufacturer<br>Nonprofit<br>Government |
| Henon et al. (1995)[16] | CEA | Not stated (FRA) | None | 5 years | Not done | Weeks | Group I vs II: 138 weeks;<br>Group I vs III: 20 weeks | Group I vs II: US$10,145;<br>Group I vs. III: US$19,270 | None stated |
| Duncan et al. (1996)[6] | CMA | Hospital (GBR) | None | Transplantation phase | NA | Inpatient hospital days;<br>AE-free days | Equal[a] | 3,031£ | Manufacturer |

**Table 3.** (Continued)

| Study | Type | Perspective (country) | Indirect costs | Time horizon | Discount rate | Effectiveness outcomes | Incremental Effects | Incremental Costs | Funding source |
|---|---|---|---|---|---|---|---|---|---|
| Jagannath et al. (1997)[22] | CMA | Not stated (USA) | Productivity loss | 2 months | NA | AEs | Equal[a] | US$-13,172 | Government |
| Nord et al. (1997)[20] | CUA | Societal (NOR, DNK, SWE) | Productivity loss | Death or censored | costs: 5% | QALYs | 0.125 QALYs | NOK 87,600 (US$ 13,700) | Manufacturer |
| Trippoli et al. (1997)[12] | CMA | Not stated (ITA) | None | Lifetime | Not done | LYs | Equal | US$42,000 | None stated |
| Trippoli et al. (1998)[8] | CEA | Societal (ITA) | None | Lifetime | costs: 5% effects: 5% | LYs | ABMT vs MP: 2.23 LYs; MP vs MP+IFN: Equal | ABMT vs MP: US$57,333; MP vs MP+IFN: not done | None stated |
| Dranitsaris et al. (1999)[13] | CBA | Societal (CAN) | None | 3 years | benefits: 3% | Risk of skeletal fracture; AEs | Included as costs | Patient: WTP of Can$3,364;C Society: Can$789 | Manufacturer |
| Ludwig et al. (2000)[11] | CEA | Not stated (International) | None | 1 year | NA | Months | Initial: 3.5 months; Maintenance: 7.6 months | Initial: US$161.33/wk; Maintenance: US$154.67/wk | Hospital |
| Gulbrandsen et al. (2001)[9] | CUA | Societal (NOR) | Productivity loss | 3 years | costs: 5% | QALYs | 1.2 QALYs | NOK 299,000 (US$ 32,300 ) | Nonprofit |
| Sampson et al. (2001)[17] | CEA | Provider (GBR) | None | 5 years | Not done | LYs | 0.7 LYs all patients; 0.8 LYs <60 years age; | 10,480£ | None Stated |
| Kouroukis et al. (2003)[10] | CEA | Payer (CAN) | None | Lifetime | costs: 5%; effects: 3% | Months | 19.3 months | US$30,517 | Nonprofit |

**Table 3.** (Continued)

| Study | Type | Perspective (country) | Indirect costs | Time horizon | Discount rate | Effectiveness outcomes | Incremental Effects | Incremental Costs | Funding source |
|---|---|---|---|---|---|---|---|---|---|
| Mehta et al. (2004)[7] | CEA | Payer (USA) | None | Lifetime | Not done | LYs | BMB vs BSC: 1.13 LYs; BMB[c] vs Thal: 1.1 LYs; BMB[d] vs Thal: 1.45 LYs | BMB vs BSC: US$50,797; BMB[c] vs Thal: US$54,777; BMB[d] vs Thal: US$31,551 | Manufacturer |
| Van Agthoven et al. (2004)[18] | CUA | Hospital (NLD) | None | 3 years | costs and effects: 4% | LYs; QALYs | 0.14LYs; 0.24QALYs | €13,067 | None stated |
| Reed et al. (2005)[15] | CCA | Not stated (International) | None | 13 months | NA | AEs | Equal | US$1,982[b] | Academic Manufacturer |
| Porter et al. (2007)[23] | CCA | Payer (USA) | None | Not stated | Not done | RR; PFS; OS | Equal | US-$1404 | None stated |
| Zubair et al. (2009)[24] | CBA | Not stated (USA) | Productivity loss; Transportation Accommodation | PBSC collection phase | NA | Relapse rate AEs | Included as costs | US$-7,497 | Government |

[a]Validity of assumption is questionable; [b]No significant difference; [c]In patients previously treated with thalidomide; [d]In patients not previously treated with thalidomide. AEs: adverse events; Auto-BMT: autologous bone marrow transplant; Bmib: bortezomib; BSC: best supportive care; EFS: event-free survival; IFN: interferon; LYs: life-years; M: melphalan; MP+IFN: melphalan prednisone plus interferon alpha; NA: not applicable; OS: overall survival; PBSC: peripheral blood stem cell; PFS: progression-free survival; prev: previous; QALYs: quality-adjusted life-years; QoL: quality of life; RR: risk ratio; vs: versus; Thal: thalidomide; WTP: willingness to pay.

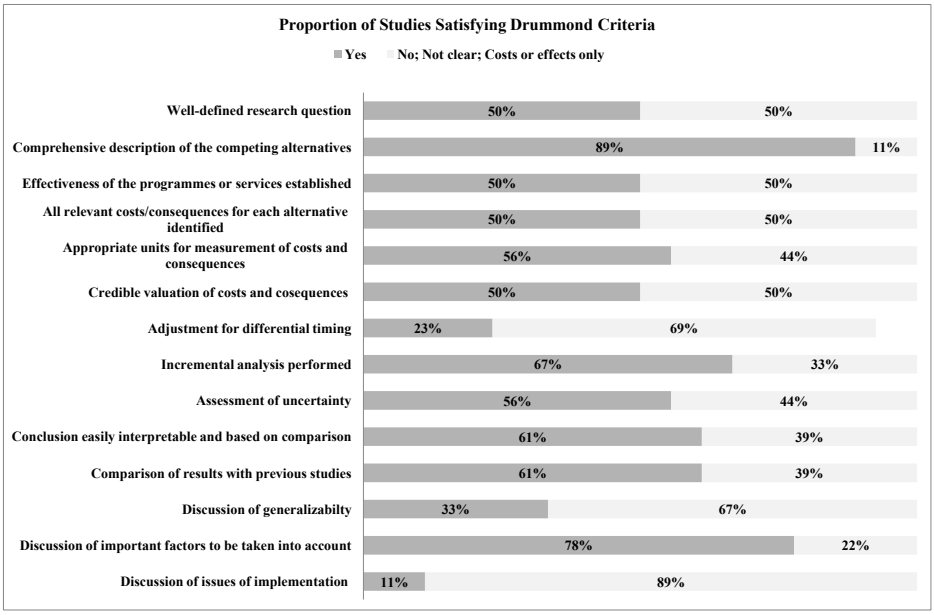



**Figure 2.** Summary of quality according to Drummond criteria.

Nine studies (50%) adequately demonstrated comparative effectiveness, mainly by means of a RCT or a meta-analysis of RCTs. The remaining studies demonstrated effectiveness of the comparator strategies by means of non-experimental comparisons.

Whether all relevant costs and consequences were identified and valued in appropriate units was difficult to determine for half (56%) of all studies. This was particularly a problem for studies that did not mention the perspective of the analysis [11, 12, 14-16, 19, 22, 24].

Nine studies (50%) valued both costs and consequences credibly [9-11, 13, 15, 16, 18, 20, 23], while three did not provide enough information regarding the valuation procedure for costs in order to judge quality [12, 17, 24]. The remaining studies did not meet this criterion for a variety of reasons. One study compared post-therapy survival time between two groups for which time in therapy began at the start of therapy for the intervention while for the alternative follow-up began at the end administration of therapy [19]. An invalid assumption of equal effectiveness between comparator groups was assumed in one study [6] despite being shown to be statistically different and for another due to incomparable patient groups[22]. The validity of the cost analysis was questionable for one study which reported an insignificant difference in costs between the two comparator groups despite the significantly fewer costly side effects reported in the trial [14]. A cost analysis assuming equal long-term costs for different comparator

treatments and using cost estimates taken from a variety of country perspectives and institutions was considered invalid [8]. Lastly, the use of a Delphi panel to estimate the costs and effectiveness of the alternative strategy was not considered credible [7].

Five studies (38%) did not incorporate discounting because the time horizon of the analysis did not extend beyond 1 year [6, 11, 15, 22, 24]. Of the studies with a time horizon greater than one year, three (23%) discounted both costs and consequences

**Table 4.** Critique of selected studies according to criteria 1 through 9 of Drummond et al. (2005).

| First author | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
|---|---|---|---|---|---|---|---|---|---|
| Donatini et al. (1994)[19] | no | yes | no | viewpoint not clear | viewpoint not clear | costs only | no | no | no |
| Laakso et al. (1994)[14] | no | yes | yes | viewpoint not clear | effects only | effects only | no | no | no |
| Henon et al. (1995)[16] | no | yes | no | viewpoint not clear | viewpoint not clear | yes | no | yes | no |
| Duncan et al. (1996)[6] | yes | yes | no | yes | yes | costs only | irrelevant | no | yes |
| Jagannath et al. (1997)[22] | no | yes | no | viewpoint not clear | viewpoint not clear | costs only | irrelevant | costs only | no |
| Nord et al. (1997)[20] | yes | yes | yes | yes | yes | yes | costs only | yes | yes |
| Trippoli et al. (1997)[12] | no | yes | yes | viewpoint not clear | effects only | effects only | no | yes | no |
| Trippoli et al. (1998)[8] | yes | yes | no | yes | yes | effects only | yes | yes | yes |
| Dranitsaris et al. (1999)[13] | yes | yes | yes | no | no | yes | effects only | yes | yes |
| Ludwig et al. (2000)[11] | no | no | yes | viewpoint not clear | viewpoint not clear | yes | irrelevant | no | no |
| Gulbrandsen et al. (2001)[9] | yes | yes | no | yes | yes | yes | costs only | yes | yes |
| Sampson et al. (2001)[17] | yes | no | yes | yes | yes | effects only | no | yes | yes |
| Kouroukis et al. (2003)[10] | yes | yes | no | yes | yes | yes | yes[a] | yes | yes |
| Mehta et al. (2004)[7] | yes | yes | no | yes | yes | costs only | no | yes | yes |
| Van Agthoven et al. (2004)[18] | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Reed et al. (2005)[15] | no | yes | yes | viewpoint not clear | yes | yes | irrelevant | yes | yes |
| Porter et al. (2007)[23] | no | yes | yes | yes | yes | yes | no | no | no |
| Zubair et al. (2009)[24] | no | yes | no | viewpoint not clear | effects only | effects only | irrelevant | yes | no |
| Total 'yes': | 50% | 89% | 50% | 50% | 56% | 50% | 23% | 67% | 56% |

[a]Discounting was incorporated in the sensitivity analysis

[8, 10, 18]. Only one of the three discounted health effects lower than costs to take into account the increase in future value of health [10].

Twelve studies (67%) performed an incremental analysis of both costs and effects [7-10, 12, 13, 15-18, 20, 24]. One study presented the incremental effectiveness between the alternative strategies but did not present the incremental costs [22].

Studies that examined uncertainty in the cost and effect estimates did so by means of a sensitivity analysis. A total of ten studies (56%) performed a sensitivity analysis on pre-determined parameters expected to have an impact on the results [6-10, 13, 15, 17, 18, 20].

Eleven studies (61%) based the conclusions on the interpretation of some overall index or ratio of costs to effectiveness, such as an incremental cost-effectiveness ratio (ICER).

Eleven studies (61%) compared the costs and effectiveness results from their study to those of previous studies [6, 9, 10, 13, 14, 16-18, 20, 22, 24]. In some cases, the authors

**Table 5.** Critique of selected studies according to sub-questions for criteria 10 of Drummond et al. (2005).

| First author | Q10a | Q10b | Q10c | Q10d | Q10e |
|---|---|---|---|---|---|
| Donatini et al. (1994)[19] | yes | no | no | yes | no |
| Laakso et al. (1994)[14] | no | yes | no | yes | no |
| Henon et al. (1995)[16] | yes | yes | yes | no | yes |
| Duncan et al. (1996)[6] | no | yes | no | no | no |
| Jagannath et al. (1997)[22] | no | yes | no | no | yes |
| Nord et al. (1997)[20] | yes | yes | no | yes | no |
| Trippoli et al. (1997)[12] | yes | no | no | no | no |
| Trippoli et al. (1998)[8] | yes | no | yes | yes | no |
| Dranitsaris et al. (1999)[13] | yes | yes | yes | yes | no |
| Ludwig et al. (2000)[11] | no | effects only | no | yes | no |
| Gulbrandsen et al. (2001)[9] | yes | yes | no | yes | no |
| Sampson et al. (2001)[17] | yes | yes | no | yes | no |
| Kouroukis et al. (2003)[10] | yes | yes | yes | yes | no |
| Mehta et al. (2004)[7] | yes | no | no | yes | no |
| Van Agthoven et al. (2004)[18] | yes | yes | yes | yes | no |
| Reed et al. (2005)[15] | no | no | no | yes | no |
| Porter et al. (2007)[23] | no | yes | yes | yes | no |
| Zubair et al. (2009)[24] | no | effects only | no | yes | no |
| Total 'yes': | 61% | 61% | 33% | 78% | 11% |

mentioned that there were no previous studies for comparison [13, 18] or made comparisons to previous studies of other indications or treatments than that under study in the analysis [8].

Discussion of generalizability was considered in six studies (33%) [8, 10, 13, 16, 18, 23]. The majority of these studies questioned the generalizability of the results to other patient groups or health care settings, namely due to differences in care and prognosis among patients in the source population compared to that in the study population.

Most studies (78%) took into account other important factors in the choice of decision under consideration, with four excluding a discussion of such considerations [6, 12, 16, 22]. Important factors included improved protocol [10, 11, 19] [7], the addition of or more precise estimate for costs [10, 13, 17, 23], and administration of treatment as an outpatient procedure or at home instead of as an inpatient [18], end of patent and reduction of time required for infusion of treatment [15], reduction in treatment-related risk for adverse events [24], and the difference in various patient subgroups [14, 20]. Two studies stated that a more relevant comparator for the alternative strategy should have been used [8, 9].

Very few studies (11%) discussed issues of implementation [16, 22]. Neither of the two studies that reported cost-savings with the intervention strategy [23, 24] discussed whether freed resources could be deployed to other worthwhile programs.

## DISCUSSION

Economic evaluations in multiple myeloma have become less frequent in the past decade. As of December 2009, two have become available in the literature since the latest review conducted by Moeremans and Annemans [4]. The past decade has also been marked by few CUAs in this indication, suggesting that a limitation persists in identifying treatment- and/or disease stage-specific utility values.

Evaluations of newer, advanced therapies such as thalidomide, bortezomib and lenalidomide are lacking despite their important contribution to improvements in survival and quality of life [25]. This suggests that evaluations of treatment for multiple myeloma are not keeping pace with the rate of advances in therapy. Given the continued expansion in the availability of alternative therapies as well as combinations of new and existing therapies, the number of economic evaluations in the indication of multiple myeloma should increase. Further advances in gene-expression profiling and single-nucleotide

polymorphism analysis are anticipated to further individualize treatment approaches for patients [26], resulting in much needed assessments into whether an individualized approach leads to improved outcomes and cost-savings. Future analyses should also assess the added value of a drug in which consideration of the sequence of all drugs between diagnosis and death. Such evaluations will be useful in formulating evidence-based guidelines for multiple myeloma in an era characterized by numerous alternative agents.

The quality of economic evaluations in multiple myeloma can be improved. The ability to judge the quality was often difficult because of inadequate description of the analysis, particularly for costs. It was unclear whether this was a result of poor quality in the methodology or documentation. Few studies incorporated standard methods expected in high quality economic evaluations of healthcare, such as assessment of uncertainty in the estimates and discounting. Many studies relied on effectiveness estimates from non-experimental studies that were often based on different patient populations. The uncommon discussion of generalizability and issues of implementation is surprising given their importance for users of the results of economic evaluations. Discussion of generalizability allows users to consider issues of transferability of the results their setting. Further, to assess the feasibility of implementation and redeployment of newer yet more expensive treatment, it is important to discuss any changes in the administration of care and the associated acquisition costs.

Low transparency and methodological weaknesses economic evaluations in the indication of hematology have been reported previously [27, 28]. Weaknesses in methodology and documentation have also been reported in critical reviews of oncology treatment, such as colorectal [29, 30] and breast cancer [31, 32]. Poor quality may not be unique to hematology but more generally a hallmark of economic evaluations in oncology. As recent evidence has demonstrated that healthcare targeting more severe diseases is more likely to be reimbursed [33], the trend for methodological weaknesses and low transparency of the results for economic evaluations in cancer may suggest that disease severity leads to different requirements for methodology. Hence, quality may be overlooked in economic evaluations within oncology. With typically a high cost per QALY threshold in oncology treatment, we argue that the highest quality possible should be required when incorporating the results of economic evaluations in decision-making.

This critical review has implications for those conducting economic evaluations of treatment for multiple myeloma. To ensure that high quality studies are performed, improvements in documentation of the viewpoint and research question being addressed are necessary. A discussion of the generalizability of the results to other settings is also

necessary. Standard methodology should be incorporated, such as discounting and sensitivity analyses, especially given the uncertainty surrounding estimates based on small number of patients and potentially incomparable groups. Lastly, there is a need for more research demonstrating the differences, or absence therefore, in the quality of life of multiple myeloma patients so that utilities can be incorporated into a CUA. This will be necessary for assessment of newer, more expensive drugs that may provide smaller margins of improved effectiveness but a meaningful gain in quality of life.

A limitation of this study is the restriction to peer-reviewed articles. Due to the exclusion of conference abstracts, more recent economic evaluations of newer, more advanced therapies for multiple myeloma may have been discarded. However, the assessment of the quality of such studies is difficult due to the limited information provided in abstracts.

As advances in treatment of multiple myeloma are expected to continue, researchers conducting economic evaluations of advances in therapy as well as users of the results should take into account the implications of this critical review. Users should be aware that the quality of such studies varies. To identify economic evaluations for which the results can be useful for decision-making, the Drummond criteria are useful for judgment of quality. For researchers conducting economic evaluations, it should be realistic to satisfy each of the Drummond criteria. Consideration of the weaknesses reported here will ensure that reimbursement decisions and treatment decisions are based on the highest quality pharmacoeconomic evidence. If the trends in quality continue, the results of economic evaluations for decision-making will be of little value.

## REFERENCES


1. Alexander DD, Mink PJ, Adami HO, et al. Multiple myeloma: a review of the epidemiologic literature. Int J Cancer 2007; 120 Suppl 12: 40-61.
2. Richardson PG, Hideshima T, Mitsiades C, Anderson KC. The emerging role of novel therapies for the treatment of relapsed myeloma. J Natl Compr Canc Netw 2007; 5(2): 149-62.
3. Wisloff F, Gulbrandsen N, Nord E. Therapeutic options in the treatment of multiple myeloma: pharmacoeconomic and quality-of-life considerations. Pharmacoeconomics 1999; 16(4): 329-41.
4. Moeremans K, Annemans L. An update: health economics of managing multiple myeloma. Eur J Cancer 2006; 42(11): 1684-91.
5. Drummond MF, Schulpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. Third ed. Oxford: Oxford University Press; 2005.
6. Duncan N, Hewetson M, Powles R, et al. An economic evaluation of peripheral blood stem cell transplantation as an alternative to autologous bone marrow transplantation in multiple myeloma. Bone Marrow Transplant 1996; 18(6): 1175-8.
7. Mehta J, Duff SB, Gupta S. Cost effectiveness of bortezomib in the treatment of advanced multiple

myeloma. Manag Care Interface 2004; 17(9): 52-61.
8. Trippoli S, Messori A, Becagli P, et al. Treatments for newly diagnosed multiple myeloma: analysis of survival data and cost-effectiveness evaluation (Structured abstract). Oncology Reports; 1998. p. 1475-82.
9. Gulbrandsen N, Wisloff F, Nord E, et al. Cost-utility analysis of high-dose melphalan with autologous blood stem cell support vs. melphalan plus prednisone in patients younger than 60 years with multiple myeloma (Structured abstract). European Journal of Haematology; 2001. p. 328-36.
10. Kouroukis CT, O'Brien BJ, Benger A, et al. Cost-effectiveness of a transplantation strategy compared to melphalan and prednisone in younger patients with multiple myeloma. Leuk Lymphoma 2003; 44(1): 29-37.
11. Ludwig H, Fritz E. Interferon in multiple myeloma: summary of treatment results and clinical implications (Structured abstract). Acta Oncologica; 2000. p. 815-21.
12. Trippoli S, Becagli P, Messori A, Tendi E. Maintenance treatment with interferon in multiple myeloma: survival meta-analysis (Structured abstract). Clinical Drug Investigation; 1997. p. 392-9.
13. Dranitsaris G. Pamidronate for the prevention of skeletal-related events in multiple myeloma: what does the public think it is worth? (Structured abstract). International Journal of Technology Assessment in Health Care; 1999. p. 108-22.
14. Laakso M, Lahtinen R, Virkkunen P, Elomaa I. Subgroup and cost-benefit analysis of the Finnish multicentre trial of clodronate in multiple myeloma. Finnish Leukaemia Group. Br J Haematol 1994; 87(4): 725-9.
15. Reed SD, Radeva JI, Glendenning GA, et al. Economic evaluation of zoledronic acid versus pamidronate for the prevention of skeletal-related events in metastatic breast cancer and multiple myeloma. Am J Clin Oncol 2005; 28(1): 8-16.
16. Henon P, Donatini B, Eisenmann JC, et al. Comparative survival, quality of life and cost-effectiveness of intensive therapy with autologous blood cell transplantation or conventional chemotherapy in multiple myeloma. Bone Marrow Transplant 1995; 16(1): 19-25.
17. Sampson FC, Beard SM, Scott F, Vandenberghe E. Cost-effectiveness of high-dose chemotherapy in first-line treatment of advanced multiple myeloma. Br J Haematol 2001; 113(4): 1015-9.
18. van Agthoven M, Segeren CM, Buijt I, et al. A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma; a prospective randomised phase III study. Eur J Cancer 2004; 40(8): 1159-69.
19. Donatini B, Henon P, Becker M, et al. Peripheral blood stem cell transplantation in multiple myeloma: Clinical benefits and cost analysis. British Journal of Medical Economics 1994; 7(I): 55-62.
20. Nord E, Wisloff F, Hjorth M, Westin J. Cost-utility analysis of melphalan plus prednisone with or without interferon-alpha2b in newly diagnosed multiple myeloma: results from a randomised controlled trial (Structured abstract). Pharmacoeconomics; 1997. p. 89-103.
21. Ludwig H, Cohen AM, Polliack A, et al. Interferon-alpha for induction and maintenance in multiple myeloma: results of two multicenter randomized trials and summary of other studies. Ann Oncol 1995; 6(5): 467-76.
22. Jagannath S, Vesole DH, Zhang M, et al. Feasibility and cost-effectiveness of outpatient autotransplants in multiple myeloma (Structured abstract). Bone Marrow Transplantation; 1997. p. 445-50.
23. Porter CA, Rifkin RM. Clinical benefits and economic analysis of pegylated liposomal doxorubicin/vincristine/dexamethasone versus doxorubicin/vincristine/dexamethasone in patients with newly diagnosed multiple myeloma. Clin Lymphoma Myeloma 2007; 7 Suppl 4: S150-5.
24. Zubair AC, Rymer R, Young J, et al. Multiple myeloma patients receiving large volume leukapher-

esis efficiently yield enough CD34+ cells to allow double transplants. J Clin Apher 2009; 24(1): 6-11.
25. Chanan-Khan AA, Borrello I, Lee KP, Reece DE. Development of target-specific treatments in multiple myeloma. Br J Haematol.
26. Munshi NC, Avet-Loiseau H. Genomics in multiple myeloma. Clin Cancer Res 2011; 17(6): 1234-42.
27. Goor KM, Schaafsma MR, Huijgens PC, van Agthoven M. Economic assessment on the management of chronic lymphocytic leukaemia. Expert Opin Pharmacother 2005; 6(7): 1179-89.
28. van Agthoven M, Uyl-de Groot CA, Sonneveld P, Hagenbeek A. Economic assessment in the management of non-Hodgkin's lymphoma. Expert Opin Pharmacother 2004; 5(12): 2529-48.
29. Garattini L, Compadri PD, Koleva D, et al. A critical review of the full economic evaluations of pharmacological treatments for colorectal cancer. J Med Econ 2008; 11(1): 177-97.
30. Krol M, Koopman M, Uyl-de Groot C, Punt CJ. A systematic review of economic analyses of pharmaceutical therapies for advanced colorectal cancer. Expert Opin Pharmacother 2007; 8(9): 1313-28.
31. Annemans L. Methodological issues in evaluating cost effectiveness of adjuvant aromatase inhibitors in early breast cancer: a need for improved modelling to aid decision making. Pharmacoeconomics 2008; 26(5): 409-23.
32. Blank PR, Dedes KJ, Szucs TD. Cost effectiveness of cytotoxic and targeted therapy for metastatic breast cancer: a critical and systematic review. Pharmacoeconomics 2010; 28(8): 629-47.
33. Koopmanschap MA, Stolk EA, Koolman X. Dear policy maker: have you made up your mind? A discrete choice experiment among policy makers and other health professionals. Int J Technol Assess Health Care 2010; 26(2): 198-204.



## APPENDIX A

### Search strategy

The following search strategy was used in all databases to find relevant articles. Search terms 1 to 3 were used to identify studies involving multiple myeloma patients. Terms 4 to 18 were used to find economic evaluations. Lastly, the results from these three categories were combined in steps 19 to 21.

### Search terms

1. multiple myeloma*
2. plasma cell myeloma*
3. plasma-cell myeloma*
4. economics
5. econom*
6. costs
7. costly
8. costing
9. pharmacoeconomics
10. pharmacoecon*

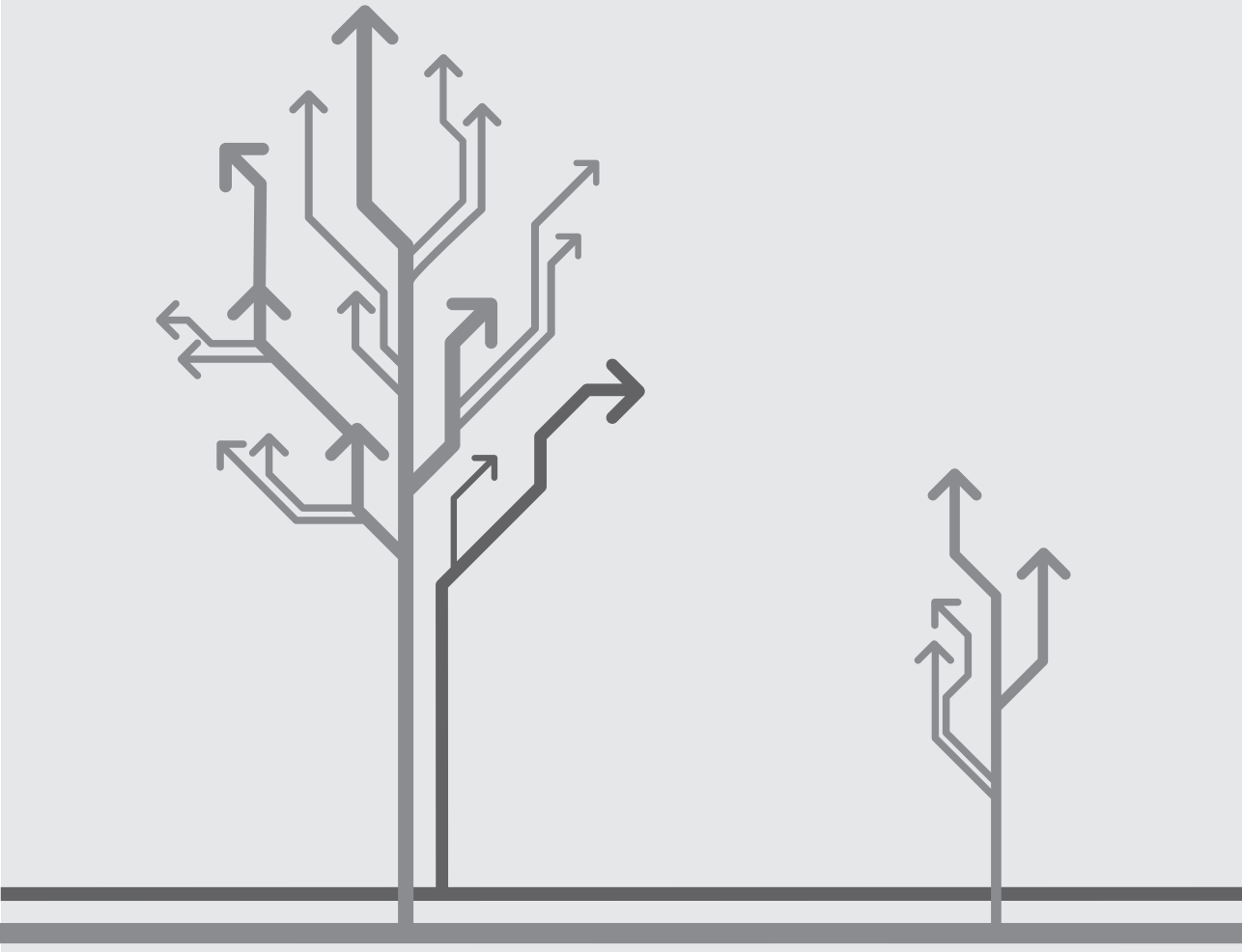
11. budget*
12. expenditure*
13. energy
14 . 12 not 13
15. “value for money”
16. cost-eff*
17. cost-ben*
18. cost-util*
19. 1 or 2 or 3
20. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 14 or 15 or 16 or 17 or 18
21. 19 and 20

# Chapter 3.

## Novel anti-cancer agents for multiple myeloma: A review of the evidence for their therapeutic and economic value

Jennifer G. Gaultney, William K. Redekop, Pieter Sonneveld, Carin A. Uyl-de Groot

*Published in: Expert Reviews in Anticancer Therapy. 2012 Jun;12(6):839-54.*



## ABSTRACT

Recent advances in oncology treatment have improved patient outcomes at the expense of increasing healthcare costs. The indication multiple myeloma is especially characterized by a recent and continuing flood of expensive novel agents. A review encompassing all elements necessary to perform an economic evaluation of novel agents for multiple myeloma was conducted for thalidomide, bortezomib and lenalidomide. Improvements in efficacy have led to a switch from conventional therapy to novel agents as standard therapy. Incremental cost-effectiveness ratios for novel agents alone or in combination with conventional agents were generally regarded to be within acceptable ranges. Conflicting results were reported for the incremental cost-effectiveness of bortezomib versus lenalidomide as unresolved questions remain regarding their comparative effectiveness. Future economic evaluations will require an assessment of the cost-effectiveness of these agents in terms of sequence within the treatment paradigm and in combination with one another.

## INTRODUCTION

Worldwide, the costs of novel anti-cancer therapies challenge patient access to advances in cancer care and controlling its financial impact on healthcare budgets [1]. The recent and continuing flood of advances in treatment of multiple myeloma is an example in oncology where the availability of expensive novel agents has presented a challenge to providers and payers to be certain of their value for money based on randomized controlled trial (RCT) data alone at the time of market access [2].

Currently, there are three novel agents available for treatment of both newly diagnosed and relapsed/refractory multiple myeloma: thalidomide, bortezomib and lenalidomide. Thalidomide and its analog lenalidomide, also known as an immunomodulatory drug (IMiDs), inhibit the progression of multiple myeloma by a number of mechanisms, such as inhibition of interleukin (IL)-6, which is a growth factor for the proliferation of myeloma cells, as well as activation of apoptotic pathways and augmenting the activity of NK-dependent cytotoxicity [3]. Bortezomib is a first-in-class proteasome inhibitor that targets the proteosome-ubiquitin pathway [4], which is known to have a considerable role in tumorigenesis [5].

All three drugs have demonstrated to be effective [6], while their acquisition and treatment-related costs can differ [7]. Based on the most recent published estimates of the average wholesale prices in the US, one 4-week cycle of thalidomide and lenalidomide is $7,362 (100 mg capsules, quantity 28) or $88,3344 USD per year (12 cycles) and $11,447 (25 mg capsules, quantity 21) or $137,364 per year (12 cycles), respectively [8]. In the same study, the cost of one 21-day cycle of bortezomib was estimated to be $6,450 (3.5 mg vial, quantity 4) or $70,950 per patient (11 cycles). These costs do vary per regimen depending on the dose per cycle and number of cycles administered. Generics are not likely to become available for the coming decade as each drug is currently protected under patent, with bortezomib being the first to expire in the year 2022 followed by thalidomide in 2023 and lenalidomide in 2026 [9]. It is expected that second- and third-generation proteasome inhibitors and immunomodulatory agents will be available in the near future [10], further adding to the number of therapy combinations and possibly increased costs.

With increasing availability of expensive novel agents, economic evaluations are needed in multiple myeloma. Elements necessary to conduct economic evaluations include evidence for the clinical effectiveness, impact on quality of life, and costs of care [11]. A review that encompasses all of these elements necessary for economic evaluations of novel anti-cancer agents in multiple myeloma is not available. The aim of this review was

to summarize and evaluate the level of evidence for efficacy, health-related quality of life, and cost-effectiveness available for novel agents in treatment of multiple myeloma.

## METHODS

This review examines published studies assessing the therapeutic value, including the efficacy and safety of each drug, health-related quality of life impact and economic evaluations of each drug. The search was performed via Medline and Embase and included all references available through September 5, 2011 and written in English. Exclusion was initially based on the title and subsequently after reading the abstract. All remaining references were then reviewed. Conference abstracts were excluded. Reviews were used solely for identifying additional sources.

The review of efficacy used the following search term combinations: 'thalidomide AND multiple myeloma', 'lenalidomide AND multiple myeloma', and 'bortezomib AND multiple myeloma'. Results were limited to studies involving humans and pivotal Phase III clinical trial results, while post-hoc analyses of such trials were excluded. For the review of health-related quality of life, the following keywords were used: 'quality of life AND multiple myeloma', 'patient reported outcomes AND multiple myeloma', 'health related quality of life AND multiple myeloma', 'patient experience AND multiple myeloma', 'utility AND multiple myeloma', and 'utilities AND multiple myeloma'. All quality of life studies were limited to novel agents and only studies reporting qualitative measures of health-related quality of life. In the searches for economic literature, the following keywords were used: 'cost AND multiple myeloma', 'economic AND multiple myeloma', and 'pharmacoeconomic AND multiple myeloma'. Only economic evaluations that examined both the costs and health impact of novel agents were included. Costs reported in economic evaluations were converted to Euros and US dollars (USD) for the cost year assumed in the analysis based on historical conversion rates [12].

Selected publications were reviewed according to the treatment phases in the disease paradigm for multiple myeloma: upfront therapy for previously untreated disease, which includes consolidation treatment, and therapy for relapsed/refractory disease. Maintenance therapy was not included in the review as it has yet to be demonstrated as beneficial in terms of survival [13].

**Table 1.** Search results per category

| Category | Full papers | Reason for exclusion (number of studies) |
|---|---|---|
| **I. Efficacy and safety** | | |
| Initial records | 56 | |
| Excluded | 29 | Review (6), Post-hoc analysis (16), Deemed low quality on account of differences in baseline characteristics among treatment groups (1), Retracted (1), Maintenance (1), Intermediate results only (1), Non-inferiority trial (1), Single-arm successor trial (1), Trial description only (1) |
| Included | 27 | |
| *Thalidomide* | 13 | |
| *Bortezomib* | 7 | |
| *Lenalidomide* | 4 | |
| *Combination* | 3 | |
| **II. Quality of life** | | |
| Initial records | 11 | |
| Excluded | 8 | Review (3), Methodology paper not reporting measures of health-related quality of life (4), No novel agents (1) |
| Included | 3 | |
| **III. Economic** | | |
| Initial records | 17 | |
| Excluded | 14 | Review (8), Partial economic evaluation excluding health benefits (6) |
| Included | 3 | |

## RESULTS

### Search

Table 1 describes the results of the search for publications. A total of 26 publications were included in the review of efficacy and safety, with 13 for thalidomide, 7 for bortezomib, 4 for lenalidomide, and 3 for combination therapy consisting of a combination of novel agents. Fewer publications were identified for the health-related quality of life and economic review, with a total of 3 and 3 publications reviewed, respectively.

### Efficacy

#### *Upfront therapy: stem cell transplant ineligible*

A number of Phase III studies have demonstrated evidence for the efficacy of novel agents in patients ineligible for stem cell transplant (SCT) (Table 2).

Thalidomide in newly diagnosed patients has been assessed extensively in patients ineligible for SCT and shown to consistently improve most outcomes when administered

**Table 2.** Studies assessing efficacy of thalidomide in previously untreated multiple myeloma patients ineligible for stem cell transplant

| | | | | | EBMT criteria for response: | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Regimen** | **n** | **Median Age** | **Median FU** | **ORR (%)** | **CR (%)** | **TTP/PFS/EFS (mos)** | **OS (mos)** | **Significant AEs during thalidomide** | **Reference** |
| ≥ 75 | MPT | 113 | NR | 47.5 mos | 62* | 7* | Median PFS: 24.1* | Median 44 mos* | Any grade PN (38 vs 22%) and grade 3 or 4 neutropenia (23 vs 9%*) | [13] |
| | MP | 116 | NR | | 31 | 1 | Median PFS: 18.5 | Median 29.1 mos | | |
| 60-85 | MPT | 129 | 72 | 38.4 mos | 68.9* | 15.6* | Median PFS: 21.8* | Median: 45 mos | Grade 3 or 4 any event (55 vs 22%), thrombosis/embolism (12 vs 2%), neurological (10 vs 1), infections (10 vs 2) | [15-16] |
| | MP | 126 | 72 | 37.7 mos | 47.6 | 3.7 | Median PFS: 14.5 | Median: 47.6 mos | | |
| | MPT | 182 | 75 | 42 mos | 57* | 13* | Median PFS: 15 | Median: 29 | Any grade constipation (43 vs 16%), neuropathy (27 vs 7%), non-neuropathy neurologic toxicity (22 vs 13%), exanthema (11 vs 4%), and non-hematologic (87 vs 68%) | [17] |
| | MP | 175 | 74 | | 40 | 4 | Median PFS: 14 | Median: 32 | | |
| ≥ 65 | MPT | 165 | 72 | 39 mos | 66* | NR | Median EFS: 13* | Median: 40 mos* | Any grade 3 or 4 events (50 vs 29%), PN (23 vs 4%) and DVT (3 vs 0%) | [18] |
| | MP | 168 | 73 | | 45 | NR | Median EFS: 9 | Median: 31 mos | | |
| All ages | TD | 145 | 72 | 28.1 mos | 66* | 2 | Median PFS / TTP: 16.7 / 21.2 | Median: 41.5 | Any grade neuropathy (72 vs 33%), constipation (33 vs 13%), psychological disturbances (37 vs 19%) | [14] |
| | MP | 143 | 72 | | 48 | 2 | Median PFS/ TTP: 20.7 / 29.1 | Median: 49.4 | | |

*Statistically significant difference between treatment groupsAE: adverse event; CR: complete response; D: dexamethasone; DVT: deep vein thrombosis; EBMT: European Group for Blood and Marrow Transplant; EFS: event-free survival; FU: follow-up; mos: months; MP: melphalan prednisone ; MPT: melphalan prednisone thalidomide; NA: not available; NR: not reported; ORR: overall response rate; OS: overall survival; PFS: progression-free survival; PN: peripheral neuropathy; TD: thalidomide dexamethasone; TTP: time-to-progression.

**Table 3.** Studies assessing efficacy of bortezomib and lenalidomide in previously untreated multiple myeloma patients ineligible for stem cell transplant

| Novel agent/ Trial | Regimen | n | Median Age | Median FU | EBMT criteria for response post-induction ORR (%) | CR (%) | TTP/PFS | OS | Significant AEs during thalidomide | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| *Bortezomib* | | | | | | | | | | |
| VISTA trial | BMB+MP | 344 | 71 | 36.7 mos | 71* | 30* | Median TTP: 24 mos* | Median: NR; 3-yr: 68.5% | Grade 3 events: 53 vs 44%; any grade PN (44 vs 5%); grade 3 or 4 gastrointestinal events (19 vs 5%) | [19-20] |
| | MP | 338 | 71 | | 35 | 4 | Median TTP: 16.6 mos | Median: 43 mos; 3-yr: 54% | | |
| *Combination* | | | | | | | | | | |
| GIMEMA trial | VMPT-VT | 254 | 71 | 23.2 mos | 89* | 38* | 3-yr PFS: 56%* | 3-yr: 89% | Grade 3/4 events VMPT-VT vs VMP: neutropenia (38 vs 28%), cardiologic events (10 vs 5%), and thromboembolic events (5 vs 2%) | [22] |
| | VMP | 257 | 71 | | 81 | 24 | 3-yr PFS: 41% | 3-yr: 87% | | |
| PETHEMA | VMP induction | 130 | 73 | 32 mos | 80 | 20 | Median PFS: 34 mos | 3-yr: 74% | Grade 3 or 4 events VTP vs VMP: treatment-related serious adverse events (31 vs 15%), patients discontinuing due to serious adverse event (17 vs 12%) and cardiac events (8 vs 0%) | [21] |
| | VTP induction | 130 | 73 | | 81 | 28 | Median PFS: 25 mos | 3-yr: 65% | | |
| *Lenalidomide* | | | | | | | | | | |
| SWOG SO232 trial | Len/Dex | 97 | 49% ≥ 65 | 47.2 mos | 78* | 26* | 1-yr PFS: 78%*; 3-yr PFS: 52%* | 3-yr: 79% | Neutropenia (42 vs 18%) and thrombocytopenia (34 vs 20%) | [23] |
| | Dex alone | 95 | 47% ≥ 65 | | 48 | 4 | 1-yr PFS: 52%; 3-yr PFS: 32% | 3-yr: 73% | | |

*Statistically significant difference between treatment groups. AE: adverse event; D: dexamethasone; CR: complete response; EBMT: European Group for Blood and Marrow Transplant; FU: follow-up; GIMEMA: Gruppo Italiano Malattie EMatologiche dell'Adulto; mos: months; MP: melphalan prednisone ; NA: not available; NR: not reported; ORR: overall response rate; OS: overall survival; PFS: progression-free survival; PN: peripheral neuropathy; RD: lenalidomide dexamethasone; TTP: time-to-progression; VMP: bortezomib melphalan prednisone; VMPT-VT: bortezomib melphalan prednisone thalidomide followed by maintenance with bortezomib thalidomide; vs: versus; VTP: bortezomib thalidomide prednisone; yr: year.

in combination with conventional therapy consisting of melphalan and prednisone (MPT) or dexamethasone (TD) versus conventional therapy consisting only of melphalan and prednisone (MP). Overall response rates (ORR) and complete response (CR) rates for MPT and TD in this setting were superior compared to MP and ranged from 57% to 69% and from 2% to 15.6%, respectively [14-19]. Median progression-free survival (PFS) and overall survival (OS) were superior for MPT compared to MP alone [14, 16-19], while TD was not shown to be superior than MP alone [15].

Bortezomib in similar patients included in the Velcade as Initial Standard Therapy (VISTA) trial [20, 21] resulted in high ORRs in this setting (71%) as well as high CR rates (30%) (Table 3). Median time-to-progression (TTP) for bortezomib was 24 months while median OS has not been reached for the VISTA trial but three-year OS was 68.5%.

The combination of bortezomib and thalidomide as first-line therapy for patients ineligible for SCT has shown the most promising results in terms of response with an ORR and CR rate ranging from 81-89% and 28-38% [22, 23]. However, it is difficult to conclude whether bortezomib therapy in combination with thalidomide is favorable compared to bortezomib without thalidomide in terms of time-to-event outcomes. Namely, the combination of bortezomib plus melphalan and prednisone (VMP) improved median PFS and 3-year OS compared to bortezomib plus thalidomide and prednisone (VTP), though this was not found to be statistically significant (34 versus 25 months; 74% versus 65%) [22]. In a separate study, VMPT-VT significantly improved progression-free survival (PFS) compared to VMP though similarly it was found that there was no significant difference in OS [23].

Lenalidomide has also shown to be effective as first-line therapy to previously untreated patients mostly ineligible for SCT or high-dose therapy [24]. High ORR and CR rates were observed, 78% and 26%, respectively, and the three-year OS was 79%.

***Upfront therapy: stem cell transplant eligible***

The use of each novel agent has been assessed for efficacy in previously untreated patients eligible for SCT (Table 4). In eligible patients, the addition of thalidomide to dexamethasone was shown to be superior to dexamethasone alone, with significantly improved ORR and CR rates [25, 26]. Time-to-event results for the comparison of TD to dexamethasone alone were only available for one of the two studies, which demonstrated an improvement of 16.1 months of survival spent without progression (median 22.6 versus 6.5 months; $P < 0.001$) [26]. Thalidomide was further shown to result in more favorable response rates in older patients when administered prior to SCT as MPT [27] than when administered in combination with adriamycin and dexamethasone (TAD) as

**Table 4.** Studies assessing efficacy of novel agents in previously untreated multiple myeloma patients eligible for stem cell transplant

| Novel agent/ Age/Trial | Regimen | n | Median (mean) Age | Median FU (range) | EBMT criteria for response post-induction ORR (%) | CR (%) | TTP/PFS/EFS | OS | Significant AEs | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| *Thalidomide* | | | | | | | | | | |
| | TD | 103 | 65 | NA | 4-mos ORR: 63%* | | NA | NA | Grade 3 or higher events (45 vs 21%), including DVT (17 vs 2%), rash (4 vs 0%), bradycardia (1 vs 0%), neuropathy (7 vs 4%) | [24] |
| | D | 104 | 65 | | 4-mos ORR: 41% | | NA | NA | | |
| | TD | 235 | (64) | 17 mos | 63* | 7.7* | Median TTP: 22.6 mos* | NA | Grade 3 or higher events (79.5 vs 64.2%) and grade 4 events (30.3 vs 22.8%) | [25] |
| | D | 235 | (64) | 18 mos | 46 | 2.6 | Median TTP: 6.5 mos | NA | | |
| 65-75 | MPT | 125 | 40% ≥70 | 36.8 mos | 76* | 13* | Median PFS: 27.5* mos | Median: 51.6* | Grade 3 or 4 PN (6 vs 0%), somnolence/ fatigue/dizziness (8 vs 0%), constipation (10 vs 0%) | [26] |
| | MP | 196 | 43% ≥70 | | 35 | 2 | Median PFS: 17.8 mos | Median: 33.2 | | |
| ≤75 years | (TT2):TAD+tandem SCT | 323 | NR | 72 mos | 60 | 19* | Median EFS: 72 mos* | 8-yr: 57% | > grade 2: thrombosis/embolism (30 vs 17%), syncope (12 vs 4%), bowel obstruction (14 vs 8%), tremor (13 vs 6%), PN (27 vs 17%) | [27-28] |
| | AD+tandem SCT | 345 | NR | | 40 | 10 | Median EFS: 49.2 mos | 8-yr: 44% | | |
| 18-65 years | TAD | 268 | 56 | 52 mos (2-86) | 71* | 3 | Median PFS: 34* mos | Median: 73 mos | Any grade 3-4 event (49 vs 37%), especially grade 3-4 gastrointestinal (10 vs 5%) and neurology (13 vs 7%) | [29] |
| | VAD | 268 | 57 | | 57 | 2 | Median PFS: 25 mos | Median: 60 mos | | |

**Table 4.** (Continued)

| Novel agent/ Age/Trial | Regimen | n | Median (mean) Age | Median FU (range) | EBMT criteria for response post-induction ORR (%) | CR (%) | TTP/PFS/EFS | OS | Significant AEs | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| Consolidation | Thalidomide-prednisolone | 114 | 57 | 36 mos | Significant improvement in depth of response from post-SCT to post-consolidation for Thalidomide arm | | 3-yr PFS: 42%* | 3-year: 86%* | Grade 3 or 4 events PN (10 vs 0%), bowel obstruction (4 vs 0%), mood disturbance (15 vs 7%), fatigue/lethargy (14 vs 2%) and other neurological toxicities (11 vs 2%) | [30] |
| | Prednisolone | 129 | 57 | | | | 3-yr PFS: 23% | 3-year: 75% | | |
| *Bortezomib* | | | | | | | | | | |
| IFM 2005-01 | VD | 240 | 56 | 32.2 mos | 79* | 5.8 | Median PFS: 36 mos | 3-yr: 81.4% | Grade 2 (20.5 vs 10.5%) and grades 3 to 4 (9.2 vs 2.5%) PN. | [31] |
| | VAD | 242 | 57 | | 63 | 1.4 | Median PFS: 29.7 mos | 3-yr: 77.4% | | |
| *Combination* | | | | | | | | | | |
| GIMEMA trial | VTD | 236 | 58 | 36 mos (22-42) | 93* | 19* | 3-yr PFS: 68%* | 3-yr: 86% | Grade 3-4 events: any event (56 vs 33%), PN (10 vs 2%) | [32] |
| | TD | 238 | 57 | | 79 | 5 | 3-yr PFS: 56% | 3-yr: 84% | | |

*Statistically significant difference between treatment groups. AD: adriamycin dexamethasone; AE: adverse event; D: dexamethasone; CR: complete response; DVT: deep vein thrombosis; EBMT: European Group for Blood and Marrow Transplant; EFS: event-free survival; FU: follow-up; GIMEMA: Gruppo Italiano Malattie EMatologiche dell'Adulto; IFM: International Myeloma Foundation; mos: months; MP: melphalan prednisone ; MPT: melphalan prednisone thalidomide; NA: not available; NR: not reported; ORR: overall response rate; OS: overall survival; PFS: progression-free survival; PN: peripheral neuropathy; SCT: stem cell transplantation; TAD: thalidomide adriamycin dexamethasone; TD: thalidomide dexamethasone; TTP: time-to-progression; VAD: bortezomib adriamycin dexamethasone; VD: bortezomib dexamethasone; vs: versus; VTD: bortezomib thalidomide dexamethasone; VTP: bortezomib thalidomide prednisone; yr: year.

part of Total Therapy 2 (ORR: 76% versus 60%) [28, 29]. However, time-to-event estimates favor Total Therapy 2 consisting of TAD compared to MPT, with a median event-free survival of 72 months for compared to a median PFS of 27.5 months, respectively. A more recent trial enrolling younger patients aged 18-65, the Haemato Oncology Foundation for Adults in the Netherlands (HOVON)-50 trial, found TAD achieved a significant improvement in response rates and PFS compared to VAD, while OS was not significantly different [30].



Consolidation therapy with thalidomide immediately after single SCT was also reported by Spencer et al. (2009) on behalf of the investigators of the phase III Australian Leukemia and Lymphoma Group (ALLG) MM6 trial [31]. Therapy consisted of alternate day prednisolone alone or in addition to thalidomide daily for a maximum of twelve months. Patients in the thalidomide arm achieved significantly higher rates of response after 12 months of maintenance therapy and an extended 3-year PFS and OS from randomization.

Bortezomib as induction therapy in SCT eligible patients resulted in improved response rates (ORR: 79% versus 60-76%) as well as time-to-event estimates (median PFS: 36 months versus 27.5-34 months) compared to TAD and MPT [32]. Further still, bortezomib in combination with thalidomide showed the most favorable results as upfront therapy in these patients [33], with an ORR of 93% and CR rate of 19% compared to the range of 60-79% and 3-19% for bortezomib or thalidomide alone. Time-to-event results are still too early but available data suggest improvements in these outcomes as well.

***Therapy for relapsed/refractory disease***

The evidence for thalidomide in the relapsed/refractory patients from Phase III trials is lacking. However, evidence in this setting is available for bortezomib and lenalidomide (Table 5).

Bortezomib has been demonstrated to be more effective than dexamethasone in the Assessment of Proteasome Inhibition for Extending Remissions (APEX) trial [34, 35] for relapsed/refractory patients. Specifically, patients treated with bortezomib versus dexamethasone experienced a median TTP and OS of 6.2 and 29.8 months versus 3.5 and 23.7 months, respectively. ORR was also significantly different at 43% compared to 18%, respectively. Bortezomib was found to be more effective when combined with pegylated liposomal dexamethasone (PLD) compared to bortezomib alone in terms of time-to-event estimates such as median PFS (9.3 versus 6.5 months) and 15-month OS (76% versus 65%), though no significant differences in ORR and CR (44 versus 41% and 4% versus 2%, respectively) rates were demonstrated [36]. Mikhael et al. (2008) reported

that bortezomib alone or in combination with dexamethasone is safe and effective in heavily pre-treated MM patients, with favorable ORR (50.9% versus 43% ) comparable to that reported in the APEX trial [37].

As for lenalidomide in the relapsed/refractory setting, both the MM-09 trial conducted in North America and the international MM-010 trial demonstrated lenalidomide in combination with dexamethasone (RD) to be superior to dexamethasone alone, with a higher ORR (60.6% versus 21.9%) and median TTP (13.4 versus 4.6), as well as OS (38 versus 31.6 months) [38-40].

**Safety**

***Thalidomide***

Common grade 3 or 4 treatment-related adverse events (AEs) in patients treated with thalidomide included constipation, neuropathy, somnolence, depression, and venous thromboembolism (VTE). The frequency varied depending on whether it is given in combination or as monotherapy. Thalidomide in combination generally increased the risk for AEs. In newly diagnosed patients, the addition of thalidomide to other active agents generally increased the incidence of AEs associated with thalidomide, including constipation, neuropathy, somnolence and VTE. When given as consolidation therapy in comparison to prednisolone, thalidomide is associated with increased incidence of grade 3 or 4 neuropathy and constipation [31]. In this same study, VTE was not significantly more common likely due to the low dose of thalidomide and the low tumor burden.

***Bortezomib***

Based on the incidence of AEs found in Phase III trials, the most common AEs related to bortezomib treatment included hematological events, such as thrombocytopenia and neutropenia, and PN. As initial treatment, VMP was associated with similar rates of hematological AEs and a higher incidence of PN compared to MP. All grade 3 or 4 gastrointestinal AEs were more common in the VMP group than in the MP group (18 versus 5%) as well as herpes zoster (13 versus 4%). For patients with relapsed/refractory disease, the APEX trial demonstrated the most relevant grade 3 or 4 AEs in the bortezomib arm to be PN (8%), thrombocytopenia (30%), and neutropenia (14%) [34]. Other AEs were also found to be associated with discontinuation of treatment, including gastrointestinal disorders, fatigue, hypercalcemia, and spinal cord compression. Some combinations of bortezomib present increased risk for additional AEs, such as the addition of pegylated liposomal doxorubicin for treatment of relapsed/refractory disease which resulted in a 5% incidence of hand-foot syndrome [36]. In addition, the combination of thalidomide

**Table 5.** Studies assessing efficacy of novel agents in relapsed/refractory multiple myeloma patients

| Novel agent/ Trial | Regimen | n | Median age | Median no. prev therapies | Median FU | EBMT criteria for response: ORR (%) | CR (%) | TTP | OS | Significant AEs | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bortezomib* | | | | | | | | | | | |
| APEX trial | V | 333 | 62 | 2 | 22 mos | 43* | 9* | Median TTP: 6.22 mos* | Median: 29.8 mos* | Grade 3: 61 vs 44%, grade 3 diarrhea (7 vs 2%), PN (7 vs 1%), thrombocytopenia (26 vs 5%), neutropenia (12 vs 1%) | [33-34] |
| | D | 336 | 61 | 2 | | 18 | 1 | Median TTP: 3.49 mos | Median: 23.7 mos | | |
| | VPLD | 318 | 62 | 66% ≥ 2 | NR | 44 | 4 | Median TTP: 9.3 mos | 15-month: 76% | Any drug-related AE (86 vs 94%); grade 3 or 4 (64 vs 80%); drug-related grade 3 or 4 (52 vs 68%); nausea (37 vs 46%); diarrhea (34 vs 43%); vomiting (19 vs 31%); neutropenia (20 vs 35%); pyrexia (22 vs 29%); anorexia (11 vs 18%); stomatitis (3 vs 18%); hand-foot syndrome (0 vs 16%) | [35] |
| | V | 318 | 62 | 66% ≥ 2 | | 41 | 2 | Median TTP: 6.5 mos* | 15-month: 65%* | | |
| Global phase 3b expanded access program | V +/- D | 638 (208 with VD) | 62.7 | 3 | Median 5 cycles (5 mos) | 50.9 | 11.2 | NR | NR | Thrombocytopenia (39%), neutropenia (16%), anemia (12%), diarrhea (7%), PN (6%) | [36] |
| *Lenalidomide* | | | | | | | | | | | |
| Pooled analysis MM-009/MM-010 trials | RD | 353 | 63 | 2 | 48 | 60.6* | 15* | Median TTP: 13.4* mos | Median: 38* mos | Any grade 3 or 4 events (83.3 vs 69.7%), neutropenia (35.4 vs 3.4%), thrombocytopenia (13 vs 6.6%), thrombotic events (15.9 vs 5.4%), anemia (10.8 vs 6) | [37-39] |
| | D | 351 | 63 | | | 21.9 | 2 | Median TTP: 4.6 mos | Median: 31.6 mos | | |

*Statistically significant difference between treatment groups. AE: adverse event; APEX: Assessment of Proteasome Inhibition for Extending Remissions; D: dexamethasone; CR: complete response; EBMT: European Group for Blood and Marrow Transplant; FU: follow-up; mos: months; NA: not available; NR: not reported; ORR: overall response rate; OS: overall survival; PFS: progression-free survival; PN: peripheral neuropathy; RD: lenalidomide dexamethasone; TTP: time-to-progression; V: bortezomib; VPLD: bortezomib plus pegylated liposomal doxorubicin; vs: versus; yr: year.

and bortezomib presents a trade-off of a higher risk for serious cardiac events which should carry weight when making treatment decisions [23].

***Lenalidomide***

Although lenalidomide is an analog of thalidomide, it has a different adverse event profile compared to thalidomide. Serious AEs associated with thalidomide, such as peripheral neuropathy (PN), constipation and somnolence, were uncommon in patients treated with lenalidomide. Common treatment-related AEs in patients treated with lenalidomide included myelosuppression and VTE. Signs of myelosuppression included neutropenia, thrombocytopenia and, to a lesser extent, anemia. VTE is particularly associated with lenalidomide in combination with dexamethasone in the absence of thromboprophylaxis. The pooled analysis of the two Phase III trials of lenalidomide in combination with dexamethasone reported the most frequent grade 3 or 4 AEs to be neutropenia (35%), anemia (11%), thrombocytopenia (13%), infection (16%), atrial fibrillation (3%) and VTE (13%), which was most likely related to the fact that thromboprophylaxis was not mandated in this trial. An increased risk of skin toxicities with

**Table 6.** Health-related quality of life studies conducted for novel agents in relapsed/refractory multiple myeloma

| Novel agent | Comparator | Trial | PRO Questionnaire | Significant differences observed in novel agent group | Reference |
|---|---|---|---|---|---|
| MPT | MP | HOVON-49 | EORTC QLQ-C30 and QLQ-MY24 | Induction: Deterioration in physical function and constipation.<br>Maintenance: Deterioration in scored for paresthesia. Improved QLQ-C30 scores for pain, insomnia, appetite loss, emotional function and future perspectives, as well as QLQ-MY24 items sick. | [41] |
| V | None | SUMMIT | EORTC QLQ-C30, QLQ-MY24, FACIT and FACT/GOG-NTX | Improved fatigue scores in patients with CR or PR. Mostly stable scores in patients with minor response or no change, and deterioration in most scores for progressive patients. Changes in scores for neuropathy-related symptoms were stable. | [42] |
| V | HDD | APEX | EORTC QLQ-C30 and FACT/GOG-NTX | Improved global health status, better physical health, role, cognitive, emotional functioning scores and NTX. Deterioration in dyspnea and sleep symptoms scores. | [43] |

APEX: Assessment of Proteasome Inhibition for Extending Remissions; CR: complete response; EORTC: European Organization for Research and Treatment of Cancer; FACIT: Functional Assessment of Chronic Illness Therapy ; FACT: Functional Assessment of Cancer Therapy; GOG: Gynecologic Oncology Group ; HDD: high-dose dexamethasone; HOVON: Haemato Oncology Foundation for Adults in the Netherlands; MPT: melphalan prednisone thalidomide; MP: melphalan prednisone; NTX: neurotoxicity scale; PR: partial response; PRO: patient-reported outcomes; QLQ-C30: quality of life questionnaire for cancer version 3; QLQ-MY24: quality of life questionnaire myeloma-specific module; SUMMIT: Study of Uncontrolled Multiple Myeloma Managed with Proteasome Inhibition Therapy; V: bortezomib.

lenalidomide has also been reported [41] and was added to the list of precautions section of its labeling in 2009.

**Health-related quality of life**

Three studies are available in the literature reporting health-related quality of life data for multiple myeloma patients treated with novel agents, one for thalidomide as upfront treatment[42] and two for bortezomib in relapsed/refractory treatment [43, 44] (Table 6).

***Previously untreated patients***

Alongside the HOVON-49 study, which assessed MPT versus MP in newly diagnosed elderly patients [19], all patients completed the European Organization for Research and Treatment of Cancer (EORTC) quality of life questionnaire (QLQ)-C30 and the myeloma-specific module (QLQ-MY24). Quality of life was measured at baseline and then 3, 9, 12 and 18 months after start of cycle 1. Results showed that both arms improved patient-reported outcome scores for overall global health, fatigue, pain, side effects, insomnia and appetite loss [42]. However, during thalidomide maintenance, MPT showed improved scores for pain, insomnia, appetite loss and sick. MPT was however associated with greater risk for parasthesis consistent with a dose effect. The authors concluded that despite greater toxicities with MPT, it was not reflected in the overall QLQ-C30 scores.

***Relapsed/refractory patients***

Using the Study of Uncontrolled Multiple Myeloma Managed with Proteasome Inhibition Therapy (SUMMIT) trial data which was a Phase II trial assessing single agent bortezomib in relapsed/refractory patients, it was demonstrated that changes in patient-reported outcome scores measured by the EORTC QLQ-C30, QLQ-MY24, Functional Assessment of Chronic Illness Therapy (FACIT) and Functional Assessment of Cancer Therapy/Gynecologic Oncology Group (FACT/GOG) neurotoxicity scale (NTX) showed statistically significant differences between response groups [43]. Improvements in patient-reported outcomes were seen in patients with CR or PR; mostly stable scores were seen in patients with minor response or no change; and deterioration was observed in most scores for progressive patients. Changes in scores for neuropathy-related symptoms were stable. Fatigue scores significantly improved with CR and PR. When looking at various minimal important difference thresholds, the proportion of patients improved exceeded 35% for several domains within all change group definitions. In summary, this study demonstrated the complementary value of patient-reported outcome assessments in further interpreting clinical response, impact of AEs, and patient prognosis in trials for multiple myeloma.

**Table 7.** Economic evaluations of novel agents in multiple myeloma assessing both costs and health benefits

| Comparators | Country | Perspective | Cost year | Discounting | Efficacy data source | Total costs and health effects | ICER | Reference |
|---|---|---|---|---|---|---|---|---|
| V vs BSC vs T | US | Healthcare payer | 2003 | None | Phase II trial and Delphi panel | V (full cohort): 1.33 LYs, $65,222 (€57.656);<br>V (received previous T): 1.31 LYs, $69,200 (€61,173);<br>V (no previous T): 2.33 LYs, $68,816 (€60,833);<br>BSC: 0.21 LYs, $14,423 (€12,750);<br>T: 0.7 LYs, $37,265 (€32,942) | V (full cohort) vs BSC: $45,356 (€40,095)/LY;<br>V (received previous T) vs BSC: $49,797 (€44,021)/LY;<br>V (no previous T) vs Thal: $21,483 (€18,991)/LY | [44] |
| VD vs RD or D | Sweden | NR | 2010 | 3% costs and effects | Phase III trial | VD: 4.78 LYs, 2.94 QALYs , SEK1,904,462 ($265,337 or €199,003);<br>RD: 4.51 LYs or 2.91 QALYs, SEK2,450,588 ($342,608 or €256,956);<br>D: 3.72 LYs, 2.26 QALYs, SEK1,278,854 ($178,790 or €134,093) | VD vs D: SEK 590,029 ($82,489 or €61,867) /LY,<br>SEK 902,874 ($126,228 or €94,671) /QALY;<br>VD vs RD: dominant | [45] |
| RD vs VD | Norway | Third-party payer | NR (assume 2010) | 4% costs and effects | Phase III trial | RD: 4.06 LYs, 2.95 QALYs, NOK689,207 ($114,868or €86,151);<br>VD: 3.11 LYs, 2.19 QALYs, NOK500,962 ($83,493 or €62,620) | RD vs VD: NOK198,714 ($33,119 or €24,839)/LY, NOK247,978 ($41,329or €30,997)/QALY. | [46] |

BSC: best supportive care; D: dexamethasone; ICER: incremental cost-effectiveness ratio; LYs: life-years; NOK: Norwegian krone; NR: not reported; QALYs: quality-adjusted life-years; RD: lenalidomide dexamethasone; SEK: Swedish krone; T: thalidomide; US: United States; V: bortezomib.

Phase III APEX trial patients were also surveyed using the EORTC QLQ-C30 and FACT-NTX questionnaires, administered at baseline and every 6 weeks up to 42 weeks [44]. Bortezomib patients showed significantly higher global health status compared to dexamethasone patients. Bortezomib was also associated with better NTX score, despite higher rates of PN. Improved health-related quality of life for bortezomib treated patients was shown to be partially explained by improved survival.

**Economic evaluations**

Three published economic evaluations from the US and the EU have been published assessing the cost-effectiveness of novel agents in relapsed/refractory multiple myeloma: one comparing bortezomib to thalidomide or best supportive care (BSC) in the US [45] and two comparing bortezomib to lenalidomide, both in combination with dexamethasone, in Sweden [46] and Norway [47] (Table 7).



From the US payer perspective, bortezomib alone in patients with relapsed/refractory disease was compared to BSC based on the results of a Phase II trial. Irrespective of previous treatment, bortezomib treatment resulted in a gain of 1.33 life-years (LYs) at a total lifetime cost of $65,222 (USD 2003) or €57,656 (Euro 2003) compared with a gain of 0.21 LYs at a total lifetime cost of $14,423 (USD 2003) or €12,750 (Euro 2003) for BSC [45]. For bortezomib in patients with previous thalidomide treatment, the results were similar with a gain of 1.31 LYs at a total lifetime cost of $69,200 (USD 2003) or €61,173 (Euro 2003). The costs and health benefits for bortezomib in thalidomide-naïve patients were the most favorable with a gain of 2.33 LYs at a cost of $69,200 (USD 2003) or €60,833 (Euro 2003) compared to a gain of 0.7 LYs at a cost $37,264 (USD 2003) or €32,942 (Euro 2003) for thalidomide. For less than twice the estimated lifetime costs of thalidomide, bortezomib treatment prior to thalidomide in the relapsed/refractory setting was estimated to result in more than three times the health benefits as that with thalidomide. The results for bortezomib were therefore concluded to be within an acceptable range of cost-effectiveness. It is important to note, however, that the effectiveness assessments for BSC and thalidomide therapy were not based on actual patient-level data but instead the responses obtained from a Delphi panel of experts.

Two indirect comparisons of bortezomib to lenalidomide have been conducted. The first study was conducted from the perspective of Sweden, which found bortezomib plus dexamethasone to be dominant compared to RD by means of a partitioned-survival analysis modeling study using the results of the Phase III APEX and MM-010/MM-011 trials [46]. The results showed that bortezomib plus dexamethasone resulted in 4.78 LYs or 2.94 quality-adjusted LYs (QALYs) at a total lifetime cost per patient of $265,337 (USD 2010) or €199,003 (Euro 2010). RD resulted in fewer health benefits and higher costs

at 4.51 LYs or 2.91 QALYs at a total lifetime cost per patient of $342,608 (USD 2010) or €256,956 (Euro 2010). Dexamethasone alone was less costly but resulted in fewer health benefits with 3.72 LYs or 2.26 QALYs at a total lifetime cost per patient of $178,790 (USD 2010) or €134,093 (Euro 2010). The results were found to be sensitive to a number of parameters, including: (i) utility prior to and following relapse, (ii) bortezomib chemotherapy costs, and (iii) number of bortezomib administrations.

A more recent study conducting the same comparisons was performed from the third-party payer perspective of Norway by means of a discrete event simulation (DES) [48] also using the APEX and MM-009/MM-010 trial results [47]. The results for health benefits were in contrast to that found in the previous study as RD was found to result in greater LYs and QALYs compared to bortezomib plus dexamethasone (4.06 LYs and 2.95 QALYs versus 3.11 LYs and 2.19 QALYs, respectively). Costs in Norway were also higher for RD compared to bortezomib plus dexamethasone, with total lifetime costs per patient of $114,868 (USD 2010) or €86,151 (Euro 2010) and $83,493 (USD 2010) or €62,620 (Euro 2010), respectively. Results were sensitive to the four-year time horizon and post-progression survival.

## DISCUSSION

Advances in treatment have resulted in a switch from conventional therapy to novel agents as standard therapy for multiple myeloma. Each novel agent has been demonstrated to be more effective compared to conventional therapy as upfront therapy. The most favorable results as upfront therapy have been shown for combination therapy of two or more drugs. In relapsed/refractory patients, lenalidomide and bortezomib are more effective compared to conventional agents, while no Phase III trials have been reported for thalidomide in this setting. No direct evidence that one is safer and more effective than the other has been demonstrated but more generally that one may be more appropriate for a particular patient group at a particular stage of the disease.

For economic evaluations, it is important to address the question of whether there is an improvement or a loss in patient-reported outcomes for novel agents, as it may be that case future agents will lead to improved quality of life rather than extended life. Patient-reported outcome studies have demonstrated improvements in health-related quality of life during thalidomide maintenance therapy and for relapsed/refractory patients responding to bortezomib, while evidence is needed to establish an improvement with lenalidomide. The findings that the impact of toxicities was not reflected in the quality of life measures for either thalidomide or bortezomib suggests that the questionnaires

currently used may not be sensitive enough to detect their impact on loss of quality of life. Nonetheless, the results of these studies are useful for the conduct of a cost-utility analysis, as they can be used to generate a summary measure (i.e., utility) that represent the gains and loss in quality of life along with the survival gains associated with a particular treatment within a particular patient group (i.e., QALYs).

Despite the availability of these patient-reported outcome studies, there remain few studies reporting utility values. The utilities that have been used in cost-utility analyses [46, 47] were all taken from Van Agthoven et al. (2004), which was a study conducted for treatment of multiple myeloma prior to the availability of novel agents [49]. Even the analyses conducted for market access in the UK submitted by the manufacturers of bortezomib and lenalidomide also relied on the utility estimates from Van Agthoven et al. (2004) in their cost-utility analyses [50, 51]. Given the evidence from the trials on bortezomib for a significant improvement in responding patients, health-related quality of life in future studies should be assessed according to response to a particular treatment, as it may be more relevant to apply different utility values according to response to therapy, especially when comparing novel agents that have similarly favorable efficacy. Fortunately, research is underway addressing this need by evaluating methods for deriving utilities for cancer from the EORTC QLQ-C30 [52], which is the most commonly used health-related quality of life survey method in cancer clinical trials in Europe and Canada. Further research should also address whether the cancer-specific questionnaire QLQ-C30 is adequately sensitive for measuring health-related quality of life in multiple myeloma patients and whether a disease-specific is more appropriate.

There is little information about the cost, let alone the cost-effectiveness, of novel agents for multiple myeloma outside of the US and EU in the published literature. The few economic evaluations that are available were conducted from the perspective of the healthcare systems in the US and the EU. No economic evaluations were available for upfront therapy and only three available for relapsed/refractory therapy. The economic value of these novel agents for treatment of multiple myeloma patients outside of the US and EU as well as for their use as upfront therapy in newly diagnosed patients represent important avenues that deserve further investigation.

The incremental cost-effectiveness ratio (ICER) for each drug relative to the other have demonstrated to be within acceptable threshold limits, though the evidence for bortezomib compared to lenalidomide is still questionable. From the US perspective, the cost-effectiveness ratios were more or less similar for thalidomide and bortezomib [45], while evidence from comparisons of bortezomib and lenalidomide in combination with dexamethasone found conflicting results. For bortezomib versus thalidomide,

bortezomib in patients previously untreated with thalidomide was in a cost-effective range below a threshold of €20,000 per LY. Economic evaluations of bortezomib compared to lenalidomide in the treatment of relapsed/refractory disease, however, have demonstrated bortezomib as either dominant (i.e., improved health at reduced costs) in the Swedish situation [46] or less costly and less effective in Norway [47]. These findings suggest that bortezomib as relapsed/refractory therapy would be favored over lenalidomide from the Swedish perspective as it offers a 'win-win' situation. In Norway, however, lenalidomide would be considered more cost-effective assuming a threshold of approximately equal to or slightly below €30,000 per LY or QALY. Both studies did find that bortezomib was less costly compared to RD and this should be consistent in other countries since reference prices for reimbursement are typically based on that in the US, which is also known to be higher for lenalidomide [8]. However, the ICER for bortezomib compared to lenalidomide is at this point still uncertain on account of the conflicting results for incremental effectiveness.

The conflicting effectiveness results from these two studies are surprising given that both studies modeled the same comparisons using data from the same trials. The explanation for this is likely to be the difference in modeling methods and assumptions. Hornberger et al. (2010) performed a partitioned survival analysis using the published estimates for the trials as reported in the literature [46]. In contrast, Moller et al. (2011) performed a DES which was based on patient-level data for lenalidomide and published literature estimates for the bortezomib strategy [47]. Both studies over-estimated the median OS in terms of LYs compared to the results of the trials that reported a median OS of 2.5 LYs for bortezomib and 3.2 LYs for lenalidomide. The degree of over-estimation is less severe for the results reported by Moller et al. (2011) (V: 3.11 LYs; RD: 4.06) [47], while the Hornberger et al. (2010) analysis greatly overestimated the effectiveness (V: 4.78 LYs; RD: 4.51) [46]. Heterogeneity in outcomes in multiple myeloma patients has been demonstrated[53]. The use of partitioned survival analysis by Hornberger et al. (2010), however, may have limited the ability to include such heterogeneity as this method does not allow examination of the effect of covariates on the survival curves [54]. Studies assessing the precision of DES models have demonstrated that DES models more accurately predict survival on account of improved precision. Greater precision is gained because more variables influencing outcomes can be added to a DES model [55], which is especially relevant when modeling outcomes for diseases such as multiple myeloma with a great deal of heterogeneity and limited follow-up data [56].

Not only do differences in the modeling methods contribute to the conflicting results, but also differences in the assumptions for post-progression survival (PPS). The DES study assumed that PPS was a function of the best response achieved, which was specific to

response type but irrespective of initial treatment for relapsed/refractory disease, based on a fitted model using the patient-level data from the MM-009/-010 trial. In contrast, Hornberger et al. (2010) used the hazard ratios reported in the literature to estimate cycle-specific probabilities for TTP and OS which did not incorporate differences in survival associated with best response. As best response rates were more favorable in the RD arm of the MM-009/-010 study compared to the bortezomib arm of the APEX trial (i.e., ORR: 60% versus 43%, respectively), the PPS calculated in the DES model incorporated greater gains in survival for the RD strategy compared to bortezomib. The assumption of gains in survival with achievement of better response was partially valid in the model by Moller et al. (2011), as the prognostic value of response to treatment with novel agents in relapsed/refractory therapy has been demonstrated [57, 58]. However, the assumption that the survival gains by response type does not differ by therapy was invalid as the degree of benefit has been shown to differ by therapy [59]. Nonetheless, both studies should be questioned namely on account of the fact that both studies performed a cost-utility analysis by means of an indirect comparison, which is susceptible to various differences in the trials under comparison that could render them incomparable (i.e., patient characteristics and setting) and instead may explain the conflicting results [60, 61].



Recent evidence on the efficacy of novel agents requires mentioning as it could impact their cost-effectiveness. Existing economic evaluations were based on data from trials administering a higher dose of dexamethasone and a more intensive treatment scheme for bortezomib. However, a low-dose approach for dexamethasone is now the standard of care [62] as it has been found to improve outcomes by lowering the risk for toxicity-related dose modifications or treatment interruptions [63]. Bortezomib has also demonstrated a better toxicity profile and no difference in efficacy when administered once weekly instead of twice weekly [22, 23]. Economic evaluations incorporating more recent evidence for low-dose dexamethasone and once weekly bortezomib could improve the cost-effectiveness ratios for novel agent regimens. Furthermore, a recent review of Phase III trials of treatments for multiple myeloma showed an excess of second primary cancers among patients with multiple myeloma receiving lenalidomide maintenance therapy [64]. This finding could have an unfavorable impact on its cost-effectiveness. However, the risk for second malignancies after multiple myeloma was established decades ago prior to the availability of lenalidomide, and it is generally regarded that the risk of dying from multiple myeloma is considerably higher than the risk of the development of a second cancer [65]. The risk-benefit ratio for lenalidomide is therefore in favor of treatment, hence having no impact on its cost-effectiveness ratio.

Another important issue to consider when assessing the cost-effectiveness of novel agents in multiple myeloma is whether or not the focus should be on estimating the mean cost-effectiveness results from the population level. The majority of analyses comparing novel agent regimens to one another tend to be sensitive to the hazard ratio or survival difference between the novel agent strategies and the utility of progression. This suggests that the extent to which one agent improves health outcomes compared to the other at the population level is small. More meaningful improvements in cost-effectiveness could be found for particular patient groups accounting for their heterogeneity in prognosis, treatment response, and sensitivity to side effects or the presence of molecular biomarkers.

Additional targeted agents are expected to be available in the near future depending on the results of ongoing clinical trials [66]. Two promising agents demonstrating favorable efficacy results in Phase I and II studies include the third-generation IMiD, pomalidomide [67], and the second-generation proteasome inhibitor, carfilzomib [68]. Pomalidomide has shown high response rates in patients refractory to lenalidomide or bortezomib [69, 70], and like its predecessor it was well tolerated with manageable adverse events and no grade 3 neuropathy and no thromboembolic events. Pomalidomide was granted orphan status by the EU in 2009 while in the US its approval is pending Phase III trial results expected to become available September 2012 [71]. Carfilzomib has demonstrated significant activity in heavily pre-treated patients relapsing or refractory to previous therapy with novel agents [61, 72], also with a manageable toxicity profile, consisting mainly of myelosuppression and very little neuropathy. Its manufacturer recently filed for accelerated FDA approval based on early findings although Phase III trial results are not expected to become available until December 2013 [73]. However, the FDA eventually rejected accelerated approval status, instead granting a standard review with a set target date of July 2012 [74]. The reasons cited included a preference for Phase III data and the fact that the trial was a single arm study. The same requirements for data are likely to be expected for future agents seeking approval for multiple myeloma.

## CONCLUSION

The arrival of second- and third-generation novel agents will require evidence demonstrating improvements in health that will be compared to the costs and health gains of standard therapy, which is currently dominated by thalidomide, bortezomib and lenalidomide. The gaps in knowledge regarding health-related quality of life improvements should be addressed for current and future novel agents as it may be the case that newer agents improve quality of life rather than extend life and this may also be

associated with best response to therapy. Further unresolved questions remain regarding the comparative effectiveness as well as the cost-effectiveness of currently available novel agents in terms of appropriate sequencing tailored to the patient and when used in combination with one another.

## EXPERT COMMENTARY

The cost-effectiveness of all three therapies when used alone or in combination with conventional agents is generally considered to be within acceptable limits. However, the incremental cost-effectiveness of one compared to the other is less known as the evidence on their comparative effectiveness is lacking. Existing modeling studies have not been able to overcome the data limitations as we see conflicting results in studies based on efficacy from similar trials. The assumptions for health benefits with novel agents incorporated in existing modeling studies are influential and often overestimate the lifetime health gains. Modeling studies in this indication should therefore be carefully criticized by experts on the validity of their assumptions so as to avoid informing decision-making based on unrealistic estimates of health benefits.

## FIVE-YEAR VIEW

The treatment paradigm for multiple myeloma is quickly evolving in ways that will likely increase the total costs of treatment while also extending life expectancy, although the costs could decrease with the finding of a cure that replaces chronic treatment. First, the availability of treatment options for multiple myeloma will continue to expand with additional targeted agents expected to be available in the near future. In addition to pomalidomide and carfilzomib, FDA approval is expected in the coming years for other investigational agents demonstrating promising efficacy results in early trials: vorinostat [75], and perifosine [76]. Phase III data for vorinostat, which is an inhibitor of the histone deacetylase, were recently presented and shown to be favorable [77]. Results of the Phase III study for perifosine, which is an inhibitor of heat-shock protein 90, are expected as early as September 2012 [78, 107]. Two other drugs also on the horizon include an inhibitor of histone deacetylase, panobinostat [79, 80], and a monoclonal antibody, erlotuzumab [81], for which Phase III trial results are expected as early as September 2013 [82] and March 2014 [83], respectively. Payers will assess whether these drugs are demonstrated to be cost-effective in the near future. Given the trend toward increasing research and development costs of novel agents, unit prices for these agents are likely to be similar to that of bortezomib and lenalidomide but could potentially be even

higher. Second, there is also a move toward combination therapy of three or more drugs including at least one novel agent. With demonstration of improved outcomes based on drug synergy [84], patients will be eligible for a combination of novel agents. The additional costs of combination therapy with more than one expensive novel agent could further increase the lifetime costs of treatment. Third, the indication for novel agents may expand further to asymptomatic smoldering multiple myeloma (SMM) patients [85], which is an asymptomatic precursor disease with a high propensity to progress to symptomatic multiple myeloma. In particular, high-risk SMM patients are an identifiable group that may benefit from early treatment. Recent evidence suggests possible benefits for these patients with treatment of novel agents, such as immunomodulatory drugs, on account of their toxicity and efficacy profiles[86, 87], which will lead to a further increase in the overall economic burden of multiple myeloma treatment. Finally, treatment stratification based on both clinical and molecular stratification is currently being explored by means of gene-expression profiling [88]. Through demonstration of the prognostic impact of molecular markers, biomarkers and diagnostic approaches for identification of risk groups will facilitate a more personalized treatment approach for future patients. Amid these changes in the treatment paradigm, future economic evaluations of treatment for multiple myeloma will focus on assessing the health gains and costs of treatment with novel agents for more narrowly defined patient groups.

## REFERENCES

1. McCabe C, Bergmann L, Bosanquet N, et al. Market and patient access to new oncology products in Europe: a current, multidisciplinary perspective. Ann Oncol 2009; 20(3): 403-12.
2. Garber AM, McClellan MB. Satisfaction guaranteed--"payment by results" for biologic agents. N Engl J Med 2007; 357(16): 1575-7.
3. Anderson KC. Lenalidomide and thalidomide: mechanisms of action--similarities and differences. Semin Hematol 2005; 42(4 Suppl 4): S3-8.
4. Kane RC, Bross PF, Farrell AT, Pazdur R. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. Oncologist 2003; 8(6): 508-13.
5. Fuchs SY. The role of ubiquitin-proteasome pathway in oncogenic signaling. Cancer Biol Ther 2002; 1(4): 337-41.
6. Richardson PG, Hideshima T, Mitsiades C, Anderson KC. The emerging role of novel therapies for the treatment of relapsed myeloma. J Natl Compr Canc Netw 2007; 5(2): 149-62.
7. 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. J Clin Pharm Ther 2011; 36(1): 19-26.
8. Bonkowski J, Vermeulen LC, Kolesar JM. The clinical utility of lenalidomide in multiple myeloma and myelodysplastic syndromes. J Oncol Pharm Pract 2009; 16(4): 223-32.
9. Knox C, Law V, Jewison T, et al. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res 2011; 39(Database issue): D1035-41.

10. Mitsiades CS, Hideshima T, Chauhan D, et al. Emerging treatments for multiple myeloma: beyond immunomodulatory drugs and bortezomib. Semin Hematol 2009; 46(2): 166-75.
11. Drummond MF, Schulpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. Third ed. Oxford: Oxford University Press; 2005.
12. European Central Bank. Statistical Data Warehouse. [cited 2011 October 27]; Available from: http://sdw.ecb.europa.eu/
13. Palumbo A, Attal M, Roussel M. Shifts in the Therapeutic Paradigm for Patients Newly Diagnosed with Multiple Myeloma: Maintenance Therapy and Overall Survival. Clinical Cancer Research 2011; 17(6): 1253-63.
14. Hulin C, Facon T, Rodon P, et al. Efficacy of melphalan and prednisone plus thalidomide in patients older than 75 years with newly diagnosed multiple myeloma: IFM 01/01 trial. J Clin Oncol 2009; 27(22): 3664-70.
15. Ludwig H, Hajek R, Tothova E, et al. Thalidomide-dexamethasone compared with melphalan-prednisolone in elderly patients with multiple myeloma. Blood 2009; 113(15): 3435-42.
16. Palumbo A, Bringhen S, Caravita T, et al. Oral melphalan and prednisone chemotherapy plus thalidomide compared with melphalan and prednisone alone in elderly patients with multiple myeloma: Randomised controlled trial. Lancet 2006; 367(9513): 825-31.
17. Palumbo A, Bringhen S, Liberati AM, et al. Oral melphalan, prednisone, and thalidomide in elderly patients with multiple myeloma: Updated results of a randomized controlled trial. Blood 2008; 112(8): 3107-14.
18. Waage A, Gimsing P, Fayers P, et al. Melphalan and prednisone plus thalidomide or placebo in elderly patients with multiple myeloma. Blood 2010; 116(9): 1405-12.
19. Wijermans P, Schaafsma M, Termorshuizen F, et al. Phase III study of the value of thalidomide added to melphalan plus prednisone in elderly patients with newly diagnosed multiple myeloma: the HOVON 49 Study. J Clin Oncol 2010; 28(19): 3160-6.
20. Mateos MV, Richardson PG, Schlag R, et al. Bortezomib plus melphalan and prednisone compared with melphalan and prednisone in previously untreated multiple myeloma: updated follow-up and impact of subsequent therapy in the phase III VISTA trial. J Clin Oncol 2010; 28(13): 2259-66.
21. San Miguel JF, Schlag R, Khuageva NK, et al. Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. N Engl J Med 2008; 359(9): 906-17.
22. Mateos MV, Oriol A, Martinez-Lopez J, et al. Bortezomib, melphalan, and prednisone versus bortezomib, thalidomide, and prednisone as induction therapy followed by maintenance treatment with bortezomib and thalidomide versus bortezomib and prednisone in elderly patients with untreated multiple myeloma: a randomised trial. Lancet Oncol 2010; 11(10): 934-41.
23. Palumbo A, Bringhen S, Rossi D. Bortezomib-melphalan-prednisone-thalidomide followed by maintenance with bortezomib-tahlidomide versus bortezomib-melphalan-prednisone for initial treatment of multiple myeloma: a randomised controlled trial. . J Clin Oncol 2010.
24. Zonder JA, Crowley J, Hussein MA, et al. Lenalidomide and high-dose dexamethasone compared with dexamethasone as initial therapy for multiple myeloma: A randomized Southwest Oncology Group trial (S0232). Blood 2010; 116(26): 5838-41.
25. Rajkumar SV, Blood E, Vesole D, et al. Phase III clinical trial of thalidomide plus dexamethasone compared with dexamethasone alone in newly diagnosed multiple myeloma: a clinical trial coordinated by the Eastern Cooperative Oncology Group. J Clin Oncol 2006; 24(3): 431-6.
26. Rajkumar SV, Rosinol L, Hussein M, et al. Multicenter, randomized, double-blind, placebo-controlled study of thalidomide plus dexamethasone compared with dexamethasone as initial therapy for newly diagnosed multiple myeloma. J Clin Oncol 2008; 26(13): 2171-7.

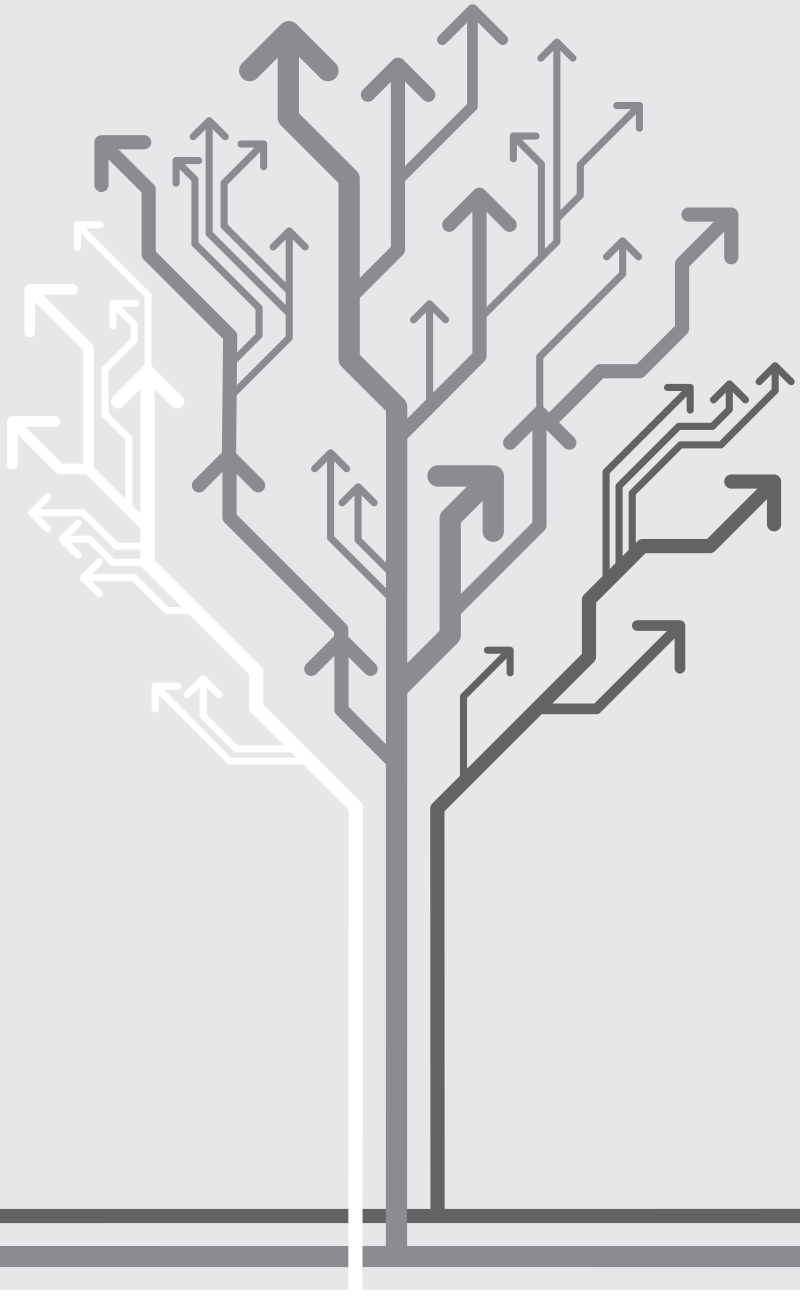
27. Facon T, Mary JY, Hulin C, et al. Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): a randomised trial. Lancet 2007; 370(9594): 1209-18.
28. Barlogie B, Pineda-Roman M, van Rhee F, et al. Thalidomide arm of Total Therapy 2 improves complete remission duration and survival in myeloma patients with metaphase cytogenetic abnormalities. Blood 2008; 112(8): 3115-21.
29. Barlogie B, Tricot G, Anaissie E, et al. Thalidomide and hematopoietic-cell transplantation for multiple myeloma. N Engl J Med 2006; 354(10): 1021-30.
30. Lokhorst HM, van der Holt B, Zweegman S, et al. A randomized phase 3 study on the effect of thalidomide combined with adriamycin, dexamethasone, and high-dose melphalan, followed by thalidomide maintenance in patients with multiple myeloma. Blood 2010; 115(6): 1113-20.
31. Spencer A, Prince HM, Roberts AW, et al. Consolidation therapy with low-dose thalidomide and prednisolone prolongs the survival of multiple myeloma patients undergoing a single autologous stem-cell transplantation procedure. J Clin Oncol 2009; 27(11): 1788-93.
32. Harousseau JL, Attal M, Avet-Loiseau H, et al. Bortezomib plus dexamethasone is superior to vincristine plus doxorubicin plus dexamethasone as induction treatment prior to autologous stem-cell transplantation in newly diagnosed multiple myeloma: results of the IFM 2005-01 phase III trial. J Clin Oncol 2010; 28(30): 4621-9.
33. Cavo M, Tacchetti P, Patriarca F, et al. Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: A randomised phase 3 study. The Lancet 2010; 376(9758): 2075-85.
34. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med 2005; 352(24): 2487-98.
35. 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.
36. Orlowski RZ, Nagler A, Sonneveld P, et al. Randomized phase III study of pegylated liposomal doxorubicin plus bortezomib compared with bortezomib alone in relapsed or refractory multiple myeloma: combination therapy improves time to progression. J Clin Oncol 2007; 25(25): 3892-901.
37. Mikhael JR, Belch AR, Prince HM, et al. High response rate to bortezomib with or without dexamethasone in patients with relapsed or refractory multiple myeloma: results of a global phase 3b expanded access program. Br J Haematol 2008; 144(2): 169-75.
38. Weber DM, Chen C, Niesvizky R, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. N Engl J Med 2007; 357(21): 2133-42.
39. Dimopoulos M, Spencer A, Attal M, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. N Engl J Med 2007; 357(21): 2123-32.
40. Dimopoulos MA, Chen C, Spencer A, et al. Long-term follow-up on overall survival from the MM-009 and MM-010 phase III trials of lenalidomide plus dexamethasone in patients with relapsed or refractory multiple myeloma. Leukemia 2009; 23(11): 2147-52.
41. Castaneda CP, Brandenburg NA, Bwire R, et al. Erythema multiforme/Stevens-Johnson syndrome/toxic epidermal necrolysis in lenalidomide-treated patients. J Clin Oncol 2009; 27(1): 156-7.
42. Verelst SGR, Termorshuizen F, Uyl-de Groot CA, et al. Effect of thalidomide with melphalan and prednisone on health-related quality of life (HRQoL) in elderly patients with newly diagnosed multiple myeloma: a prospective analysis in a randomized trial. Annals of Hematology 2011: 1-13.
43. Dubois D, Dhawan R, van de Velde H, et al. Descriptive and prognostic value of patient-reported

outcomes: the bortezomib experience in relapsed and refractory multiple myeloma. J Clin Oncol 2006; 24(6): 976-82.
44. Lee SJ, Richardson PG, Sonneveld P, et al. Bortezomib is associated with better health-related quality of life than high-dose dexamethasone in patients with relapsed multiple myeloma: Results from the APEX study. British Journal of Haematology 2008; 143(4): 511-9.
45. 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.
46. Hornberger J, Rickert J, Dhawan R, et al. The cost-effectiveness of bortezomib in relapsed/refractory multiple myeloma: Swedish perspective. Eur J Haematol 2010; 85(6): 484-91.
47. Moller J, Nicklasson L, Murthy A. Cost-effectiveness of novel relapsed-refractory multiple myeloma therapies in Norway: lenalidomide plus dexamethasone vs bortezomib. J Med Econ. 2011;14(6):690-7.
48. Bernard P. Zeigler, Herbert Praehofer, and Tag Gon Kim. Theory of modeling and simulation: Integrating discrete event and continuous complex dynamic systems Second ed. Sand Diego, California: Academic Press; 2000.
49. Van Agthoven M, Segeren CM, Buijt I, et al. A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma; a prospective randomised phase III study. Eur J Cancer 2004; 40(8): 1159-69.
50. Evidence Review Group Report for NICE's Single Technology Appraisal process. Bortezomib for the treatment of multiple myeloma patients. July 2006 [cited 2011 October 27]; Available from: http://www.hta.ac.uk/erg/reports/1637.pdf
51. Evidence Review Group Report for NICE's Single Technology Appraisal process. The Clinical and Cost-Effectiveness of Lenalidomide for multiple myeloma in people who have received at least one prior therapy: An evidence review of the submission from Celgene. October 2008 [cited 2011 October 27]; Available from: http://www.hta.ac.uk/erg/reports/1746.pdf
52. Rowen D, Brazier J, Young T, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health 14(5): 721-31.
53. Mitsiades CS, Davies FE, Laubach JP, et al. Future directions of next-generation novel therapies, combination approaches, and the development of personalized medicine in myeloma. J Clin Oncol 2011; 29(14): 1916-23.
54. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. Stat Med 1990; 9(11): 1259-76.
55. Simpson KN, Strassburger A, Jones WJ, et al. Comparison of Markov model and discrete-event simulation techniques for HIV. Pharmacoeconomics 2009; 27(2): 159-65.
56. Heeg BM, Damen J, Buskens E, et al. Modelling approaches: the case of schizophrenia. Pharmacoeconomics 2008; 26(8): 633-48.
57. Moreau P, Attal M, Pegourie B, et al. Achievement of VGPR to induction therapy is an important prognostic factor for longer PFS in the IFM 2005-01 trial. Blood 2010; 117(11): 3041-4.
58. Niesvizky R, Richardson PG, Rajkumar SV, et al. The relationship between quality of response and clinical benefit for patients treated on the bortezomib arm of the international, randomized, phase 3 APEX trial in relapsed multiple myeloma. Br J Haematol 2008; 143(1): 46-53.
59. Harousseau JL, Attal M, Avet-Loiseau H. The role of complete response in multiple myeloma. Blood 2009; 114(15): 3139-46.
60. Gartlehner G, Moore CG. Direct versus indirect comparisons: a summary of the evidence. Int J Technol Assess Health Care 2008; 24(2): 170-7.
61. Siegel DSD, Martin T, Wang M, et al. PX-171-003-A1, an open-label, single-arm, phase (Ph) II study

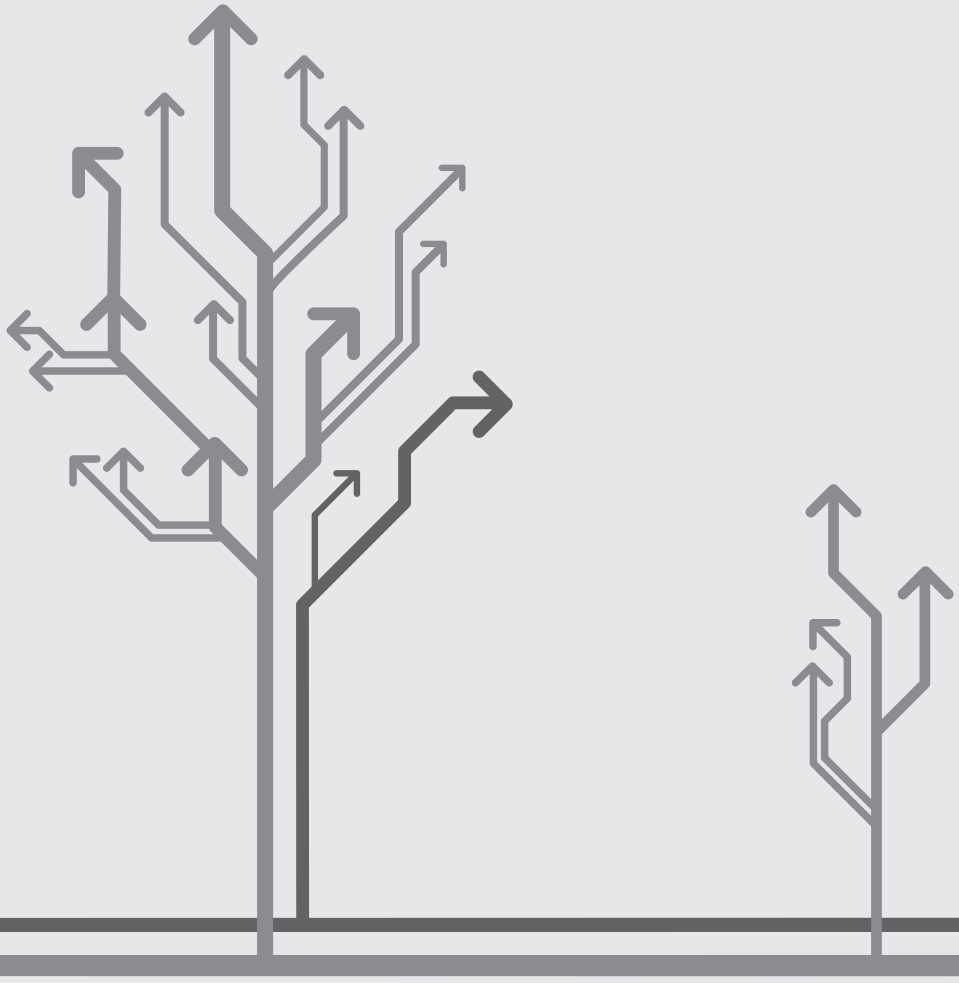
of carfilzomib (CFZ) in patients (pts) with relapsed and refractory multiple myeloma (R/R MM): Long-term follow-up and subgroup analysis. American Society of Clinical Oncology Annual Meeting; 2011; Chicago; 2011.

62. Rajkumar SV. Multiple myeloma: 2012 update on diagnosis, risk-stratification, and management. Am J Hematol 2012; 87(1): 78-88.
63. Rajkumar SV, Jacobus S, Callander NS, et al. Lenalidomide plus high-dose dexamethasone versus lenalidomide plus low-dose dexamethasone as initial therapy for newly diagnosed multiple myeloma: an open-label randomised controlled trial. Lancet Oncol 2010; 11(1): 29-37.
64. Dimopoulos MA, Richardson PG, Brandenburg N, et al. A review of second primary malignancy in patients with relapsed or refractory multiple myeloma treated with lenalidomide. Blood 2012.
65. Landgren O, Thomas A, Mailankody S. Myeloma and second primary cancers. N Engl J Med 2011; 365(23): 2241-2.
66. Mahindra A, Cirstea D, Raje N. Novel therapeutic targets for multiple myeloma. Future Oncol 2010; 6(3): 407-18.
67. Schey S, Ramasamy K. Pomalidomide therapy for myeloma. Expert Opin Investig Drugs 2011; 20(5): 691-700.
68. Khan ML, Stewart AK. Carfilzomib: a novel second-generation proteasome inhibitor. Future Oncol 2011; 7(5): 607-12.
69. Lacy MQ, Hayman SR, Gertz MA, et al. Pomalidomide (CC4047) plus low-dose dexamethasone as therapy for relapsed multiple myeloma. J Clin Oncol 2009; 27(30): 5008-14.
70. Mikhael JR, Rajkumar V, Roy V, et al. Efficacy of pomalidomide plus low-dose dexamethasone in multiple myeloma patients despite previous use of lenalidomide. American Society of Clincal Oncology Annual Meeting; 2011; Chicago; 2011.
71. Celgene Corporation. A Phase III, Multicenter, Randomized, Open-Label Study to Compare the Efficacy and Safety of Pomalidomide in Combination With Low-Dose Dexamethasone Versus High-Dose Dexamethasone in Subjects With Refractory Multiple Myeloma or Relapsed and Refractory Multiple Myeloma. Bethesda (MD): National Library of Medicine (US), 2012 [cited 2012 March 12]; Available from: http://clinicaltrials.gov/show/NCT01311687
72. Stewart AK, Kaufman JL, Jakubowiak AJ, et al. The effect of carfilzomib (CFZ) in patients (Pts) with bortezomib (BTZ)-naive relapsed or refractory multiple myeloma (MM): Updated results from the PX-171-004 study. American Society of Clinical Oncology Annual Meeting; 2011; Chicago; 2011.
73. Onyx Therapeutics Inc. Phase III Study Comparing Carfilzomib, Lenalidomide, and Dexamethasone (CRd) Versus Lenalidomide and Dexamethasone (Rd) in Subjects With Relapsed Multiple Myeloma. Bethesda (MD), 2012 [cited 2012 March 12]; Available from: http://clinicaltrials.gov/show/NCT01080391
74. Onyx Pharmaceuticals Inc. Onyx Pharmaceuticals Announces FDA Standard Review Designation of Carfilzomib New Drug Application for the Treatment of Relapsed and Refractory Multiple Myeloma. December 2011 [cited 2012 March 2012]; Available from: http://www.onyx.com/view.cfm/551/onyx-pharmaceuticals-announces-fda-standard-review-designation-of-carfilzomib-new-drug-application-for-the-treatment-of-relapsed-and-refractory-multiple-myeloma
75. Dimopoulos MA, Jagannath S, Yoon S-S, et al. Vantage 088: Vorinostat in Combination with Bortezomib in Patients with Relapsed/Refractory Multiple Myeloma: Results of a Global, Randomized Phase 3 Trial. ASH Annual Meeting Abstracts 2011; 118(21): 811-.
76. Richardson PG, Wolf J, Jakubowiak A, et al. Perifosine plus bortezomib and dexamethasone in patients with relapsed/refractory multiple myeloma previously treated with bortezomib: results of a multicenter phase I/II trial. J Clin Oncol 2011; 29(32): 4243-9.

77. Wolf JL, Jakubowiak AJ, Zonder JA, et al. Perifosine Plus Bortezomib and Dexamethasone in Relapsed / Refractory Multiple Myeloma Patients Previously Treated with Bortezomib: Final Results of a Phase I/II Trial. American Society of Hematology Annual Meeting; 2011; San Diego; 2011.
78. Keryx/AOI Pharmaceuticals Inc. A Phase III Randomized Study to Assess the Efficacy and Safety of Perifosine Added to the Combination of Bortezomib and Dexamethasone in Multiple Myeloma Patients. Bethesda (MD): National Library of Medicine (US), 2012 [cited 2012 March 12]; Available from: http://clinicaltrials.gov/show/NCT01002248
79. Wolf JL, Siegel D, Goldschmidt H, et al. Phase II trial of the pan-deacetylase inhibitor panobinostat as a single agent in advanced relapsed/refractory multiple myeloma. Leuk Lymphoma 2012.
80. Offidani M, Polloni C, Cavallo F, et al. Phase II study of Melphalan, Thalidomide and Prednisone (MPT) combined with oral Panobinostat in patients with relapsed/refractory MM. Leuk Lymphoma 2012.
81. Jakubowiak AJ, Benson DM, Bensinger W, et al. Phase I Trial of Anti-CS1 Monoclonal Antibody Elotuzumab in Combination With Bortezomib in the Treatment of Relapsed/Refractory Multiple Myeloma. J Clin Oncol 2012.
82. Novartis Pharmaceuticals. A Multicenter, Randomized, Double Blind, Placebo Controlled Phase III Study of Panobinostat in Combination With Bortezomib and Dexamethasone in Patients With Relapsed Multiple Myeloma. Bethesda (MD): National Library of Medicine (US), 2012 [cited 2012 March 12]; Available from: http://clinicaltrials.gov/show/NCT01023308
83. Bristol-Myers Squibb. Phase III, Randomized, Open Label Trial of Lenalidomide/Dexamethasone With or Without Elotuzumab in Relapsed or Refractory Multiple Myeloma. Bethesda (MD): National Library of Medicine (US), 2012 [cited 2012 March 2012]; Available from: http://clinicaltrials.gov/show/NCT01239797
84. Jagannath S, Dimopoulos MA, Lonial S. Combined proteasome and histone deacetylase inhibition: A promising synergy for patients with relapsed/refractory multiple myeloma. Leuk Res 2010; 34(9): 1111-8.
85. Korde N, Kristinsson SY, Landgren O. Monoclonal gammopathy of undetermined significance (MGUS) and smoldering multiple myeloma (SMM): novel biological insights and development of early treatment strategies. Blood 2011; 117(21): 5573-81.
86. Barlogie B, van Rhee F, Shaughnessy JD, Jr., et al. Seven-year median time to progression with thalidomide for smoldering myeloma: partial response identifies subset requiring earlier salvage therapy for symptomatic disease. Blood 2008; 112(8): 3122-5.
87. Mateos M-V, Lopez-Corral L, Hernandez MT, et al. Multicenter, Randomized, Open-Label, Phase III Trial of Lenalidomide-Dexamethasone (Len/dex) Vs Therapeutic Abstention in Smoldering Multiple Myeloma at High Risk of Progression to Symptomatic MM: Results of the First Interim Analysis. ASH Annual Meeting Abstracts 2009; 114(22): 614-.
88. Johnson SK, Heuck CJ, Albino AP, et al. The use of molecular-based risk stratification and pharmacogenomics for outcome prediction and personalized therapeutic management of multiple myeloma. Int J Hematol 2011; 94(4): 321-33.



# Part 2 – Outcomes research of bortezomib in advanced multiple myeloma

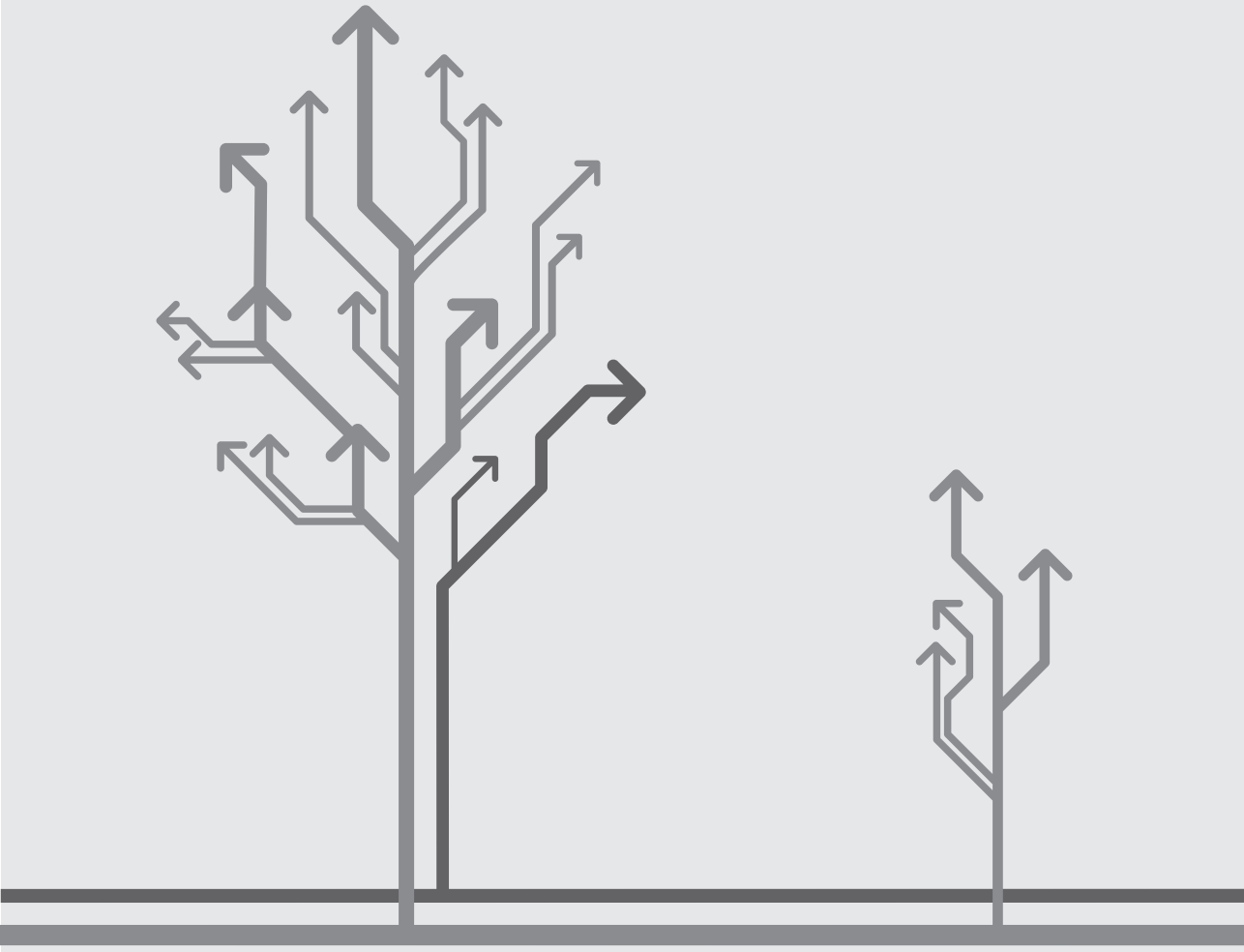


# Chapter 4.

## Practical implications of outcomes research into the appropriate use and effectiveness of novel cancer therapies: Experience with bortezomib in advanced multiple myeloma

Jennifer G. Gaultney, Margreet G. Franken, Carin A. Uyl-de Groot, William K. Redekop, Peter C. Huijgens, Bronno van der Holt, Henk M. Lokhorst, Pieter Sonneveld

*Submitted.*



## ABSTRACT

Decision makers more often request outcomes research of novel cancer drugs to help resolve uncertainty of their health benefits and budget impact during the initial reimbursement decision phase. Given the known limitations to observational data in demonstrating treatment effects, we assessed its usefulness in evaluating the appropriate use and effectiveness of bortezomib in advanced multiple myeloma patients treated in Dutch daily practice. Data were retrospectively collected from the pivotal trial patients (APEX; n=333) and two groups of daily practice patients treated with bortezomib following progression from HOVON-50 (n=201): patients treated as of May 2009 (HO50-1; n=72) and June 2012 (HO50-2; n=129). Prognosis, treatment, and effectiveness were compared. Prognosis was less favorable for HO50-1 compared to APEX and HO50-2. Lower planned dosages (20%), frequent dose modification (57%), premature discontinuation (65%) and treatment variation were observed in HO50-1 versus APEX. Data on serum β2-microglobulin (65%), c-reactive protein (72%), response rates (15%) and progression dates (14%) were missing in patient charts. Estimating progression-free survival was not feasible. Observational data was useful for assessing drug use and overall survival in initially eligible daily practice patients. However, the drug's real-world effectiveness was not representative of the target patient population but was instead limited to a sample of patients with unfavorable prognosis. Decision-making incorporating inferences about effectiveness of novel drugs in multiple myeloma requires careful consideration of the possibilities as well as the limitations of outcomes research during the drug's uptake period.

## INTRODUCTION

Multiple myeloma (MM) is a malignant plasma cell disorder which accounts for 1% of all cancer diagnoses worldwide and 13% of all hematologic malignancies[1]. The survival of MM patients has improved substantially in the past decade partly due to the introduction of novel agents [2],[3], among which includes bortezomib (Velcade®).

The pivotal phase III APEX (Assessment of Proteasome Inhibition for Extending Remissions) trial demonstrated superior efficacy of bortezomib compared to dexamethasone in advanced MM [4, 5], leading to its approval in 2003 and 2004 by the FDA [6] and European Medicines Agency (EMA) [7], respectively. Decision makers in the Netherlands and the UK were confronted with uncertainty regarding its effectiveness in daily practice and overall budget impact and hence access to care. Consequently, a conditional reimbursement was instituted by means of a performance-based scheme in the UK [8] and a coverage with evidence development scheme in the Netherlands [9, 10]. Both policies allowed access to the drug while requiring outcomes research to be performed to assess appropriate drug use (e.g., patient characteristics, types of treatments, dosages, and dose modifications), effectiveness and costs in daily practice.

Outcomes research is useful for facilitating a feedback loop to clinicians about shortfalls in daily practice so that policies can be implemented to improve quality and access to care [11, 12]. It does however have its limitations as it is based on observational data often collected during the diffusion phase of a new technology, lending it susceptible to bias [13]. Decision making that incorporates biased effectiveness data from daily practice could lead to the wrong decision and ultimately hinder rather than improve society's access to innovations in healthcare. Therefore, it is important to inform decision makers involved in reimbursement decisions of the feasibility and usefulness of observational data to assess the therapeutic value of a new drug in daily practice.

To demonstrate the value of outcomes research in evaluating a drug's appropriate use and real-world effectiveness, we conducted an outcomes research study of bortezomib in advanced MM in the Netherlands following its EMA approval, as outlined in the Dutch policy rule. We report here the evidence generated in daily practice compared to the trial and draw conclusions regarding the usefulness of observational data to assess the drug's therapeutic value in Dutch patients.

## METHODS

Data on two patient cohorts receiving bortezomib for treatment of advanced MM were collected: one from daily practice patients treated in the Netherlands following progression/relapse or refractory disease from an RCT for upfront therapy (n=201) [14], and one from the pivotal phase III APEX trial (n=333) [4, 5].

### Patient groups

Information about the inclusion/exclusion criteria for each patient group is available in Table 1.

**Table 1.** Description of patient groups and study methods.

| | Trial setting (n=333) | Daily practice (n=201) | |
|---|---|---|---|
| | **APEX (n=333)[4,5]** | **HO50-1 (n=72)** | **HO50-2 (n=129)** |
| Eligibility criteria<br>For inclusion: | Measurable progressive disease after one to three previous treatments; Karnofsky ≥ 60; Platelet count ≥50.000 per cubic millimeter; Hemoglobin ≥ 7.5 g/dl; Neutrophil count ≥ 750 per cubic millimeter; Creatinine clearance ≥ 20ml/min | Relapsed or refractory to HOVON 50 treatment protocol for first line; Measurable progressive disease after one or more previous treatments; Received bortezomib for relapsed/refractory disease in daily practice | Relapsed or refractory to HOVON 50 treatment protocol for first line; Measurable progressive disease after one or more previous treatments; Received bortezomib for relapsed/refractory disease in daily practice |
| For exclusion: | Previous bortezomib or refractory to high dose dexamethasone; Grade 2 peripheral neuropathy; Any clinically significant coexisting illness unrelated to myeloma | Secondary malignancy (excluding basal cell carcinoma); Received bortezomib for relapsed/ refractory disease under a controlled trial setting | None |
| Study design | Randomized controlled trial | Retrospective observational design | Retrospective observational design |
| Data collection | Prospective | Detailed case reports using medical charts review | Data extracted from trial follow-up data collected by the HOVON Data Centre |
| Date of follow-up | June 2002 to March 2006 | January 2001 and May 2009 with date of last contacted updated as of June 2012 | January 2001 and June 2012 |
| Data points available for comparison | Baseline prognostic factors, treatment, adverse events, efficacy in terms of response rates, TTP and OS. | Baseline prognostic factors, treatment, adverse events, effectiveness in terms of response rates, TTP and OS. | Baseline prognostic factors and effectiveness in terms of response rates, TTP and OS. |

#### *Daily practice patients*

Daily practice patients were selected from 556 patients enrolled from November 2001 to June 2005 in the phase III Dutch-Belgian Cooperative Trial Group for Hematology

Oncology (HOVON)-50 trial [14] which investigated the treatment effect of thalidomide, adriamycin and dexamethasone (TAD) versus vincristine, adriamycin and dexamethasone (VAD) in newly diagnosed Durie-Salmon stage II/III MM patients aged 18 – 65 years. In this trial, TAD improved event-free survival and PFS compared to VAD while patients randomized to TAD had strongly reduced survival after progression/relapse. Patients relapsing from the HOVON-50 regime were then treated with treatment options available for advanced disease at the discretion of the treating physician, e.g. with bortezomib, in everyday clinical practice.

To demonstrate the pitfalls of assessing effectiveness during the initial years of the drugs availability in daily practice, patients in this cohort consisted of two groups: daily practice sample 1 (HO50-1) (n=72) and daily practice sample 2 (HO50-2) (n=129). HO50-1 included the original group of patients for which the outcomes research study was performed in the years 2006-2009. For this group, detailed case reports using medical charts were retrospectively collected. Data was collected from start of HOVON-50 therapy until last known follow-up between January 2001 and May 2009.



All HOVON-50 patients who had received bortezomib for relapsed/refractory disease before June 2012 were extracted from trial follow-up data collected by the HOVON Data Centre (HDC). A total of 193 patients were identified which also included 64 of the HO50-1 patients, for which the date of last contact and follow-up status were updated. The remaining 129 patients were included in the HO50-2 group. Data for these patients were taken from the follow-up data which is periodically collected by the HDC. Since the primary purpose of collecting data after progression/relapse by the HDC was for tracking overall survival for all patients treated in the trial, it did not provide detail regarding drug use and safety.

***APEX trial patients***

The trial patient population was represented by the bortezomib arm of the APEX trial, which enrolled advanced patients from June 2002 to October 2003. To allow for statistical comparisons, the manufacturer of bortezomib (Janssen Pharmaceuticals, Inc.) provided the mean and standard deviation of continuous data on baseline prognostic factors in the bortezomib arm as well as the aggregate level data on the number of events and patients still at risk per month that was used to calculate the median TTP and OS reported in the publication.

## Analysis

***Patient prognosis***

Baseline characteristics at start of bortezomib treatment were compared, where possible, between HO50-1 patients and APEX and HO50-2 patients.

***Treatment and safety***

Duration and total dosages of bortezomib treatment received were compared between HO50-1 and APEX patients as well as reported toxicities and rates of dose modification and discontinuation of treatment. Comparisons between HO50-1 and HO50-2 patients were not performed since detailed data on treatment in follow-up was not available at the HDC.

***Effectiveness***

Outcomes for effectiveness included response rates, overall survival (OS) and time to progression (TTP), as defined by the European Group for Blood and Marrow Transplant [15]. OS and TTP were calculated from start of bortezomib treatment. Comparison of effectiveness outcomes to other endpoints frequently reported in the trial such as time to and duration of response was not possible since they were most often not recorded in medical charts.

The comparison of effectiveness to efficacy reported to the Dutch Healthcare Board in 2009 was based on the data for HO50-1 patients available up until this time [16]. To show the limitations of the data when restricting the follow-up period to 3 years, updated survival time for HO50-1 patients available in the HDC data as of June 2012 was also shown.

Data was summarized as the mean and standard deviation for continuous variables and as numbers (with percentages) of subjects for categorical variables. For baseline comparisons between the HO50-1 patients versus the APEX and HO50-2 patients, the independent sample t-test was used to test for significant difference in means and chi square test for differences in proportions. The median and 95% confidence interval for TTP and OS in the HO50-1 and HO50-2 patients were computed by the Kaplan-Meier estimator. Since patient-level data was not available for the APEX patients, estimates for the median and 95% confidence interval for TTP and OS for these patients were taken from the data reported in the publication. Consequently, survival curves which compared the TTP and OS between the three patients groups were presented instead of Kaplan-Meier curves, which do not provide the patient-level detail regarding censoring. The Wilcoxon statistic was used to test for equality of survival between the HO50-1

versus APEX and HO50-2 patients. To compute the Wilcoxon statistic for the comparison of HO50-1 versus APEX patients using aggregate data, the survival times were interval-censored for both patient groups, while this was not necessary for the comparison of HO50-1 versus HO50-2 since patient-level data was available. All p-values are 2-sided, and a significant level $\alpha = 0.05$ was used. Analyses were conducted with the statistical software program SAS, version 9.1 (SAS Institute Inc., Cary, NC).

## RESULTS

### Baseline prognosis

#### *HO50-1 versus APEX*

Not all prognostic factors at start of bortezomib treatment were recorded in medical charts, such as serum B2-microglobulin (65% missing) and c-reactive protein (72% missing) (Table 2). Those available showed that HO50-1 patients were younger ($p<0.01$), and presented with a significantly shorter time since diagnosis (mean 3.3 versus 4.2 years; $p<0.01$) despite similar number of previous therapies. Types of previous treatment regimens received were significantly more extensive with HO50-1 patients having been more heavily pre-treated.

#### *HO50-1 versus HO50-2*

HO50-1 patients were more often female (56% versus 69%; $p=0.05$) and younger at the start of bortezomib treatment (mean: 58 versus 60 years; $p=0.05$). HO50-1 patients were not a representative sample of all patients treated in the HOVON-50 trial since they more often received VAD compared to the HO50-2 patients (60% versus 46%). HO50-1 patients also presented with a significantly shorter time since diagnosis than HO50-2 patients (3.3 versus 3.9 years; $p<0.001$) suggesting that the HO50-1 patients were a cohort of patients progressing first from first-line treatment.

### Treatment and safety

#### *HO50-1 versus APEX patients*

Data in medical charts on administration of bortezomib for HO50-1 patients were detailed. Bortezomib was administered as monotherapy in 32% of all administrations, primarily in combination with one or more other drugs (68%), particularly dexamethasone (61%).

In APEX patients, bortezomib dosages (1.3 mg/m$^2$) were administered four times per three-week cycles up to eleven cycles. Approximately 18% and 39% of the HO50-1 and trial patients, respectively, received the planned eight cycles set forth in the trial. The median length of bortezomib therapy was four cycles in HO50-1 patients compared to

**Table 2.** Comparison of baseline characteristics between daily practice and APEX patients

| **Baseline Characteristic** | **APEX4,5 (n=333)** | **HO50-1 (n=72)** | **P-value† (HO50-1 vs. APEX)** | **HO50-2 (n=129)** | **P-value† (HO50-1 vs. HO50-2)** |
|---|---|---|---|---|---|
| Male gender | 56% | 56% | 0.96 | 69% | 0.05** |
| Age at start bortezomib* | 61 (9.8) | 58 (7.8) | 0.02** | 60 (7.2) | 0.05** |
| HOVON50 treatment arm | | | | | 0.05** |
| TAD/VAD | | 40%/60% | | 54%/46% | |
| Received ASCT in HOVON50 | | 22% | | 14% | 0.13 |
| Time since diagnosis (yr)* | 4.2 (3.3) | 3.3 (1.1) | <0.01** | 3.9 (1.9) | <0.001** |
| Missing | 0,6% | 0% | | 0% | |
| Serum B2 (mg/liter)* | 5.1 (5.3) | 3.9 (3.1) | 0.04** | 3.5 (2.1) | 0.60 |
| Missing | 3% | 65% | | 82% | |
| C-reactive protein (mg/liter)* | 10.9 (19.3) | 18.7 (34.5) | 0,33 | 17.4 (25.2) | 0.89 |
| Missing | 10% | 72% | | 81% | |
| Hemoglobin (g/liter)* | 109 (16.8) | 114 (20.9) | 0,08 | 119 (19.4) | 0.12 |
| Missing | 1% | 0% | | 43% | |
| Platelet count (cells/mm$^3$ x 10$^5$)* | 1.98 (8.8) | 1.72 (1.0) | 0,74 | 1.88 (8.2) | 0.29 |
| Missing | 1% | 0% | | 43% | |
| Number of previous lines | | | 0,43 | | 0.11 |
| 1 | 40% | 35% | | 46% | |
| 2 or 3 | 56% | 58% | | 47% | |
| ≥4 | 4% | 7% | | 7% | |
| Type of previous therapy | | | | | |
| Corticosteroids | 98% | 100% | 0.22 | NA | |
| Alkylating agents | 91% | 98% | 0.03** | NA | |
| Anthracyclines | 77% | 98% | <0.01** | NA | |
| Thalidomide | 48% | 40% | 0,85 | NA | |
| Vinca alkaloids | 75% | 100% | <0.01** | NA | |
| SCT or other high-dose therapy | 67% | 94% | <0.01** | NA | |
| Experimental or other | 3% | 6% | 0.28 | NA | |

*Reported as the mean (SD). †The independent sample t-test was used to test for significant difference in means and chi square test for differences in proportions. The log-rank statistic was used to test for differences in median survival measures. **P-values were significant at alpha = 0.05. ASCT: autologous stem cell transplantation; NA: not available; TAD: thalidomide adriamycin dexamethasone; VAD: vincristine adriamycin dexamethasone.

six in the trial. The planned daily practice dosage per cycle was on average lower (20%) for HO50-1 as well as the cumulative total dose received.

Comparison of adverse events was difficult because severity grades for adverse events were often not stated in patient charts. The incidence of any adverse event in HO50-1 patients was less (~80%) than the incidence seen with the trial patients (~100%). A qualitative assessment showed that the type of adverse events was similar to that in the trial, with no unexpected adverse events occurring in HO50-1 patients. Dose modifications occurred in 57% of HO50-1 patients, of which 80% listed an adverse event as a contributing factor. Excluding patients discontinuing due to intercurrent death (n=7) or still being treated at the time of data collection (n=3), 21% of HO50-1 patients (n=15) stopped for reasons of stable disease or end of planned protocol while 65% (n=47) discontinued bortezomib early. The most common reasons were progression of disease (45%) and toxicity (31%), of which 60% were due to peripheral neuropathy. Other reasons for discontinuation included patient condition, refusal, and poor response. Discontinuation due to treatment-related adverse events occurred in 24% of HO50-1 patients compared to 37% in the APEX trial.



***HO50-1 versus HO50-2***

A comparison between HO50-1 and HO50-2 patients in terms of treatment and safety was not performed since this level of detail was not available in the HDC follow-up data.

**Effectiveness**

***HO50-1 versus APEX***

Comparisons of effectiveness and efficacy outcomes for HO50-1 and trial patients are shown in Table 3. Available follow-up for HO50-1 patients as of May 2009 was lower compared to APEX patients (median: 15.9 versus 22 months). Overall response rates were not significantly different between the two groups (50% versus 38%), namely on account of similar rates for complete response. Other response categories were however different, with HO50-1 patients more often achieving a very good, partial or minor response (54% versus 41%; $p<0.01$), less often achieving no change (3% versus 43%; $p<0.01$) and more frequently progressive disease (19% versus 7%; $p<0.01$) or were not evaluated (15% versus 3%; $p<0.01$). A total of 51 HO50-1 patients achieved progression (71%) for which 7 (14%) the date of progression was missing from the patient chart requiring censoring at start of next treatment. TTP in daily practice was found to be similar to the trial (median: 6.8 versus 6.2 months; $p=0.70$). The survival curves depicted in Figure 1 show that TTP was similar in early points in time but eventually diverge in favor of the trial patients. TTP was not significantly different for HO50-1 patients compared to the trial when stratifying by HOVON-50 treatment.

Figure 2 shows the discrepancy in OS for HO50-1 patients when right-censored as of May 2009 and June 2012. Using the data available as of May 2009, the median OS (17.2 versus 29.8 months; p<0.01) for HO50-1 patients was significantly shorter compared to trial patients with HO50-1 patients consistently experiencing poorer survival, especially in patients receiving VAD versus TAD as upfront therapy (median: 16.1 versus 29.8; p=0.02).

**Table 3.** Effectiveness outcomes for bortezomib in daily practice versus APEX patients

| | **APEX[4,5] (n=333)** | **HO50-1 (n=72)** | **P-value† (HO50-1 vs. APEX)** | **HO50-2 (n=129)** | **P-value† (HO50-1 vs. HO50-2)** |
|---|---|---|---|---|---|
| Mean total follow-up as of May 2009 (months) | 18.2 | 13.7 | <0.01** | | |
| Patients still in follow-up as of May 2009 | 56% | 51% | 0.49 | | |
| Mean total follow-up as of June 2012 (months) | | 23.9 | <0.01** | 23 | 0.61 |
| Patients still in follow-up as of June 2012 | | 29% | <0.01** | 34% | 0.47 |
| Overall response (CR/VGPR/PR) | 38% | 50% | 0.10 | NR | |
| CR | 6% | 8% | 0.46 | 3% | 0.97 |
| VGPR/PR/MR | 41% | 54% | <0.001** | 53% | 0.92 |
| No change | 43% | 3% | <0.001** | 10% | 0.06 |
| Progressive disease (and relapse) | 7% | 19% | <0.001** | 25% | 0.39 |
| Not evaluated | 3% | 15% | <0.001** | 9% | 0.14 |
| VAD arm | | | | 65% | 0.96 |
| | **Median (95% CI)** | | **P-value‡** | **Median (95% CI)** | **P-value‡** |
| Time to progression all patients (months) | 6.2 (NR) | 6.8 (5.4 - 8.5) | 0.70 | 7.8 (6.1 - 9.3) | 0.01** |
| TAD arm | | 7.4 (5.4 - 9.4) | 0.33 | 8.0 (6.1 - 11.5) | 0.40 |
| VAD arm | | 6.5 (3.6 - 8.5) | 0.59 | 7.7 (5.2 - 10.2) | 0.19 |
| Overall survival all patients as of May 2009 (months) | 29.8 (23.2-NE) | 17.2 (12.9 - 32.5) | <0.01** | | |
| TAD arm | | 17.2 (11 - 32.5) | 0.09 | | |
| VAD arm | | 16.1 (10.9 - NE) | 0.02 | | |
| Overall survival all patients as of June 2012 (months) | 29.8 (23.2-NE) | 23.1 (14 - 34.1) | 0.06 | 24.1 (17.7 - 30.6) | 0.72 |
| TAD arm | | 21.1 (12.9 - 38.1) | 0.14 | 22.1 (16.2 - 33.5) | 0.88 |
| VAD arm | | 25.7 (12.5 - 36.3) | 0.20 | 28.7 (17.7 - 31.1) | 0.76 |

†The independent sample t-test was used to test for significant difference in means and chi square test for differences in proportions. ‡Wilcoxan statistic was used to compare hazard functions. **P-values were significant at alpha = 0.05. CR: complete response; MR: minor response; NE: not evaluable; NR: not reported; PR: partial response; TAD: thalidomide adriamycin dexamethasone; VAD: vincristine adriamycin dexamethasone; VGPR: very good partial response.

As of June 2012, the median OS (23.1 versus 29.8; p=0.06) for daily practice patients was slightly lower for HO50-1 patients compared to the trial patients but was not influenced by HOVON-50 treatment.

***HO50-1 versus HO50-2***

As of June 2012, the median follow-up available for HO50-1 and HO50-2 patients was similar (20.2 versus 19.2 months) (Table 3). HO50-1 patients achieved similar responses

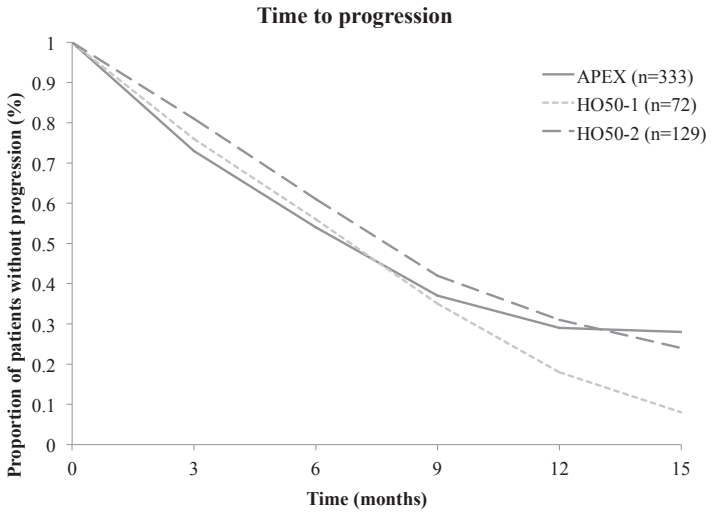



**Figure 1.** Time to progression curves for APEX and daily practice patients.

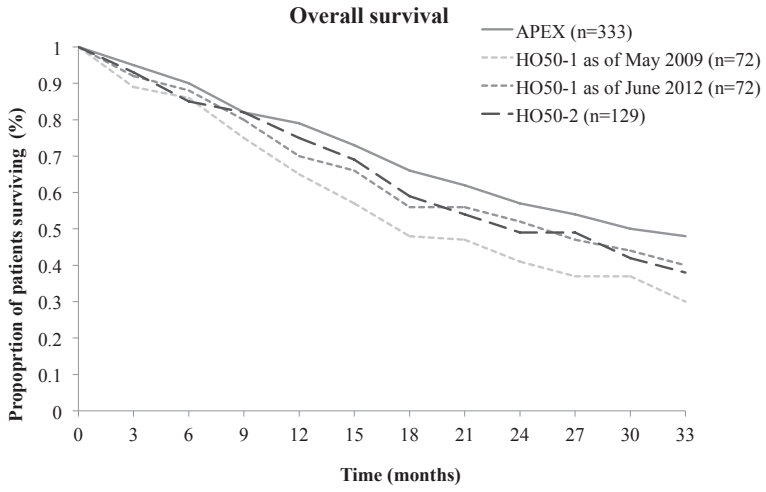



**Figure 2.** Overall survival curves for APEX and daily practice patients.

to HO50-2 patients. Despite similar response profiles, TTP for HO50-1 patients was lower (median: 6.8 versus 7.8 months; p=0.01) (Figure 1), which remained consistent when stratifying by HOVON-50 treatment arm. Patients receiving VAD in first line, regardless of group, progressed sooner than patients receiving TAD. In terms of survival, there was no significant difference in median OS (23.1 versus 24.1 months; p=0.72) (Figure 2). OS from start of bortezomib was less favorable for patients receiving TAD during first-line for both groups.

## DISCUSSION

Outcomes research into the use and effectiveness of bortezomib in advanced MM generated evidence on the dynamics of care in daily practice and the limitations to observational data. It was useful for addressing *who* received bortezomib and *how* bortezomib was administered in daily practice. Our results show that prognosis was less favorable for the first daily practice patients treated with bortezomib (i.e., HO50-1) compared to the pivotal trial and daily practice patients eligible at a later point in time (i.e., HO50-2). An important limitation however was that not all prognostic factors could fully be assessed due to missing values in medical charts. Further differences could have been present according to other prognostic factors found to substantially contribute to adverse outcomes, such as cytogenetic markers and comorbidities [17, 18], but were missing entirely in patient charts. Evidence regarding use revealed more conservative treatment (i.e., lower planned dosages and frequent dose modification) in HO50-1 patients alongside treatment variation and premature discontinuation, which are treatment practices that could reduce the treatment effect.

We showed that the real-world effectiveness of bortezomib did not confirm its efficacy and demonstrated the limitations to the evidence on effectiveness. It is likely that the discrepancy between efficacy and effectiveness is explained by differences in prognosis and use of the drug. The discrepancy in the available data on response and progression rates could also be due to a number of differences between daily practice and a trial setting such as the frequency of follow-up, criteria or the methods used to evaluate response or progression. A proper evaluation of TTP using observational data was further challenged by missing data, lending this outcome susceptible to overestimation since missing data required censoring the date of progression at the date of start of next treatment. Estimating time-to-next-treatment instead of progression rates in daily practice is probably more appropriate, as recommended by Durie et al. (2006) [19], but would not allow for comparisons with trial endpoints. Given these caveats of observational data, decision makers should be warned about making invalid comparisons between

trial-based and daily practice outcomes using response and progression rates. OS, on the other hand, is a more reliable outcome measure for comparisons, since it is clear in the charts when death has been reached. We do recommend however that OS be evaluated for MM patients in terms of previous treatment, which was shown to be an independent predictor [14, 20].

The timeframe impacted the results as it influenced patient selection in terms of sample size and prognosis. A small number of patients treated with bortezomib could initially be identified in daily practice (i.e., n=72) which is related to the low prevalence of the disease and the slow rate of bortezomib uptake in daily practice [21]. The patients included in the HOVON-50 trial were initially the most favorable patients in daily practice since they would otherwise not have been eligible for this trial. However, the worst off patients from this trial were selected in the HO50-1 group since they were the early failures. By extending the timeframe from 2009 to 2012, as demonstrated in the HO50-2 group, more patients with a more favorable prognosis became eligible for bortezomib, thus demonstrating better outcomes. A recent population-based study reported similar findings with better outcomes in patients with late relapse due to increased use of salvage high dose therapy and introduction of new drugs [22]. We believe that an insufficient timeframe led to the problem of incomplete follow-up, resulting in OS estimates not representative of the entire cohort but instead those experiencing early death. Since the time of our study, the Dutch authority has extended the evaluation period from three to four years presumably, in part, due to the realization that more time is needed to collected the required data [23]. There remains however no consensus on the timeframe needed to conduct outcomes research since the timeframe is largely dependent on the decision problem and indication.

The study limitations pertain to its retrospective design, which precluded prospective data collection on patient-reported outcomes and known prognostic factors generally not included in medical charts. The drawbacks of any observational study, be they retrospective, prospective or registry-based, is that there is no set treatment protocol and physicians are free to manage patients[24], lending outcomes to be affected by a number of factors unrelated to the safety and efficacy of an intervention[25]. Despite its high costs, the advantage of a population-based registry, however, is that it includes *all* patients, which avoids the problem of introducing selection bias when sampling from the target population. On the other hand, provider-introduced selection bias (i.e., confounding by indication) remains a possibility even for registry studies. Since conducting the present study, a population-based registry for hematological malignancies has been implemented in the Netherlands [26], but whether it overcomes the limitations shown here is currently being evaluated.

The implications of these findings are most relevant for decision makers requesting outcomes research for new drugs in MM and researchers performing such studies. The future of treatment for these patients remains promising with new drugs expected to come available within the coming years [27]. Like bortezomib, these drugs will likely be approved first for the indication of advanced MM, as this is the current trend for introducing innovation into the treatment paradigm. Decision makers requesting future outcomes research studies should be made aware of the extensive treatment variation that in exists in MM and how it influences the usefulness of collecting outcomes research data. Consideration should also be given to the influence of the dynamics in care during a new drug's uptake in daily practice. In the future, decision makers should carefully select the outcomes measures that are necessary for informed decision making about drug reimbursement instead of requesting a comprehensive set of all possible outcomes. Clinicians and researchers performing these studies should be involved in outcomes research from the start since they can inform decision makers about the type and quality of evidence that can be generated in daily practice.

In conclusion, an observational study in advanced MM generated evidence on the use and health impact of a new drug in the first eligible patients. However, the drug's real-world effectiveness was not representative of the target patient population but was instead limited to a sample of patients with unfavorable prognosis. When making inferences regarding the effectiveness of novel drugs in MM, decision makers should give careful consideration beforehand to the possibilities as well as the limitations of outcomes research during the drug's uptake period.

## ACKNOWLEDGEMENTS

The authors would like to thank Kevin Liu and Suzanne Viselli, both of Janssen Research and Development, Raritan NJ (USA), Don Robinson Jr., Janssen Global Services, Malvern PA (USA), and Michel van Agthoven, Janssen-Cilag, Tilburg (NE) for providing the supplemental APEX data, all participating hospitals and hospital data-managers and cancer registration (IK) staff, and Danielle Ammerlaan, Hedwig Blommestein, 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.

## REFERENCES


1. Palumbo A, Anderson K. Multiple myeloma. N Engl J Med. 2011;364(11):1046-60.
2. Kumar SK, Rajkumar SV, Dispenzieri A, Lacy MQ, Hayman SR, Buadi FK, et al. Improved survival in multiple myeloma and the impact of novel therapies. Blood. 2008;111(5):2516-20.
3. Brenner H, Gondos A, Pulte D. Recent major improvement in long-term survival of younger patients with multiple myeloma. Blood. 2008;111(5):2521-6.
4. Richardson PG, Sonneveld P, Schuster M, Irwin D, Stadtmauer E, Facon T, 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.
5. Richardson PG, Sonneveld P, Schuster MW, Irwin D, Stadtmauer EA, Facon T, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med. 2005;352(24):2487-98.
6. Kane RC, Bross PF, Farrell AT, Pazdur R. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. Oncologist. 2003;8(6):508-13.
7. Committee for Medicinal Products for Human Use (CHMP). Velcade: European public assessment report - Summary for the public. European Medicines Agency (EMA); 2009.
8. Williamson S. Patient access schemes for high-cost cancer medicines. Lancet Oncol. 2010;11(2):111-2.
9. Kuijpers MR, Toenders WGM. Assessment procedure for inpatient drugs [Procedure beoordeling intramurale geneesmiddelen]. Diemen: Dutch Health Care Insurance Board; 2006.
10. Delwel G. Guidance for outcomes research [Leidraad voor uitkomstenonderzoek]. Diemen: Dutch Health Care Insurance Board; 2008.
11. Krumholz HM. Outcomes research: generating evidence for best practice and policies. Circulation. 2008;118(3):309-18.
12. Marko NF, Weil RJ. The role of observational investigations in comparative effectiveness research. Value Health. 2010;13(8):989-97.
13. Lu CY. Observational studies: a review of study designs, challenges and strategies to reduce confounding. Int J Clin Pract. 2009;63(5):691-7.
14. Lokhorst HM, van der Holt B, Zweegman S, Vellenga E, Croockewit S, van Oers MH, et al. A randomized phase 3 study on the effect of thalidomide combined with adriamycin, dexamethasone, and high-dose melphalan, followed by thalidomide maintenance in patients with multiple myeloma. Blood. 2010;115(6):1113-20.
15. Blade J, Samson D, Reece D, Apperley J, Bjorkstrand B, Gahrton G, et al. Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT. European Group for Blood and Marrow Transplant. Br J Haematol. 1998;102(5):1115-23.
16. Franken MG, Gaultney JG, Blommestein HM, Tan SS, van Gils CWM, Huijgens PC, et al. Pilot outcomes research: effects and costs of bortezomib in relapsed or refractory multiple myeloma Dutch Healthcare Insurance Board (CVZ) 2010.
17. Stewart AK, Bergsagel PL, Greipp PR, Dispenzieri A, Gertz MA, Hayman SR, et al. A practical guide to defining high-risk myeloma for clinical trials, patient counseling and choice of therapy. Leukemia. 2007;21(3):529-34.
18. Offidani M, Corvatta L, Polloni C, Centurioni R, Visani G, Brunori M, et al. Assessment of Vulnerability Measures and Their Effect on Survival in a Real-Life Population of Multiple Myeloma Patients Registered at Marche Region Multiple Myeloma Registry. Clin Lymphoma Myeloma Leuk. 2012.
19. Durie BG, Harousseau JL, Miguel JS, Blade J, Barlogie B, Anderson K, et al. International uniform

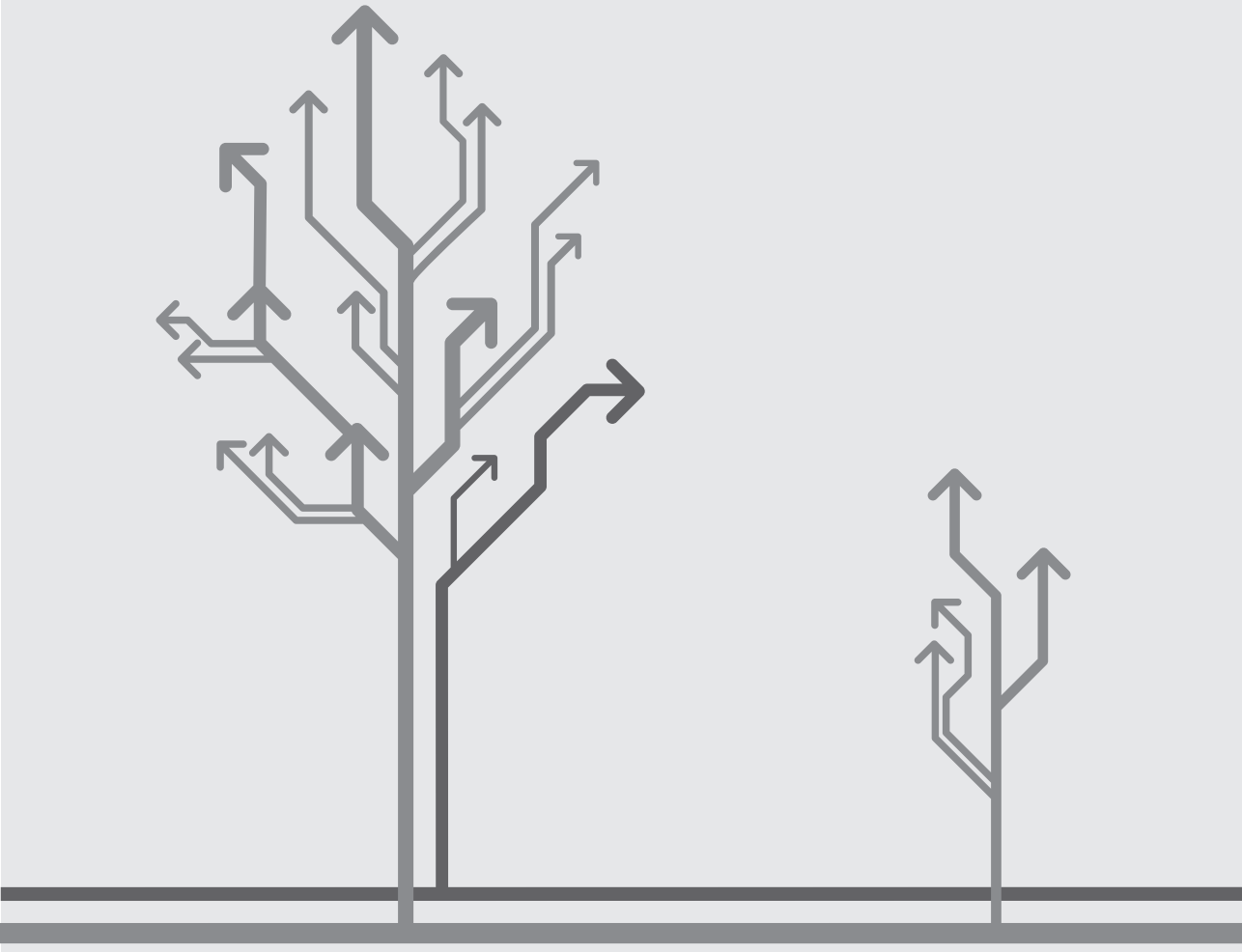
response criteria for multiple myeloma. Leukemia. 2006;20(9):1467-73.
20. Vogl DT, Stadtmauer EA, Richardson PG, Sonneveld P, Schuster MW, Irwin D, et al. Impact of prior therapies on the relative efficacy of bortezomib compared with dexamethasone in patients with relapsed/refractory multiple myeloma. Br J Haematol. 2009;147(4):531-4.
21. Blommestein H, Franken M, Uyl-De Groot C, editors. Accessibility and utilization of bortezomib in dutch daily practice. Internatioanl Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2010; Prague, Czech Republic: Value in Health.
22. Vangsted AJ, Klausen TW, Andersen NF, Abildgaard N, Gang AO, Gregersen H, et al. Improved survival of multiple myeloma patients with late relapse after high-dose treatment and stem cell support, a population-based study of 348 patients in Denmark in 1994-2004. Eur J Haematol. 2010;85(3):209-16.
23. Delwel G. Guidance for outcomes research [Leidraad voor uitkomstenonderzoek]. Diemen: Dutch Health Care Insurance Board, 2008.
24. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. Pharmacoeconomics. 1999;15(5):423-34.
25. Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. BMJ. 2007;335(7621):648-50.
26. Huijgens PC, Posthuma EFM, Coebergh JWW, Van de Poll-Franse LV, Uyl-de Groot CA, Sonneveld P. A population-based registry for hemato-oncology. [Een 'population-based' registry voor hemato-onologie.] Nederlands Tijdschrift voor Hematologie 2010;7:321-5.
27. Mahindra A, Cirstea D, Raje N. Novel therapeutic targets for multiple myeloma. Future Oncol. 2010;6(3):407-18.

# Chapter 5.

## Real-world healthcare costs of relapsed/refractory multiple myeloma during the era of novel cancer agents

Jennifer G. Gaultney, Margreet G. Franken, Siok Swan Tan, William K. Redekop, Peter C. Huijgens, Pieter Sonneveld, Carin A. Uyl-de Groot

*Published in: Journal of Clinical Pharmacy and Therapeutics. 2013 Feb;38(1):41-7.*



## ABSTRACT

High costs of novel agents increasingly put pressure on limited health care budgets. Demonstration of their real-world costs and cost-effectiveness is often required for reimbursement. However, few published economic evaluations of novel agents for multiple myeloma exist. Moreover, existing cost analyses were heavily based on conventionally treated patients. We investigated real-world healthcare costs of relapsed/refractory multiple myeloma in Dutch daily practice. A retrospective medical chart review was conducted for 139 patients treated between January 2001 and May 2009. Total monthly costs attributable to each cost component were described across all regimens and for bortezomib-, thalidomide- and lenalidomide-based treatment regimens. Mean monthly total costs (€3,981) varied depending on the sequence of therapy (range: €442 – €31,318). Significant cost drivers across all regimens included costs of therapy and hospital admissions. The acquisition costs for novel agents in particular accounted for 32% of mean total monthly costs. Prognostic factors associated with increased mean total monthly costs in multivariate regression analysis included low platelet counts ($p=0.01$) and worsening performance status ($p<0.001$). Mean total monthly costs of bortezomib- and lenalidomide-based regimens were significantly higher than those for thalidomide-based regimens in second, third and fourth treatment line. Real-world costs during treatment of RRMM vary greatly. Cost drivers include hospital admissions and acquisition costs of novel agents. Costs also vary by prognostic factors and treatment-related resource use. Future studies assessing the costs of combination therapy consisting of two or more novel agents are encouraged.

## INTRODUCTION

Multiple myeloma (MM) is an incurable malignancy caused by the autonomous growth of malignant plasma cells [1]. Worldwide it accounts for approximately 0.8% of all cancer diagnoses and 0.9% of cancer deaths in 2002. In Europe, the annual incidence of MM is 4.5-6 per 100,000 people with a median age at diagnosis between 63 and 70 years [2].

Relapse always occurs and patients eventually become refractory to all available treatments. Conventional therapies achieve a median survival duration of three to four years, while autologous stem cell transplant (SCT) has extended the median life expectancy from diagnosis to five years [3]. New generation therapies, such as bortezomib, thalidomide and lenalidomide, promise even further gain in life expectancy [4], while at significantly increased acquisition costs.

Given the limited budget for healthcare, many decision makers require demonstration of cost-effectiveness and budget impact and are interested in the 'real-world' costs of expensive drugs. However, few published economic evaluations exist for novel agents in MM [5]. Existing cost-effectiveness analyses were based on cost analyses using a synthesis of data sources and expert opinion and not patient-level data to estimate resource use [6-8].

Cost studies in MM have been conducted using patient-level data from the hospital perspective in Sweden and, more recently, France and Italy [9-11]. The costs reported for Swedish patients were mainly based on conventional therapies [9] and the results from Italy did not provide information on the use of novel agents [11]. A more recent cost study conducted in France found that total costs were substantial mainly due to novel agents, though this conclusion was based mainly on the increasing costs due to bortezomib [10]. It remains unclear to what extent the increased costs of treating MM are caused by the acquisition costs of therapy.

To improve upon previous research, we investigated the real-world healthcare costs and significant drivers of increased costs for treatment of relapsed/refractory MM (RRMM) in daily practice in the Netherlands during the era of novel agents.

## METHODS

### Patient population

The patient population included patients previously enrolled in the phase III HOVON-50 trial which investigated the treatment effect of thalidomide in newly diagnosed symptomatic MM patients aged 18–65 years [12]. Patients in this trial received upfront therapy with chemotherapy alone or in combination with thalidomide, followed by high dose melphalan and SCT. Patients relapsing from protocol-based upfront therapy and treated for RRMM in daily practice were included in our study.

Detailed case reports were collected retrospectively using medical charts and electronic records for patients treated between January 2001 and May 2009. Data relates to the time period when patients relapsed or progressed from upfront therapy until last known date of follow-up.

### Resource use and unit prices

Total costs for individual patients were determined by the identification of resource use and unit costs of all cost components. The analysis was based on the resource use of the full patient sample including patients still treated at the end of data collection. All costs were based on Euro 2009 cost data. Where necessary, costs were adjusted to 2009 using the general price index from the Dutch Central bureau of Statistics [13].

Unit costs of hospital visits and admissions, procedures and concomitant medication are presented in supplemental Table S1. The unit cost calculations for hospital visits and admissions were based on detailed micro-costing studies reflecting full hospital costs, including overhead costs [14]. Some unit costs were weighted for their origin to reflect the distribution of patients among university versus general hospitals in Dutch daily practice. The cost of radiotherapy, surgical procedures, and medical imaging services were valued using the fees issued by the Dutch Healthcare Authority [15]. Unit costs for laboratory services were based on a detailed inventory of the resource use of 12 patients (approximately 1,000 tests). Unit costs of concomitant treatment were acquired from the cost guidelines available from the national pharmaceutical formulary drafted by the Dutch Healthcare Insurance Board [16]. Detailed daily concomitant treatment costs were determined for 18 patients and were considered representative of the remaining 121 patients.

Unit costs for therapy regimens are shown in supplemental Table S2. The unit cost of donor lymphocyte infusions was taken from published micro-costing studies and corrected for resource use to prevent double counting [17]. Unit prices for autologous and

allogeneic SCT were taken from a cost study conducted for the Netherlands [18], which are exclusive to the costs of the pre-transplantation and transplantation phase except for hospital visits and admissions as these costs were included separately.

The price for thalidomide is substantially lower in the Netherlands compared to the rest of Europe. To allow generalizability to the European perspective, unit prices for thalidomide were based on estimates reported by previously published cost studies [9, 10]. All other prices for therapy regimens were acquired from the national pharmaceutical formulary drafted by the Dutch Healthcare Insurance Board [16].

**Statistical analyses**

To correct for differences in follow-up between censored patients and patients still being treated at the time of data collection, total costs per patient were corrected by the patient's total follow-up time and total monthly costs were reported. The association between prognostic factors at start of RRMM treatment (independent variables) and average total monthly costs across all lines (dependent variable), controlling for patient characteristics, was modeled by means of a generalized linear regression model with a gamma distribution and log link function [19]. Only prognostic factors with 10% or fewer missing values and a significant bivariate effect were entered into the multivariate model. The final model was chosen based on the likelihood ratio test for goodness-of-fit.

To compare costs between novel agent-based regimens, groups were defined as follows: 1) *bortezomib-based* included any combination with bortezomib; 2) *thalidomide-based* included any combination with thalidomide except in combinations with bortezomib and lenalidomide; 3) *lenalidomide-based* included any combination with lenalidomide except in combination with bortezomib. Statistical comparisons of continuous variables were performed using the Kruskal-Wallis test for two or more related samples followed by pair-wise tests with the Mann-Whitney test to identify which groups differed significantly from the others. The Fisher's exact text and Cochran-Armitage test for trend were used for categorical or ordinal variables with two or more than two categories, respectively. Statistically significant differences were concluded for comparisons with a p-values less than or equal to an alpha of 0.05. All statistical analyses were conducted with the statistical software program SAS, version 9.1 (SAS Institute Inc., Cary, NC).

**Table 1.** Patient characteristics and prognostic factors at start of relapsed/refractory treatment

| | Frequency (%) | Mean | Median | Range | Missing |
|---|---|---|---|---|---|
| **Patient characteristics** | | | | | |
| Male Gender | 58% | | | | 0% |
| Age | | 57 | 60 | 34-69 | 0% |
| Durie-Salmon stage at diagnosis: | | | | | 0% |
| I A/B | 2% | | | | |
| IIA | 21% | | | | |
| IIB | 1% | | | | |
| IIIA | 66% | | | | |
| IIIB | 10% | | | | |
| Months from diagnosis to RRMM : | | | | | 0% |
| All patients (n=139) | | 25 | 24.9 | 1.9-61.4 | |
| First-line therapy | | | | | |
| TAD (-SCT) (n=39) | | 25.6 | 24.9 | 2.1-55.6 | |
| TAD (+SCT) (n=8) | | 29 | 25.6 | 18.5-54.7 | |
| VAD (-SCT) (n=71) | | 23.2 | 24.6 | 1.9-57.9 | |
| VAD (+SCT) (n=21) | | 28.8 | 25 | 13-61.4 | |
| Best response to initial therapy: | | | | | 6% |
| Complete response | 12% | | | | |
| Partial response | 66% | | | | |
| Minor response | 7% | | | | |
| No change | 4% | | | | |
| Progressive disease | 5% | | | | |
| **Prognostic factors** | | | | | |
| Albumin (g/l) [mean (sd)] | | 38.9 | 39 | 16.6-58.9 | 34% |
| Serum B2 (mg/l) | | 3.8 | 2.8 | 1.1-16.7 | 69% |
| C-reactive protein (mg/l) | | 23.4 | 6 | 1-171 | 68% |
| Hemoglobin (mmol/l) | | 7.5 | 7.7 | 21-10 | 31% |
| Platelet count (*10'^9/l) | | 220 | 191 | 10-828 | 2% |
| Creatinine clearance (mmol/l) | | 6.84 | 6 | 1-16 | 82% |
| Plasma cell infiltration > 50% | 15% | | | | 51% |
| Present with neurotoxicity | 35% | | | | 6% |
| WHO performance status: | | | | | 8% |
| Asymptomatic | 44% | | | | |
| Symptomatic but ambulatory | 35% | | | | |
| In bed < 50% | 8% | | | | |
| In bed > 50% | 3% | | | | |
| Bedridden | 3% | | | | |

RRMM: relapsed/refractory multiple myeloma; SCT: stem cell transplantation; SD: standard deviation; TAD: thalidomide, adriamycin and dexamethasone; VAD: vincristine, adriamycin and dexamethasone.

**Table 2.** Breakdown of mean monthly costs during relapsed/refractory treatment

| | 2nd line (n=139) | | 3rd line (n=90) | | 4th line (n=54) | | All lines (n=139)* | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Min-Max | Mean | Min-Max | Mean | Min-Max | Mean (%) | Min-Max |
| Months of follow-up | 13.8 | 0.3 - 63 | 8.4 | 0.8 - 42.7 | 7.1 | 1.1 - 22.2 | 24 | 0.6 - 70.1 |
| Alive at data collection | | | | | | | 49% | |
| **Mean Monthly Costs (Euro 2009)** | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| **Hospital visits** | 338 | 294 | 531 | 326 | 470 | 350 | 364 | 258 |
| Outpatient | 191 | 132 | 266 | 162 | 260 | 224 | 201 | 123 |
| Daycare | 138 | 210 | 254 | 267 | 198 | 254 | 152 | 176 |
| Emergency room | 7 | 22 | 10 | 25 | 9 | 20 | 9 | 22 |
| Telephone consults | 2 | 3 | 2 | 3 | 4 | 11 | 2 | 2 |
| **Hospital admissions** | 1007 | 2,686 | 1,065 | 1,818 | 925 | 1,714 | 1,201 | 2,268 |
| Intensive care unit | 101 | 607 | 81 | 476 | 147 | 995 | 168 | 683 |
| Hematology or surgery ward | 905 | 2,552 | 985 | 1,741 | 778 | 1,112 | 1,032 | 2,066 |
| **Radiotherapy** | 124 | 465 | 108 | 238 | 75 | 197 | 83 | 146 |
| **Surgery** | 51 | 265 | 60 | 243 | 40 | 288 | 63 | 223 |
| **Diagnostics** | 402 | 802 | 485 | 422 | 433 | 446 | 432 | 757 |
| **Total concomitant medication costs** | 512 | 788 | 571 | 428 | 587 | 503 | 557 | 676 |
| **Acute** | | | | | | | | |
| Anti-bacterial / fungal / viral (acute infection) | 13 | 34 | 25 | 44 | 14 | 21 | 15 | 32 |
| Anemia (red blood cell transfusion) | 50 | 267 | 53 | 141 | 87 | 270 | 60 | 205 |
| Thrombocytopenia (platelet infusion) | 79 | 509 | 52 | 226 | 57 | 212 | 82 | 462 |
| **Chronic/Prophylactic** | | | | | | | | |
| Anti-bacterial / fungal / viral (prophylaxis) | 78 | 84 | 98 | 83 | 83 | 84 | 86 | 74 |
| Biphosponates | 109 | 52 | 107 | 54 | 111 | 50 | 107 | 47 |
| Anemia (erythropoietin injection) | 8 | 26 | 11 | 30 | 16 | 36 | 11 | 27 |
| Neurotoxicity | 14 | 63 | 32 | 91 | 27 | 84 | 20 | 67 |
| Gastro-intestinal | 138 | 144 | 163 | 143 | 160 | 144 | 149 | 127 |
| Analgesics | 24 | 25 | 31 | 25 | 32 | 25 | 26 | 22 |
| **Active therapy** | 822 | 975 | 1,782 | 1,787 | 2,090 | 2,155 | 1,106 | 979 |
| Bortezomib | 363 | 954 | 1,049 | 1,600 | 602 | 1,293 | 465 | 738 |
| Thalidomide | 315 | 419 | 117 | 296 | 76 | 242 | 219 | 287 |
| Lenalidomide | 59 | 379 | 533 | 1,406 | 1,376 | 2,203 | 363 | 762 |
| Dexamethasone | 6 | 13 | 6 | 8 | 7 | 7 | 5 | 8 |
| Adriamycin | 5 | 21 | 6 | 21 | 6 | 24 | 5 | 16 |

**Table 2.** (Continued)

| | 2nd line (n=139) | | 3rd line (n=90) | | 4th line (n=54) | | All lines (n=139)* | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Min-Max | Mean | Min-Max | Mean | Min-Max | Mean (%) | Min-Max |
| Vincristine | 1 | 6 | 1 | 5 | 1 | 7 | 1 | 5 |
| Melphalan | 3 | 9 | 2 | 10 | 4 | 15 | 3 | 7 |
| Prednisone | 1 | 3 | 0 | 1 | 1 | 5 | 1 | 2 |
| Cyclophosphamide | 0 | 1 | 1 | 4 | 2 | 6 | 1 | 1 |
| Donor leukocyte infusions | 68 | 249 | 66 | 205 | 15 | 49 | 43 | 111 |
| Interferon alpha | 0 | 0 | 1 | 6 | 0 | 0 | 1 | 3 |
| Other/experimental | 1 | 15 | 0 | 0 | 1 | 7 | 1 | 5 |
| **Stem cell transplantation** | 211 | 825 | 191 | 838 | 65 | 346 | 173 | 447 |
| **Total monthly costs** | 3,469 | 4,408 | 4,792 | 2,983 | 4,685 | 3,017 | 3,981 | 3,538 |

SD: standard deviation. *These costs are calculated per month over the patient's entire follow-up

## RESULTS

### Baseline characteristics

Baseline characteristics of all patients are shown in Table 1. The majority of patients received VAD (n=92; 65%) as upfront treatment and 20% (n=29) received SCT as part of upfront therapy. Many patients responded to upfront treatment, with 66% having achieved a partial response. Regarding known prognostic factors, most patients presented at start of RRMM treatment with favorable performance status.

### Costs during relapsed/refractory treatment

A total of 139, 90, and 54 patients received second, third, and fourth line treatment with 20 patients receiving 5 or more lines of therapy. Mean monthly costs incurred during all treatment regimens are presented in Table 2. Generally, costs varied substantially within the patient population with mean total costs during the entire follow-up of approximately €3,981 per month (range: €442 – €31,318) or €72,968 per patient (range: €1,423 – €232,685). Mean total costs during second and third line were approximately €3,469 per month (range: €353 – €31,318) or €30,982 per patient (range: €1,067 – €169,132) and €4,792 per month (range: €421 – €13,437) or €33,319 per patient (range: €1,193 – €191,124), respectively. Mean total costs during fourth line were €4,685 per month (range: €490 – €15,881) or €32,889 per patient (range: €1,055 – €144,967) which was higher compared to second and third line due to more frequent use of novel agents, especially lenalidomide. Generally, costs of therapy and hospital admissions were drivers of increased costs across all regimens. Supplemental Table S3 provides additional information describing the distribution of these costs, such as the median, range, and the parameters for the gamma distribution.

There was large variation in costs for therapy regimens within the lines as well as across lines. Thalidomide was most commonly administered in second line (n=71; 75%

**Table 3.** Generalized linear model results describing the association between prognostic factors and total monthly costs

| | Average total monthly costs in RRMM (n=126*) | | | | | |
|---|---|---|---|---|---|---|
| Independent variable | β | SE | P-value | $e^{\beta}$ | 95% CI | |
| Constant | 9.0705 | 0.3927 | <0.001 | 8695 | 4024 | 18034 |
| Age start RRMM (years) | -0.0186 | 0.0065 | 0.0043 | 0.982 | 0.969 | 0.994 |
| Death at time of data collection (yes/no) | 0.4278 | 0.0984 | <0.001 | 1.534 | 1.265 | 1.860 |
| HOVON-50 thalidomide maintenance** | 0.2956 | 0.1082 | 0.0063 | 1.344 | 1.087 | 1.661 |
| Platelet count start RRMM (*10^9/l) | -0.0009 | 0.0003 | 0.0107 | 0.999 | 0.999 | 1.000 |
| Worsening WHO performance status | 0.1959 | 0.0499 | <0.001 | 1.216 | 1.103 | 1.341 |

*Fewer observations included in the model due to missing values
**Reference category is none or interferon alpha
CI: confidence interval; SE: standard error; RRMM: relapsed/refractory multiple myeloma; WHO: world health organization.

all administrations); bortezomib was administered most often in third line (n=35; 44% all administrations); and lenalidomide was most often administered in fourth line (n=20; 40% all administrations). A treatment sequence pattern of treatment initially with thalidomide followed by bortezomib explained these findings as the majority of patients receiving thalidomide in second line had previously received non-thalidomide-based regimens as upfront therapy (i.e., VAD) (90%), which was statistically significantly higher compared to the patients receiving bortezomib (44%) (p<0.0001).

### Regression of prognostic factors on costs

Controlling for age, survival status at data collection, and upfront treatment, an association was found between known prognostic factors at start of RRMM and mean total monthly costs of RRMM across all lines (Table 3). The most statistically significant associations were observed for the prognostic factors WHO performance status and platelet count. Worsening WHO performance status was associated with a 22% increase in mean total monthly costs (P<0.001). A higher platelet count was associated with decreased mean total monthly costs (p=0.0107).

### Costs during novel agent regimens

Mean monthly costs during bortezomib- , thalidomide- and lenalidomide-based regimens are described in Table 4. Mean total monthly costs were significantly lower in second, third and fourth line during thalidomide-based regimens (€2,684, €3,393 and €4,308, respectively) compared to bortezomib-based (€4,814, €5,966 and €6,260, respectively) and lenalidomide-based (€4,215, €5,029 and €5,114, respectively) regimens. The lower costs for thalidomide-based regimens was due to the lower acquisition costs for novel agents (€592, €700 and €672) compared to bortezomib-based (€2,089, €2,718 and €2,825) and lenalidomide-based (€2,049, €3,423 and €3,651) regimens. Mean total monthly and acquisition costs for novel agents did not differ significantly between

**Table 4.** Breakdown of mean monthly costs for bortezomib-, thalidomide- and lenalidomide-based regimens during relapsed/refractory treatment

| | 2nd line | | | 3rd line | | | 4th line | | |
|---|---|---|---|---|---|---|---|---|---|
| **Targeted therapy** | **Bortezomib (n=25)** | **Thalidomide (n=71)** | **Lenalidomide (n=4)** | **Bortezomib (n=35)** | **Thalidomide (n=14)** | **Lenalidomide (n=14)** | **Bortezomib (n=12)** | **Thalidomide (n=6)** | **Lenalidomide (n=20)** |
| | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** | **Mean (range)** |
| Months in this line | 9.9 (2.4 - 19.5) | 15.7 (0.6 - 62.8) | 15.8 (10.6-23.4) | 7.2 (1.2 - 17.6) | 11.4 (4 - 42.7) | 8.2 (0.8 - 20.3) | 6.3 (1.1 - 16.4) | 6.5 (3.3 - 13.6) | 8.3 (1.9 - 22.2) |
| Alive at data collection | 36% | 18% | 100% | 9% | 21% | 36% | 25% | 17% | 80% |
| **Mean Monthly Costs (Euro 2009)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** |
| Hospital visits | 606 (257)* | 239 (130) | 128 (81) | 793 (300)* | 317 (154) | 314 (197) | 859 (321)* | 298 (113) | 286 (160) |
| Hospital admissions | 920 (2,223) | 791 (1,934) | 366 (422) | 845 (1,677) | 914 (960) | 386 (780) | 970 (1,078) | 1,888 (4,283) | 308 (405) |
| Procedures | 398 (370) | 487 (633) | 347 (396) | 697 (617) | 706 (566) | 439 (367) | 708 (607) | 662 (502) | 414 (307) |
| Concomitant medication | 733 (1,332) | 386 (238) | 209 (301) | 602 (331) | 516 (312) | 381 (211) | 880 (840) | 403 (349) | 415 (276) |
| Novel agents | 2,089 (1,262) | 592 (401)† | 2,049 (1,086) | 2,718 (1,515) | 700 (360)† | 3,423 (1,696) | 2,825 (1,342) | 672 (375)† | 3,651 (2,193) |
| Bortezomib | 2,019 (1,322) | 0 | 0 | 2,698 (1,457) | 0 | 0 | 2708 (1,356) | 0 | 0 |
| Thalidomide | 71 (269) | 592 (401) | 0 | 20 (117) | 700 (360) | 0 | 7 (24) | 672 (375) | 0 |
| Lenalidomide | 0 (0) | 0 | 2,049 (1,086) | 0 | 0 | 3,423 (1,696) | 110 (380)** | 0 | 3,651 (2,193) |
| Other therapy | 67 (130) | 64 (169) | 101 (171) | 83 (199) | 20 (25) | 86 (201) | 18 (52) | 23 (27) | 41 (79) |
| Stem cell transplantation | 0 (0) | 125 (394) | 1,015 (1,548)* | 227 (1,032) | 221 (655) | 0 | 0 | 362 (887) | 0 |
| **Total monthly costs** | 4,814 (3,759) | 2,684 (2,678)† | 4,215 (2,132) | 5,966 (2,677) | 3,393 (1,797)† | 5,029 (2,285) | 6,260 (2,311) | 4,308 (5,726)† | 5,114 (2,403) |

Due to violation of the normality assumption, the nonparametric Kruskal-Wallis test for two or more related samples was used to compare the three groups. For cost components that differed significantly as $\alpha$=0.05, pairwise tests by means of the Mann-Whitney test were performed to identify which groups differed significantly. *Cost component was significantly higher compared to all other groups in pair-wise comparisons (p-value<0.05). †Cost component was significantly lower compared to all other groups in pair-wise comparisons (p-value<0.05).**One patient received combination therapy of bortezomib and lenalidomide which was primarily bortezomib-based. SD: standard deviation.

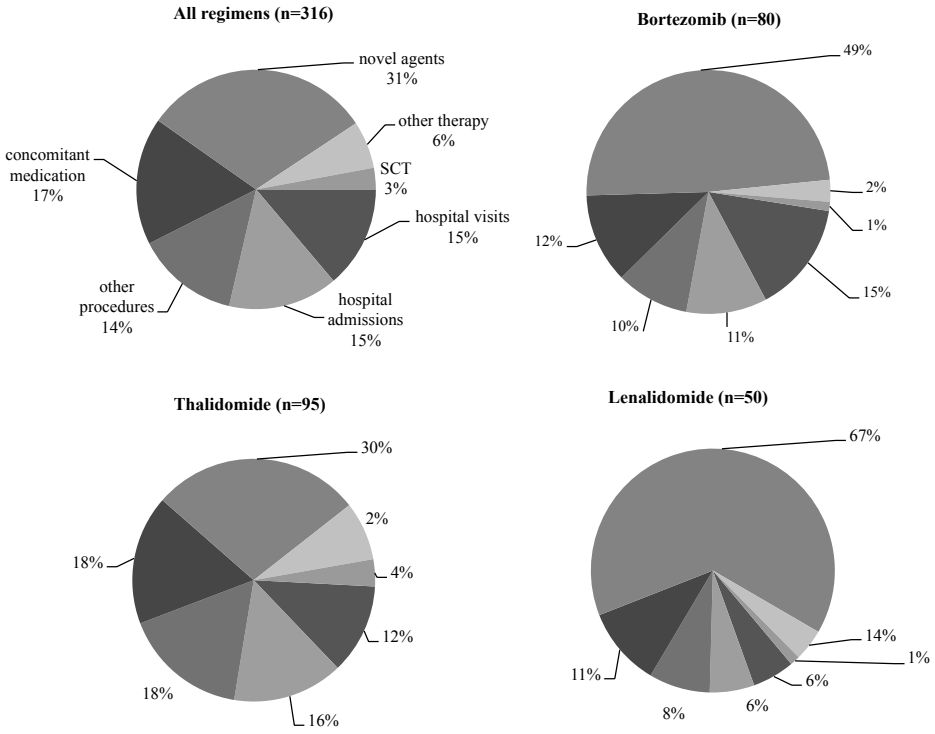



**Figure 1.** Proportion of mean total monthly costs during all regimens and separately during bortezomib-, thalidomide-, and lenalidomide-based regimens by main cost categories.

bortezomib-based and lenalidomide-based regimens. Higher mean total monthly costs during bortezomib-based regimens were observed for hospital visits. Costs during lenalidomide regimens were mainly attributable to significantly higher acquisition costs of lenalidomide.

Figure 1 further illustrates that the structure of total monthly costs also differed across novel agent regimens. Approximately 67% and 49% of mean monthly costs incurred during lenalidomide- and bortezomib-based treatment regimens were attributable to the acquisition costs of the drug, compared to 30% during thalidomide-based treatment and 32% average across all regimens. Costs attributable to hospital visits and hospital admissions were important cost drivers during bortezomib-based (15% and 11%, respectively) and thalidomide-based regimens (12% and 16%, respectively) and less influential during lenalidomide-based (6% and 6%, respectively) regimens.

## DISCUSSION

Costs during treatment of RRMM varied greatly and were largely attributable to hospital admissions and acquisition costs of novel agents. Costs varied also by treatment-related resource use. The variation is not surprising since healthcare costs are known to be skewed with costs generally increasing as prognosis worsens. Known prognostic factors such as poor performance status and low platelet count were indeed associated with increased costs. Other relationships may also exist but this could not be assessed with the data collected in our study. Given that MM as a heterogeneous disease with high-risk patients in particular continuing to have a poor outcome despite improvements in treatment and survival [20], further investigation should address whether the cost to health benefit ratio in these patients is more favorable for regimens containing existing novel agents versus those containing experimental therapies.

Total costs during bortezomib-based and lenalidomide-based regimens are significantly higher compared to thalidomide-based regimens. The costs during bortezomib-based regimens were high on account of hospital visits and acquisition costs of the drug, while lenalidomide-based regimens were costly due to the acquisition costs of the drug alone. Since bortezomib was only administered intravenously in this study, hospital visits were required. Some cost-savings could result with subcutaneous administration of bortezomib [21] or by reducing waste of bortezomib by using unused reconstituted bortezomib within two weeks [22] or adjustment of vial size [23]. Whether these options lead to cost-savings should be addressed in future studies.

Published existing economic evaluations have compared bortezomib to best supportive care or thalidomide [7] and lenalidomide [6, 8]. None of these studies based the cost analysis on real-world estimates. As for existing costs studies, the costs seen here are similar to the costs reported for Sweden [9] and France [10]. In Sweden, the costs also varied between individuals with inpatient stays (35%) and therapy-induced adverse events (42%) constituting a large proportion of mean total monthly costs [9]. Costs of chemotherapy in Sweden were similarly found to increase over treatment lines. Mean total monthly costs were, however, lower compared to our study (i.e., €2,770 versus €3,981), most likely due to less frequent use of novel agents. The more recent study from France reported higher costs compared to Sweden mainly due to greater use of novel agents (73%) [10]. Similar to our study, a large proportion of the total monthly costs reported in the French study were attributable to the costs of chemotherapy drugs (66%) and inpatient hospital stays (15.4%). Mean total monthly costs from France were, however, slightly lower compared to our study (i.e., €3,130 versus €3,981). This is

probably due to differences in methodology as our study implemented the method of micro-costing, which results in a more expansive yet accurate list of resource use [24].

In contrast to the healthcare costs in Italy [11], the mean total costs per patient (€1,171 per month or €14,053 per year) were considerably lower compared to our results (€3,981 per month or €47,772 per year). The authors did report a similar finding that the majority of the costs were attributable to therapy and hospital admissions. However, their cost estimates are not comparable for a number of reasons affecting transferability. First, concomitant medication costs were not included. Second, actual differences in costs may be present on account of differences in the patient group and treatment practice. Third, lower tariffs were applied. Finally, random variation could account for the differences since the Italian study was a single-centre study.

The transferability of real-world cost studies across countries can be assessed by the patterns observed for resource use and cost drivers over time. Hospital admissions were consistently found to account for a large portion of the total costs. In addition, therapy costs are increasingly becoming an important driver of increased costs. Similar to the French study [10], the high degree of variation in costs was mainly due to the period of data collection (2001-2009) which was characterized by rapid changes in therapy options. The EMA approved bortezomib and lenalidomide in 2004 and 2007, respectively. As a result, the patients included in our study represent a cohort treated during the early years of the availability of both novel and expensive agents. This explains why the use of bortezomib was more common in third line compared to second line as well as the generally low numbers of patients receiving lenalidomide-based regimens. As a consequence of the low numbers of patients receiving lenalidomide, comparisons between lenalidomide-based regimens and regimens with other novel agents should be judged carefully.

Variation in treatment and associated costs will remain a hallmark of RRMM. An abundance of novel agents will characterize the future treatment paradigm for MM [25], and with demonstration of improved outcomes based on drug synergy [26], patients will also be eligible for a combination of novel agents. The additional costs of combination therapy with expensive novel agents could be substantial and should be assessed in future real-world studies.

## ACKNOWLEDGEMENTS

The authors would like to thank all participating hospitals, data-managers, cancer registration (IK) staff and Danielle Ammerlaan, Hedwig Blommestein, Aat van Klooster, Annemieke Leunis, Arewika Mamakanjan, Zjwan Rashaan, Bruno Vieyra, and Anneloes van Walsem for their contribution to this study.

## REFERENCES

1. Alexander DD, Mink PJ, Adami HO, et al. Multiple myeloma: a review of the epidemiologic literature. Int J Cancer 2007; 120 Suppl 12: 40-61.
2. Harousseau JL, Dreyling M. Multiple myeloma: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2009; 20 Suppl 4: 97-9.
3. Harousseau JL, Moreau P. Autologous hematopoietic stem-cell transplantation for multiple myeloma. N Engl J Med 2009; 360(25): 2645-54.
4. Richardson PG, Hideshima T, Mitsiades C, Anderson KC. The emerging role of novel therapies for the treatment of relapsed myeloma. J Natl Compr Canc Netw 2007; 5(2): 149-62.
5. Gaultney JG, Redekop KW, Sonneveld P, Uyl-de Groot CA. Critical review of economic evaluations in multiple myeloma: An overview of the economic evidence and the trends and quality in the methodology. European Journal of Cancer Forthcoming.
6. Hornberger J, Rickert J, Dhawan R, et al. The cost-effectiveness of bortezomib in relapsed/refractory multiple myeloma: Swedish perspective. Eur J Haematol 2011; 85(6): 484-91.
7. Mehta J, Duff SB, Gupta S. Cost effectiveness of bortezomib in the treatment of advanced multiple myeloma. Manag Care Interface 2004; 17(9): 52-61.
8. Moller J, Nicklasson L, Murthy A. Cost-effectiveness of novel relapsed-refractory multiple myeloma therapies in Norway: lenalidomide plus dexamethasone vs bortezomib. J Med Econ 2011; 14(6).
9. Ghatnekar O, Alvegard 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. Clin Ther 2008; 30(9): 1704-13.
10. 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. J Clin Pharm Ther 2011; 36(1): 19-26.
11. Koleva D, Cortelazzo S, Toldo C, Garattini L. Healthcare costs of multiple myeloma: an Italian study. Eur J Cancer Care (Engl) 2011; 20(3): 330-6.
12. Lokhorst HM, van der Holt B, Zweegman S, et al. A randomized phase 3 study on the effect of thalidomide combined with adriamycin, dexamethasone, and high-dose melphalan, followed by thalidomide maintenance in patients with multiple myeloma. Blood 2010; 115(6): 1113-20.
13. Centraal Bureau voor de Statistiek. Statline database for all statistics relevant to the Netherlands [cited 2010 July 1]; Available from: http://statline.cbs.nl/
14. Tan SS, Van Gils CW, Franken MG, et al. The unit costs of inpatient hospital days, outpatient visits, and daycare treatments in the fields of oncology and hematology. Value Health 2010; 13(6): 712-9.
15. Nederlandse Zorgautoriteit. Search program for DBC amounts and other amounts due to spe-

cialist medical care visits for health care., [cited 2010 July 1]; Available from: http://ctg.bit-ic.nl/Nzatarieven/top.do
16. Dutch Health Care Insurance Board. Farmacotherapeutisch Kompas., [cited 2010 July 1]; Available from: http://www.fk.cvz.nl/
17. van Agthoven M, Vellenga E, Fibbe WE, et al. Cost analysis and quality of life assessment comparing patients undergoing autologous peripheral blood stem cell transplantation or autologous bone marrow transplantation for refractory or relapsed non-Hodgkin's lymphoma or Hodgkin's disease. a prospective randomised trial. Eur J Cancer 2001; 37(14): 1781-9.
18. Blommestein HM, Verelst SGR, Cornelissen JJ, et al. Cost of Autologous and Allogeneic Stem Cell Transplantations for Hematological Diseases [Kostprijzen van autologe en allogene hematopoietische stamceltransplantaties voor hematologische ziekten]: institute of Health Policy and Management; 2010.
19. Briggs A, Claxton K, Schulpher M. Decision modelling for health economic evaluation. New York: Oxford University Press; 2006.
20. Richardson PG, Laubach J, Mitsiades CS, et al. Managing multiple myeloma: the emerging role of novel therapies and adapting combination treatment for higher risk settings. Br J Haematol 2011.
21. Moreau P, Pylypenko H, Grosicki S, et al. Subcutaneous versus intravenous administration of bortezomib in patients with relapsed multiple myeloma: a randomised, phase 3, non-inferiority study. Lancet Oncol 2011; 12(5): 431-40.
22. Vanderloo JP, Pomplun ML, Vermeulen LC, Kolesar JM. Stability of unused reconstituted bortezomib in original manufacturer vials. J Oncol Pharm Pract 2010; 17(4): 400-2.
23. Clark L, Castro AP, Fortes AF, et al. Ideal vial size for bortezomib: real-world data on waste and cost reduction in treatment of multiple myeloma in Brazil. Value Health 2011; 14(5 Suppl 1): S82-4.
24. Swindle R, Lukas CV, Meyer DA, et al. Cost analysis in the Department of Veterans Affairs: consensus and future directions. Med Care 1999; 37(4 Suppl Va): AS3-8.
25. Mahindra A, Cirstea D, Raje N. Novel therapeutic targets for multiple myeloma. Future Oncol 2010; 6(3): 407-18.
26. Jagannath S, Dimopoulos MA, Lonial S. Combined proteasome and histone deacetylase inhibition: A promising synergy for patients with relapsed/refractory multiple myeloma. Leuk Res 2010; 34(9): 1111-8.

## SUPPLEMENTAL DATA

**Table S1.** Unit costs for hospital resource use (Euro 2009)

| Resource Use | Unit Price | Unit |
|---|---|---|
| Hematological outpatient visit * | € 110 | day |
| Hematological daycare visit* | € 167 | day |
| Emergency room visit | € 109 | day |
| Intensive care unit | € 2,080 | day |
| Hematological inpatient ward* | € 516 | day |
| Consultation by telephone | € 13 | day |
| Radiotherapy standard | € 1,656 | per session |
| Radiotherapy intensive | € 7,971 | per session |
| X-ray | € 52 | per session |
| Skeletal scan | € 177 | per session |
| CT scan | € 208 | per session |
| MRI | € 270 | per session |
| Radionucleide scan | € 350 | per session |
| PET scan (total body) | € 1,411 | per session |
| Laboratory (per test) | € 53 | per test |
| Cytology testing | € 47 | per test |
| Echo (ultrasound) | € 86 | per test |
| Bacterial culture | € 14 | per test |
| Viral culture | € 27 | per test |
| Concomitant medication | | |
| **Acute** | | |
| Antibacterial / antifungal / antiviral | € 211.50 | per 10 days |
| Anemia (red blood cell transfusion) | € 204.00 | per unit |
| Thrombocytopenia (platelet infusion) | € 492.80 | per unit |
| **Chronic/Prophylactic** | | |
| Antibacterial / antifungal / antiviral | € 5.56 | day |
| Biphosponates | € 4.45 | day |
| Anemia (erythropoietin injection)** | € 3.21 | day |
| Neurotoxicity | € 9.57 | day |
| Gastro-intestinal | € 2.17 | day |
| Analgesics | € 1.70 | day |

*Weighting factor of 67:33 for general and university hospitals;

**Unit cost of injection given every three weeks was converted to a daily price.

**Table S2.** Unit costs of therapy regimens for treatment of multiple myeloma (Euro 2009)

| Therapy | Unit | Price |
|---|---|---|
| **Novel agents** | | |
| Bortezomib (iv) | 3.5 mg | € 954.52 |
| Thalidomide (oral) | 1 mg | € 0.24 |
| Lenalidomide (oral) | 5 mg | € 210.27 |
| | 10 mg | € 222.26 |
| | 15 mg | € 233.69 |
| | 25 mg | € 256.66 |
| **Other therapy** | | |
| Dexamethasone (oral) | 1 mg | € 0.06 |
| Dexamethasone (iv) | 5 mg | € 2.66 |
| | 20 mg | € 7.81 |
| Doxorubicin (iv) | 1 mg | € 1.48 |
| Vincristine (iv) | 1 mg | € 8.82 |
| Melphalan (iv) | 50 mg | € 52.25 |
| Melphalan (oral) | 2 mg | € 0.47 |
| Prednisone (iv) | 25 mg | € 2.42 |
| Prednisone (oral) | 5 mg | € 0.05 |
| Cyclophosphamide (oral) | 50 mg | € 0.14 |
| Cyclophosphamide (iv) | 200 mg | € 11.21 |
| | 500 mg | € 18.15 |
| | 1000 mg | € 34.38 |
| Interferon alpha (subcutaneous) | 1.0 x 10^6 IE | € 8.89 |
| Fludarabine (iv) | 50 mg | € 266.94 |
| Lomustine (oral) | 40 mg | € 23.39 |
| Cytarabine (iv) | 1600 mg | € 72.31 |
| Etoposide (iv) | 200 mg | € 40.25 |
| Carboplatin (iv) | 150 mg | € 100.65 |
| Donor lymphocyte infusion | related | € 2,342.03 |
| | unrelated | € 7,876.64 |
| **Pre-transplantation and transplantation costs (excluding hospital visits and admissions)** | | |
| Autologous PBSCT | | € 11,149.00 |
| Allogeneic PBSCT | | € 29,652.00 |
| Allogeneic MUD | | € 65,209.00 |

iv: intravenous; MUD: matched unrelated donor; PBSCT: peripheral blood stem cell transplantation.

**Table S3.** Additional descriptive statistics for mean monthly costs during relapsed/refractory treatment

| Cost category | 2nd line | | | | | 3rd line | | | | | 4th line | | | | | All lines | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Range | | Gamma parameters | | | Range | | Gamma parameters | | | Range | | Gamma parameters | | | Range | | Gamma parameters | |
| | Median | Min | Max | Alpha* | Beta** | Median | Min | Max | Alpha* | Beta** | Median | Min | Max | Alpha* | Beta** | Median | Min | Max | Alpha* | Beta** |
| **Hospital visits** | 254 | 0 | 2,241 | 183.7 | 1.8 | 456 | 32 | 1,461 | 238.8 | 2.2 | 365 | 101 | 1,638 | 97.4 | 4.8 | 318 | 0 | 2,241 | 276.7 | 1.3 |
| **Hospital admissions** | 0 | 0 | 18,756 | 19.5 | 51.5 | 266 | 0 | 8,988 | 30.9 | 34.5 | 432 | 0 | 10,611 | 16.0 | 58.8 | 381 | 0 | 16,153 | 39.0 | 30.8 |
| **Radiotherapy** | 0 | 0 | 4,969 | 9.9 | 12.5 | 0 | 0 | 1,202 | 18.5 | 5.8 | 0 | 0 | 920 | 7.8 | 9.6 | 0 | 0 | 1,080 | 44.9 | 1.8 |
| **Surgery** | 0 | 0 | 2,723 | 5.1 | 9.9 | 0 | 0 | 1,793 | 5.5 | 10.9 | 0 | 0 | 1,854 | 1.0 | 38.4 | 0 | 0 | 1,965 | 11.1 | 5.7 |
| **Diagnostics** | 197 | 25 | 7,938 | 34.9 | 11.5 | 341 | 0 | 1,863 | 118.9 | 4.1 | 282 | 43 | 1,818 | 50.9 | 8.5 | 248 | 25 | 7,938 | 45.3 | 9.5 |
| **Concomitant medication** | 421 | 0 | 6,852 | 58.7 | 8.7 | 551 | 0 | 3,179 | 160.2 | 3.6 | 515 | 0 | 3,383 | 73.5 | 8.0 | 463 | 0 | 6,129 | 94.4 | 5.9 |
| **Therapy** | 550 | 0 | 5,121 | 98.7 | 8.3 | 1,362 | 0 | 7,300 | 88.7 | 20.1 | 1,686 | 6 | 9,930 | 50.8 | 41.1 | 805 | 0 | 5,313 | 177.4 | 6.2 |
| Bortezomib | 0 | 0 | 5,091 | 20.1 | 18.0 | 0 | 0 | 6,608 | 38.7 | 27.1 | 0 | 0 | 4,405 | 11.7 | 51.4 | 129 | 0 | 4,492 | 55.2 | 8.4 |
| Thalidomide | 122 | 0 | 2,077 | 78.6 | 4.0 | 0 | 0 | 1,480 | 14.1 | 8.3 | 0 | 0 | 1,290 | 5.3 | 14.3 | 114 | 0 | 1,440 | 80.9 | 2.7 |
| Lenalidomide | 0 | 0 | 3,593 | 3.4 | 17.5 | 0 | 0 | 6,738 | 12.9 | 41.2 | 0 | 0 | 9,919 | 21.1 | 65.3 | 0 | 0 | 3,932 | 31.5 | 11.5 |
| Non-novel agents | 10 | 0 | 2,007 | 17.0 | 5.1 | 9 | 0 | 1,319 | 14.2 | 5.8 | 11 | 0 | 281 | 20.7 | 1.9 | 14 | 0 | 757 | 38.4 | 1.5 |
| **Stem cell transplantation** | 0 | 0 | 6,962 | 9.1 | 23.2 | 0 | 0 | 5,772 | 4.7 | 40.9 | 0 | 0 | 2,172 | 1.9 | 34.1 | 0 | 0 | 3,268 | 20.8 | 8.3 |
| **Total monthly costs** | 2,112 | 353 | 31,318 | 86 | 40 | 4,412 | 421 | 13,437 | 232 | 21 | 3,866 | 490 | 15,881 | 130 | 36 | 3,226 | 442 | 31,318 | 176 | 22.6 |

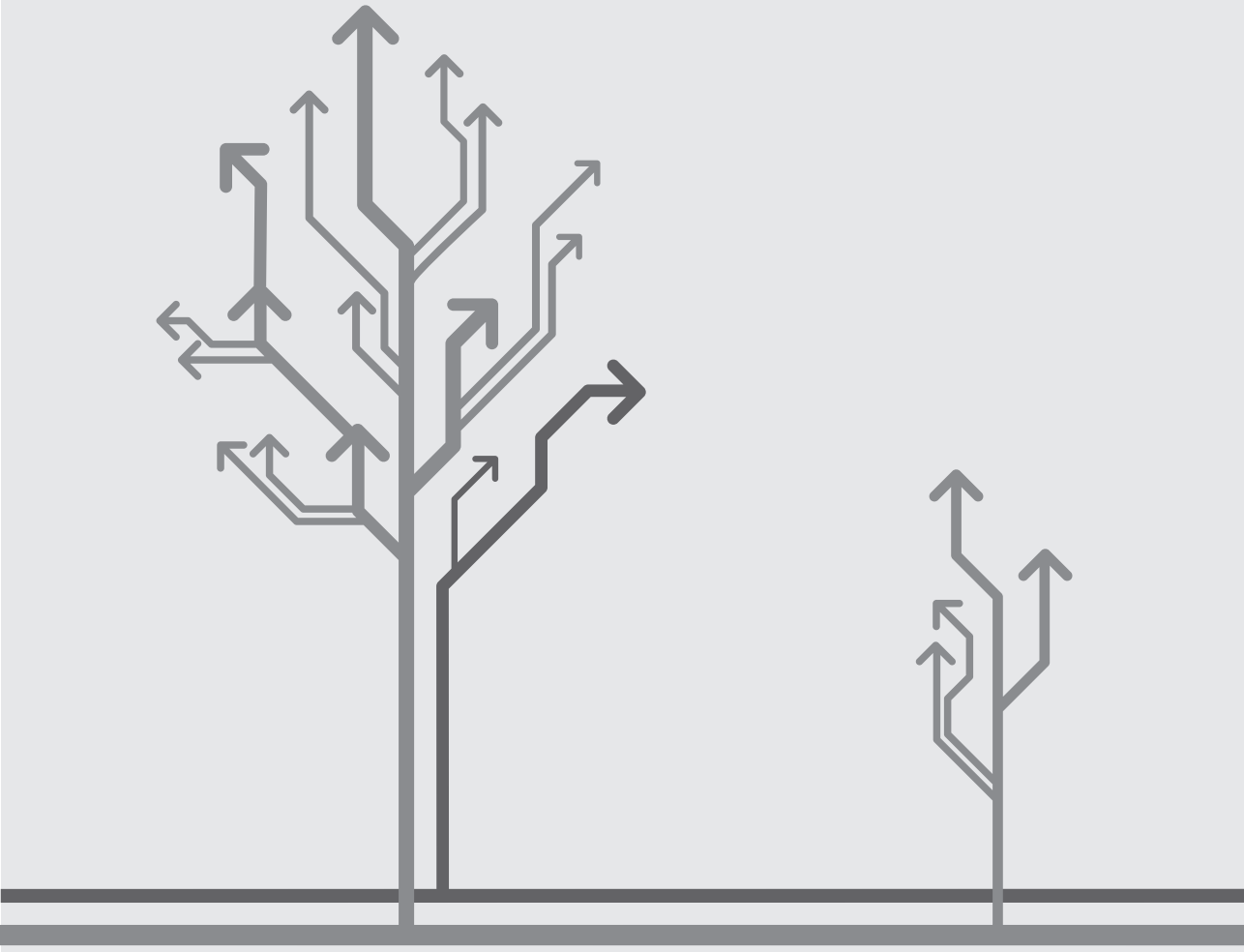
*Alpha = (mean/standard error)^2. **Beta = (standard error^2)/mean

# Chapter 6.

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

Margreet G. Franken, Jennifer G. Gaultney, Hedwig M. Blommestein,
Peter C. Huijgens, Pieter Sonneveld, William K. Redekop, Carin A. Uyl-de Groot

*Revision submitted to Value in Health.*



## 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 of bortezomib reduced the initial policymaker uncertainty regarding real-world cost-effectiveness. Our cohort study included 139 relapsed/ refractory multiple myeloma patients 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] versus 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. However, patients treated with bortezomib 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 the use of bortezomib, its effects and costs in everyday practice. However, much uncertainty remained regarding its real-world incremental 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.

## INTRODUCTION

Rising health care expenditures are making it extremely difficult to manage early access to promising innovative 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 new, often expensive, drugs. These policies address clinical and/or finance uncertainty, for example by means of finance-based [1, 2] or outcome-based [2, 3] risk sharing agreements such as coverage with evidence development schemes [4-6] or outcomes research requirements [7, 8].

Although outcomes research and evidence development requirements increasingly seem an attractive policy option, many unanswered questions remain regarding their actual value [2, 6, 9] and feasibility [10-12]. In The Netherlands, outcomes research requirements were first implemented in 2006 for expensive inpatient drugs. From 2013 onwards, this policy will be 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 [13], namely to gather data in everyday clinical 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 [14]. 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 [15], which found bortezomib to be superior to high dose dexamethasone in terms of increased time to progression (6.22 versus 3.49 months), response rates (38% versus 18%), response duration (8 versus 5.6 months), and one-year survival rate (80% versus 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 was available [16]. 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 clinical 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 paper describes our experiences in The Netherlands in performing outcomes research of bortezomib in relapsed/ refractory multiple myeloma. We investigated whether outcomes research of bortezomib 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 newly diagnosed Durie-Salmon stage II/ III multiple myeloma patients aged 18 – 65 years [17]. Patients who went off-protocol from this trial regime no longer received protocol-based therapy and were therefore eligible for our outcomes research study as 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 versus 0.18) in the APEX trial [15]. 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 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.

### Drug use and real-world cost-effectiveness

To assess drug use in everyday practice, we examined baseline patient characteristics, types of treatments received, dosages and dose modifications. To estimate overall sur-

vival (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 fol-

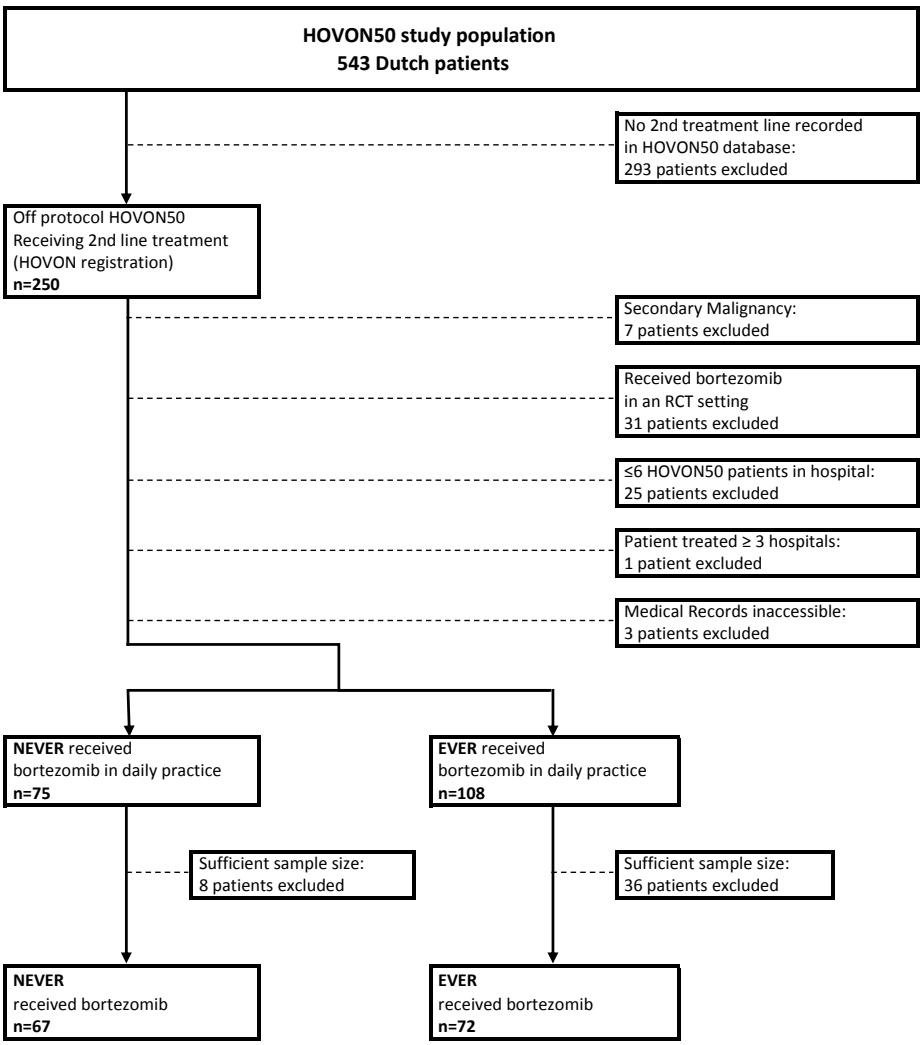



**Figure 1.** Flowchart of the patient selection process.

lowing 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 [18].

For cost-effectiveness of bortezomib in everyday practice, we investigated the feasibility to obtain comparable patient groups, identify treatment comparators, and estimate (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 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 received by patients. Table 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 2 reveals, a combination of treatments was common practice; more than ten drugs were given in more than twenty different combinations. The most frequent combinations were thalidomide/ dexamethasone (n=57), lenalidomide/ dexamethasone

**Table 1.** Baseline characteristics at start of relapsed/ refractory treatment

| Characteristic | Received bortezomib (n=72) | | Never received bortezomib (n=67) | | P-value † |
|---|---|---|---|---|---|
| *Patient-related characteristics* | | % Missing | | % Missing | |
| 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 |
| Months until first progression [mean (range)] | 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

**Table 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 | |

(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 comparable cycle regimes similar to the pivotal registration trial (i.e. APEX trial [15, 19]). However, patients in everyday practice received fewer treatment cycles (4 versus 6) as well as lower dosages (13%).

It was not feasible to establish a pattern for bortezomib 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%).

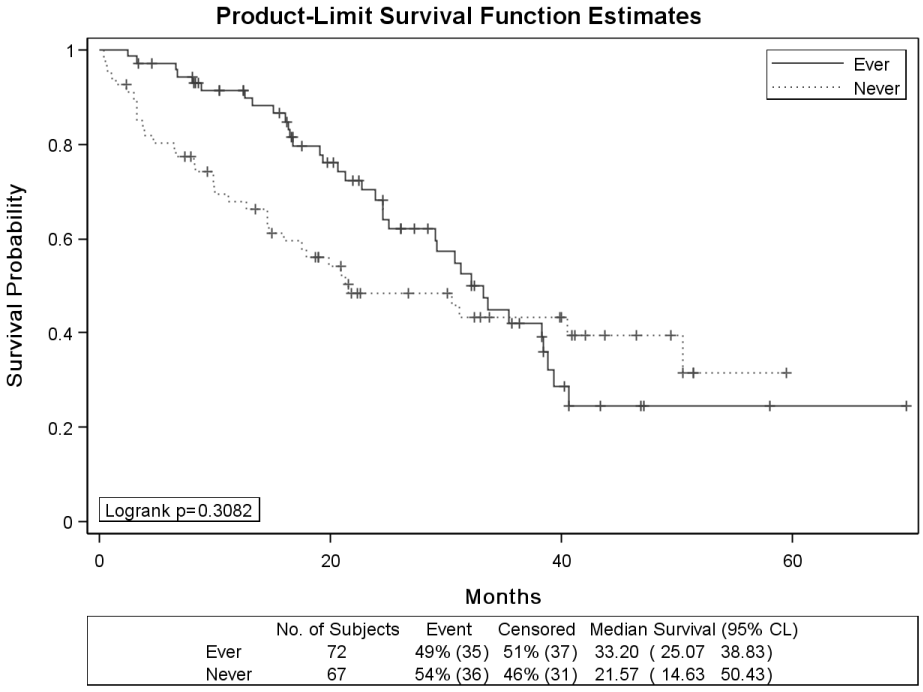



**Figure 2.** Kaplan-Meier OS curves from start of relapsed/refractory treatment stratified by receipt of bortezomib.

### Treatment effects

Policymakers generally prefer OS and quality adjusted life years (QALYs) as outcome measures in reimbursement decision making [20]. 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 bortezomib and 31 non-bortezomib patients were still alive. Kaplan-Meier curves (see Figure 2) from start of relapsed/ refractory treatment showed a longer mean (29.5 versus 28.0 months) and median (33.2 versus 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).

Previous research found that receiving thalidomide as first line treatment is associated with reduced overall survival after relapsed/ refractory treatment [21, 22]. Therefore, we stratified the Kaplan-Meier curves [not shown] by HOVON50 treatment arm (Thalidomide, Adriamycin, and Dexamethasone (TAD) arm versus 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: bortezomib patients VAD arm 30.9 [33.6] and TAD arm 27.0 [29.2] months; and non-bortezomib patients VAD arm 29.9 [31.1] and 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 for receiving bortezomib in second, third, or fourth line or later, respectively). However, as Figure 3 shows, further stratifying all 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).

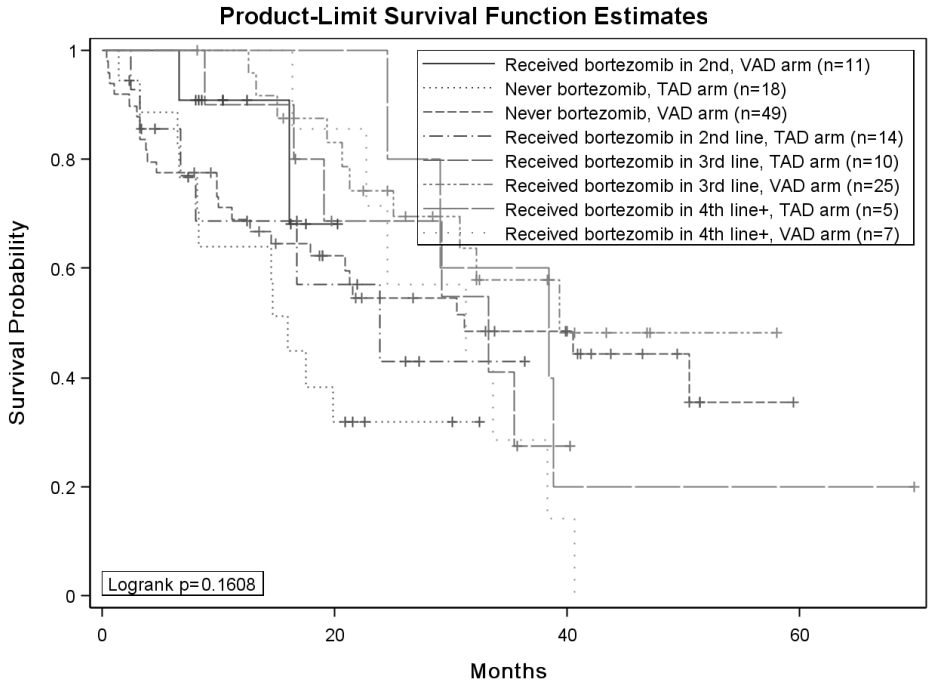



**Figure 3.** Kaplan-Meier OS curves from start of relapsed/refractory treatment stratified by receipt of bortezomib, bortezomib treatment line and Hovon50 treatment arm.

**Table 3.** Total mean costs from start of relapsed/ refractory treatment stratified by receipt of bortezomib

| | Ever received bortezomib | | Never received bortezomib | |
|---|---|---|---|---|
| | (n=72) | | (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

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 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 differences were most likely due to the fact that the majority of 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 substantial 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 influence on the total mean costs. The greatest influence was obtained by varying the unit price for inpatient hospital days (range bortezomib patients: €75,176 to €88,076; range non-bortezomib patients: €45,527 to €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. However, without the intention to make direct comparisons, 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: €95962; 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 in terms of real-world effectiveness, health care costs and cost-effectiveness at the time of the initial reimbursement decision. Although outcomes research and evidence development requirements globally seem to be popular as well as promising policy options to reduce decision maker uncertainty [4, 7, 23], 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 [15]. No data were available on long-term survival and healthcare 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 to high dose dexamethasone. In contrast, outcomes research showed that treatment in clinical practice was far more heterogeneous. Although real-world patients received fewer bortezomib treatment cycles (4 versus 6) as well as lower dosages (13%) compared to trial patients, time to progression (6.8 versus 6.22 months) and response rates (complete

response: 8% versus 6%; very good, partial, and minimal response: 55% versus 41%) seemed reasonably similar. However, one-year survival rate was lower in everyday clinical practice (66% versus 80%). (A detailed comparison between our real-world patients and trial patients is reported elsewhere [24]). Furthermore, outcomes research showed detailed healthcare costs beyond the price of bortezomib itself. Thus, outcomes research of bortezomib provided valuable information to policymakers on types of treatments received, which patients received or did not receive bortezomib, bortezomib dosages, dose modifications, (overall) survival, treatment costs, and costs per month of survival of relapsed/ refractory multiple myeloma patients in everyday clinical practice. Hence, outcomes research reduced initial policymaker uncertainty about bortezomib use, effects and health care costs in everyday practice.

However, outcomes research did not reduce the uncertainty of the incremental cost-effectiveness of bortezomib compared to other treatments. Because of extensive treatment variation, it was not possible to identify appropriate treatment comparators. Furthermore, as expected, our results confirm previous concerns [25-28] that great heterogeneity and a lack of randomisation in everyday practice resulted in incomparable patient groups. Although other observational studies successfully employed the propensity score matching technique [29-33], essential prerequisites [34], 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.

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 [35-37] and economic evaluations [38-40] were published in relapsed/ refractory multiple myeloma. However, previous cost-studies were based on conventional therapies [35], did not apply micro-costing techniques [36], nor provided information on the use of novel agents [37]. Previous economic evaluations were based on synthesising data and expert opinions and did not use patient level data [38-40]. Therefore, our outcomes research 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. However, our findings 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. For example, recent Dutch experiences also revealed insufficient data to perform a reassessment after four years of data collection (i.e., omalizumab, infliximab, and ranibizumab). A limitation of our study was that we used the HOVON50 population to select patients who received bortezomib outside of a clinical trial. Many Dutch multiple myeloma 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 in outcomes research. 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 cancer patients [41]. 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. However, if medical records are to be used for data collection, it is important to emphasise that there is a high need to improve reporting of clinical data.

Because much uncertainty remained regarding the incremental (cost-) effectiveness, the question arises if outcomes research was the best option to reduce policymaker uncertainty regarding value for money of bortezomib in relapsed/ refractory multiple myeloma. In England and Wales, for example, after first receiving a negative advice from NICE, the manufacturer and the UK department of health agreed on a performance based response-rebate scheme for bortezomib [42]. Although the Dutch minister announced the implementation of risk-sharing arrangements from 2013 onwards [43], coverage with evidence development by outcomes research has been the only legal policy option in The Netherlands.

Several taxonomies exist that classify different risk-sharing arrangements [1, 9, 23, 44, 45], 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 [46]. Also other studies [2, 3, 23, 47, 48] 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 UK 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. Moreover, recent Dutch experiences revealed insufficient data to perform a reassessment after four years of data collection (i.e., omalizumab, infliximab, and ranibizumab).

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). However, it also remains debatable if an outcomes based risk-sharing agreement, such as in the UK, would have decreased policymaker uncertainty regarding the value for money of bortezomib. 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, 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 (VOI)- to enhance outcomes of investments of valuable resources. Policymakers could also consider requesting Bayesian updating of the existing model, 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 which 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, 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 versus chronic) and characteristics of the drug.

To conclude, outcomes research of bortezomib provided valuable information to policymakers on real-world patients, types of treatments, dosages, dose modifications and

health care costs. However, assessing (incremental) effectiveness and cost-effectiveness 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 bortezomib in relapsed/refractory multiple myeloma in everyday practice, much uncertainty remained regarding its incremental cost-effectiveness. At reimbursement decision making, policymakers should carefully consider what type of evidence could lead to an acceptable reduction in uncertainty regarding the value for money of each specific drug in its re-evaluation. Instead of implementing outcomes research requirements in general, policy makers 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.



## REFERENCES

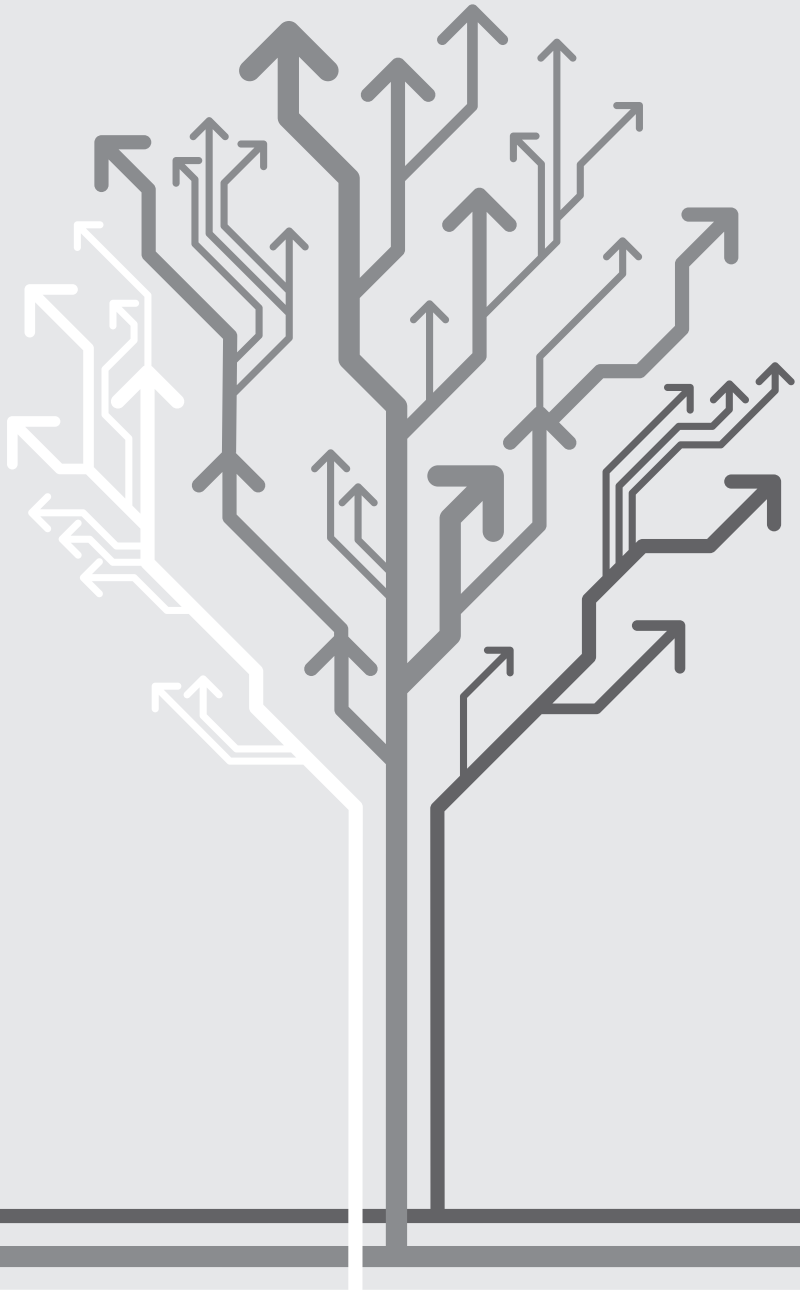
1. Carlson JJ, Sullivan SD, Garrison LP, et al. Linking payment to health outcomes: a taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. Health Policy. 2010;96 (3):179-90.
2. Adamski J, Godman B, Ofierska-Sujkowska G, et al. Risk sharing arrangements for pharmaceuticals: potential considerations and recommendations for European payers. BMC health services research. 2010;10 (1):153.
3. Carlson JJ, Garrison LP, Sullivan SD. Paying for outcomes: innovative coverage and reimbursement schemes for pharmaceuticals. J Manag Care Pharm. 2009 Oct;15 (8):683-7.
4. Tunis SR, Pearson SD. Coverage options for promising technologies: Medicare's' coverage with evidence development'. Health Aff. 2006;25(5):1218-30.
5. Mohr PE, Tunis SR. Access with evidence development: the US experience. Pharmacoeconomics. 2010;28(2):153-62.
6. Trueman P, Grainger DL, Downs KE. Coverage with Evidence Development: Applications and issues. Int J Technol Assess Health Care. 2010;26(01):79-85.
7. Krumholz HM. Outcomes research: generating evidence for best practice and policies. Circulation. 2008;118(3):309.
8. Clancy CM, Eisenberg JM. Outcomes research: measuring the end results of health care. Science.

1998;282(5387):245.
9. Towse A, Garrison LP. Can't 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.
10. 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. Eur J Cancer. 2013;49(1):8-16.
11. Chalkidou K, Tunis S, Lopert R, et al. Comparative Effectiveness Research and Evidence Based Health Policy: Experience from Four Countries. Milbank Q. 2009;87(2):339-67.
12. Rubin DB. On the limitations of comparative effectiveness research. Stat Med. 2010;29(19):1991-5.
13. Delwel GO. Guidance for Outcomes Research. Diemen: College voor Zorgverzekeringen. 2008. Report No.: 270.
14. Kuijpers MR, Toenders WGM. Assessment procedure for inpatient drugs [in Dutch: Procedure beoordeling intramurale geneesmiddelen]. Diemen: College voor Zorgverzekeringen. 2006. Report No.: 26022597.
15. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med. 2005;Jun(16):2487-98.
16. Commissie Farmaceutische Hulp. Farmacotherapeutisch rapport bortezomib (Velcade®). Diemen: College voor Zorgverzekeringen. 2005. Report No.: 25060777.
17. 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):124-7.
18. Gaultney JG, Franken MG, Tan SS, et al. Real-world healthcare costs of relapsed/refractory multiple myeloma during the era of novel cancer agents. J Clin Pharm Ther. 2013;38(1):41-7.
19. 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.
20. Pharmacoeconomic Guidelines Around The World [Online]. U.S.A.: International Society for Pharmacoeconomics and Outcomes Research. 2013; cited May 21, 2013]. Available from: http://www.ispor.org/peguidelines/index.asp.
21. 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. Br J Haematol. 2009;147(4):531-4.
22. 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):1113-20.
23. Garrsion L, Towse A, Brigss A, et al. Performance-based risk-sharing arrangements: Good practices for design, implementation and evaluation, An ISPOR Task Force Report. ISPOR. 2012. [Final draft for review].
24. Gaultney JG, Franken MG, Redekop WK, et al. PRM3: Turning the tables to address the real value of real-world observational studies of novel anti-cancer agents in multiple myeloma. Value in Health [Abstract]. 2012;15(7):A460.
25. Cox E, Martin BC, Van Staa T, et al. 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 Task Force

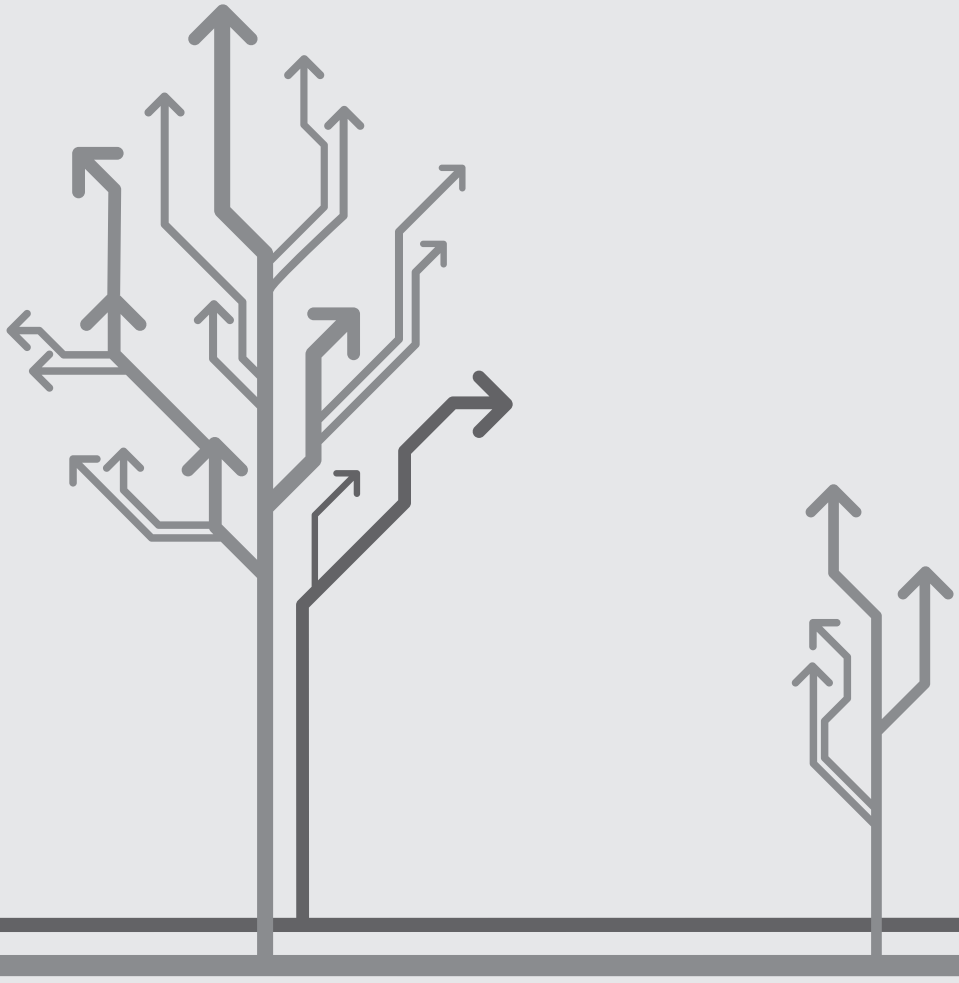
Report—Part II. Value in Health. 2009;12(8):1053-61.
26. Concato J, Lawler EV, Lew RA, et al. Observational methods in comparative effectiveness research. Am J Med. 2010;123(12S):e16-23.
27. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic 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):1062-73.
28. Psaty BM, Siscovick DS. Minimizing Bias Due to Confounding by Indication in Comparative Effectiveness Research. JAMA. 2010;304(8):897-8.
29. Connors A, Speroff T, Dawson NV, et al. The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients. JAMA. 1996;276(11):889-97.
30. Earle CC, Tsai JS, Gelber RD, et al. Effectiveness of chemotherapy for advanced lung cancer in the elderly: instrumental variable and propensity analysis. Journal of Clinical Oncology. 2001;19(4):1064-70.
31. 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.
32. Seeger JD, Walker AM, Williams PL, et al. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. Am J Cardiol. 2003; 92(12):1447-51.
33. Weintraub WS, Grau-Sepulveda MV, Weiss JM, et al. Comparative effectiveness of revascularization strategies. N Engl J Med. 2012;366(16):1467-76.
34. Weitzen S, Lapane KL, Toledano AY, et al. Principles for modeling propensity scores in medical research: a systematic literature review. Pharmacoepidemiol Drug Saf. 2004;13(12):841-53.
35. 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. Clin Ther. 2008;30(9):1704-13.
36. 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. J Clin Pharm Ther. 2011;36(1):19-26.
37. Koleva D, Cortelazzo S, Toldo C, et al. Healthcare costs of multiple myeloma: an Italian study. European Journal of Cancer Care. 2011;20(3):330-6.
38. Mehta J, Duff SB, Gupta S. Cost effectiveness of bortezomib in the treatment of advanced multiple myeloma. Manag Care Interface. 2004;17(9):52-61.
39. Hornberger J, Rickert J, Dhawan R, et al. The cost-effectiveness of bortezomib in relapsed/refractory multiple myeloma: Swedish perspective. Eur J Haematol. 2010; 85(6):484-91.
40. 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):690-7.
41. de Groot S, Blommestein H, Franken M, et al. PCN154 Are population-based registries a suitable tool for outcomes research in cancer? Experiences from four registries. Value in Health [Abstract]. 2012;15(7):A437-8.
42. National Institute for Health and Clinical Excellence. Bortezomib monotherapy for relapsed multiple myeloma. London: National Institute for Health and Clinical Excellence. 2007. Report No.: NICE Technology Appraisal Guidance 129.
43. Schippers EI. Proposals for provisional and conditional inclusion of new drugs in the drug re-

imbursement scheme [in Dutch: Voorstellen voor voorlopige en voorwaardelijke opname van nieuwe geneesmiddelen in het pakket]. Ministerial Letter. 2011; GMT-U-3065578.

44. McCabe CJ, Stafinski T, Edlin R, et al. Access with evidence development schemes: a framework for description and evaluation. Pharmacoeconomics. 2010;28(2):143-52.
45. 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):113-42.
46. Williamson S. A Report into the Uptake of Patient Access Schemes in the NHS. Northumbria Healthcare NHS Trust. North Shields, United Kingdom: Cancer Network Pharmacist Forum. 2009.
47. Neumann PJ, Chambers JD, Simon F, et al. Risk-Sharing Arrangements That Link Payment For Drugs To Health Outcomes Are Proving Hard To Implement. Health Aff. 2011;30(12):2329-37.
48. Sudlow CLM, Counsell CE. Problems with UK government's risk sharing scheme for assessing drugs for multiple sclerosis. BMJ. 2003;326(7385):388-92.



# Part 3 – Economic evaluations of risk-stratified treatment approaches

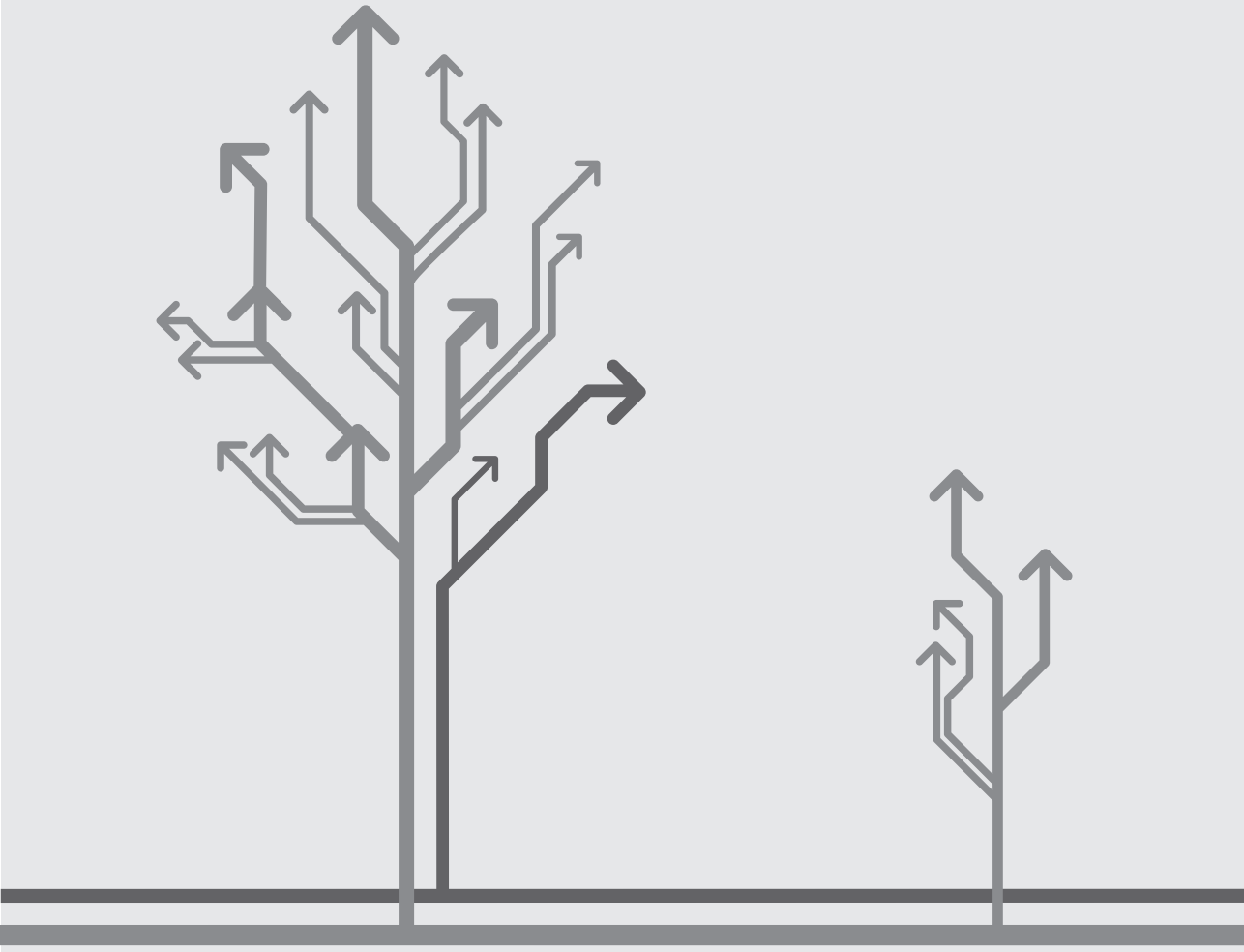


# Chapter 7.

## A modeling study to assess the potential therapeutic and economic value of risk-stratified treatment in multiple myeloma using molecular markers of adverse prognosis

Jennifer G. Gaultney, William K. Redekop, Pieter Sonneveld, Bronno van der Holt, Carin A. Uyl-de Groot

*Manuscript preparation in progress*



## ABSTRACT

Biomarkers associated with prognosis in multiple myeloma (MM) distinguish adverse prognosis, such as patient factors, markers of tumor burden, and molecular markers of tumor biology. Challenging the traditional uniform treatment (UT) approach, a risk-stratified treatment (RST) is proposed where high-risk patients receive bortezomib-based regimens while standard-risk patients receive alternative regimens. The objective was to evaluate the potential therapeutic and economic value of RST compared to UT in newly diagnosed transplant-eligible MM patients in Dutch daily practice. A Markov-type decision analytic model compared total health benefits and costs for two strategies: 1) UT where all patients received the standard of care consisting of bortezomib induction/maintenance and 2) RST where treatment was stratified according to clinical and tumor biomarkers only, molecular biomarkers only, or any biomarker. In RST, high-risk patients received bortezomib while other patients received chemotherapy and thalidomide. Input data originated from clinical trials, literature reviews, observational studies and national tariffs. Various sensitivity and scenario analyses were performed. RST dominated UT, with average health gains of 0.007-0.059 LYs (0.009-0.040 QALYs) and cost-savings of €1,842-€4,924 depending on biomarker set. A scenario analysis for RST where all high-risk patients received an experimental treatment increased health by 0.40 LYs (0.30 QALYs) and costs by €2,567 compared to UT. Influential parameters included the price of bortezomib and survival and quality-of-life-related parameters. An economic evaluation of biomarkers in the preclinical development phase provided evidence that RST in MM may improve health outcomes and lower costs. Modeling techniques made it feasible to assess the circumstances under which RST would be promising and hence guide the prioritization of designing experimental studies to evaluate clinical utility. These findings should encourage payers and users to support the clinical development and adoption of RST approaches in MM.

## INTRODUCTION

Multiple myeloma (MM) is a malignant plasma cell disorder that accounts for approximately 10% of all hematologic malignancies [1]. The median overall survival (OS) of MM has nearly doubled over the past decade, ranging from 4 to 10 years. Improvement in survival has not been realized for all patients with a subgroup of patients consistently having a poor prognosis with a median OS of 2-3 years [2].

Markers of tumor burden and patient factors in MM are useful for predicting prognosis and comparison of trial results. The International Staging System (ISS) is based on tumor burden markers serum β2 microglobulin and albumin [3]. Patient factors also have prognostic significance, such as creatinine levels as a marker of renal failure [4]. Molecular markers of tumor biology also have prognostic value, such as florescence *in situ* hybridization (FISH) which demonstrated the survival for patients with and without a cytogenetic abnormality differs significantly (median OS being 53 versus 80 months) [5]. Molecular tests based on gene expression signatures further demonstrated independent prognostic value [6-8], and therefore incorporating of molecular markers within prognostic classification system has been acknowledged by a number of key opinion leaders [9, 10]. In particular classification based on ISS stage and del(17p) and/or t(4;14) on FISH has been proposed [11]. Incorporating the patient factor creatinine and gene expression signatures could further improve the sensitivity of prognostic classification.

The clinical utility of such a prognostic classification systems depends on its impact in treatment decision making. Currently, two methods dictate the current treatment approach towards newly diagnosed MM patients: a uniform treatment (UT) approach and a risk-stratified treatment (RST) approach [12]. UT comprises of similarly treating all patients irrespective of prognostic risk while RST adapts the treatment decision making process based on prognostic risk. The rational for RST is based on the desire to preserve quality-of-life for patients with a good prognosis or patient convenience or evidence for significant differences in treatment efficacy between risk groups. For example, patients with an expected survival beyond the median may wish to avoid high-risk of early onset of irreversible toxicity, while for others is the route of administration or number of hospital visits an important factor in treatment choice. Differences in treatment efficacy have been demonstrated; high-risk patients achieve significantly longer survival with bortezomib-based induction regimens compared to non-bortezomib-based regimens, while the improvement for patients without high-risk markers were not significantly different [11, 13-18]. It is also proposed that high-risk patients receive potentially curative experimental options that are otherwise not recommended due to high rates for treatment-related mortality (TRM) such as allogeneic stem cell transplant (allo-SCT) [19]

which has a 30% TRM with the possibility of cure for some patients despite initial poor risk factors [20].

Based on these findings, we propose a RST approach for newly diagnosed transplant-eligible patients where high-risk patients are stratified to bortezomib-based induction and maintenance regimens while standard-risk patients are treated with non-bortezomib-based regimens shown to be similarly effective without the risk of peripheral neurotoxicity. Such an approach has the potential to improve the health benefits of high-risk patients while also reducing costs, such as the Netherlands, Sweden, France and the UK, where the cost of bortezomib-based regimens is substantially higher compared to alternative regimens [21-24].

We aimed to address the potential therapeutic and economic implications of RST versus UT by conducting an economic evaluation for newly diagnosed transplant-eligible patients treated in the Netherlands. We further assessed whether the cost-effectiveness of RST where high-risk is detected by existing markers for adverse prognosis differs from a detection method that incorporates molecular markers.

## METHODS

### Patient selection

Patient-level data was obtained from the HOVON 65/GMMG-HD4 trial for a total of 815 newly diagnosed transplant eligible patients randomized to either bortezomib-based induction and maintenance (PAD-B) or chemotherapy-based induction and thalidomide maintenance (VAD-T). Details regarding the study protocol and patient characteristics can be found elsewhere [17].

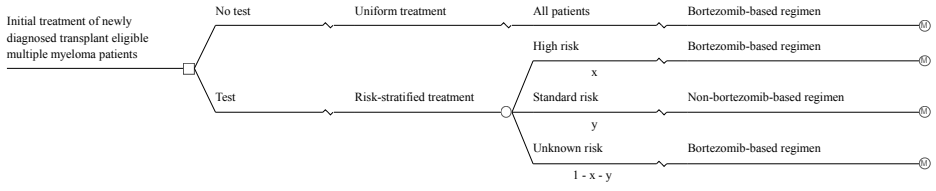



**Figure 1.** Decision analytic model comparing the costs and health benefits of the uniform treatment approach versus a risk-adapted approach.
*M denotes Markov structure and x and y denote the population prevalence of high and standard-risk prognosis, respectively.

### Decision analytic model

A decision analytic model was used to compare the total costs and health benefits for two strategies: 1) UT where all patients receive bortezomib-based regimens; and 2) RST where treatment is determined according to a patient's baseline prognosis (Figure 1). The RST approach categorized patients into three groups: 1) *high-risk* patients who were positive for one or more adverse prognostic markers; 2) *standard-risk* patients who did not fulfill the criteria for high-risk and for which data was available for all prognostic factors; and 3) *unknown-risk* patients for which high-risk criteria could not be met and data was missing for one or more prognostic factors. This third group of patients were retained in the analysis to reflect the reality of daily practice where data is often missing and treatment decisions must be made despite thereof.

High-risk disease was detected according to three alternatives: RST-1) using currently available prognostic tests in daily practice assumed to be the gold standard, such as ISS stage III, high creatinine levels (>2 mg/dL), t(4;14), or del(17p) on FISH; RST-2) using a positive result for a high-risk gene signature that detects patients with an OS less than 2 years [7]; and RST-3) using the gold standard scenario but with reclassification only for patients negative for the gold standard with a positive gene signature.

#### *Markov structure*

To model the disease process, a Markov model with a time horizon of 15 years comprised of 180 one-month cycles (Figure 2). The model assumes that a population of 408 patients enters the first cycle in the progression-free state and was at risk of two events, either death before progression or progressive disease. Patients entering the

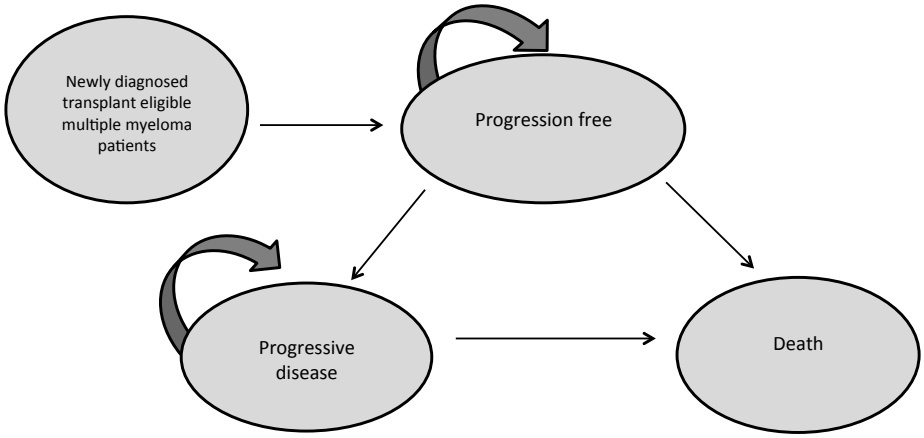



**Figure 2.** Markov structure of the decision analytic model

progressive state as the first event were then at risk for death after progression. For survival curves estimating the time to events, various distributions were assessed for model fit according to the criteria of the lowest AIC and the -2 log likelihood. A Weibull survival function for progression-free survival (PFS) estimated the probability of the first event, irrespective of death or progression. The probability of the first event type being progression or death was estimated using a logistic regression with 'event type' as the dependent variable and independent variables for treatment arm, risk group, and time to event. For progressive patients, an exponential survival function for overall survival (OS) after progression estimated the probability of death. Covariates for treatment arm and risk group were added to the survival functions to distinguish survival curves by treatment/risk subgroups.

***Input parameters***

Supplemental tables 1 through 4 (S1-4) provide details of the regression coefficients and supplemental tables 5 and 6 (S5, S6) provide details regarding the resource use, unit costs and all other input parameters assumed in the analysis.

***Transition probabilities***

The cumulative PFS and OS probability at time=$t_i$ was calculated using the following formula: $S(t_i)=\exp(-((t_i/\text{ scale})^{\wedge}\text{ shape}))$ where i=1 to 180 months and scale was parameterized as $\exp(\beta_0+(B_j x_j))$ with j corresponding to the model covariates.

The odds of the first event being progression was calculated using the following formula: odds of progression = $\exp(\beta_0+(B_j x_j)) / (1+ \exp(\beta_0+(B_j x_j))$ where j corresponds to the model covariates. It then follows that the odds of death = 1 – odds of progression.

***Unit prices and resource use***

Mean total costs were calculated from a Dutch healthcare payer perspective and were limited to the costs of the treatment protocol, cytogenetic testing, peripheral neuropathy and the costs of treatment for relapsed/refractory disease.

For calculating resource use (Table S5), the percentage of patients undergoing stem cell transplantation and the average number of cycles of induction and consolidation was based on the average across all patients selected for the analysis, which was decided based on no presence of significant difference between the two treatment arms. Maintenance therapy began in cycle (i.e., month) 12 for all patients but was made to correspond with the progression-free survival per treatment/risk patient subgroup. Costs of treatment protocol were then calculated by applying unit prices (Table S6) to the frequency or duration of resource use. Unit prices for anticancer agents were based on

**Table 1.** Description of patients according to HOVON 65/GMMG-HD4 treatment arm

| Descriptive parameters | All patients (n=815) | | PAD-B (n=408) | | VAD-T (n=407) | | P-value* |
|---|---|---|---|---|---|---|---|
| Age [mean (range)] | 56 (26 - 66) | | 56 (32 - 66) | | 56 (26 - 66) | | 0.67 |
| | No. | % | No. | % | No. | % | |
| ISS stage | | | | | | | 0.04† |
| III | 185 | 25% | 80 | 22% | 105 | 28% | |
| I/II | 559 | 75% | 292 | 78% | 267 | 72% | |
| missing | 71 | 9% | 36 | 9% | 35 | 9% | |
| t(4;14) FISH | | | | | | | 0.67 |
| positive | 68 | 14% | 35 | 14% | 33 | 13% | |
| negative | 435 | 86% | 212 | 86% | 223 | 87% | |
| missing | 312 | 38% | 161 | 39% | 151 | 37% | |
| del(17p) FISH | | | | | | | 0.12 |
| positive | 64 | 11% | 25 | 9% | 39 | 13% | |
| negative | 530 | 89% | 261 | 91% | 269 | 87% | |
| missing | 221 | 27% | 122 | 30% | 99 | 24% | |
| elevated creatinine (> 2 mg/dL) | | | | | | | 0.34 |
| positive | 80 | 10% | 36 | 9% | 44 | 11% | |
| negative | 735 | 90% | 372 | 91% | 363 | 89% | |
| missing | 0 | 0% | 0 | 0% | 0 | 0% | |
| % pts classified according to strategy RST-1: Gold standard | | | | | | | 0.07† |
| High-risk: any marker | 276 | 34% | 126 | 31% | 150 | 37% | |
| Standard-risk: no marker | 241 | 30% | 123 | 30% | 118 | 29% | |
| Unknown-risk | 298 | 37% | 159 | 39% | 139 | 34% | |
| % pts classified according to strategy RST-2: Gene signature only | | | | | | | 0.11 |
| High-risk positive | 63 | 8% | 26 | 6% | 37 | 9% | |
| High-risk negative | 224 | 27% | 119 | 29% | 105 | 26% | |
| Unknown-risk | 528 | 65% | 263 | 64% | 265 | 65% | |
| % pts classified according to strategy RST-3: Gold standard or gene signature | | | | | | | 0.10 |
| High-risk: any marker | 294 | 36% | 136 | 33% | 158 | 39% | |
| Standard-risk: no marker | 232 | 28% | 120 | 29% | 112 | 28% | |
| Unknown-risk | 289 | 35% | 152 | 37% | 137 | 34% | |
| Grade 3 PN | 135 | 17% | 88 | 22% | 47 | 12% | <0.001† |
| Grade 4 PN | 18 | 2% | 15 | 4% | 3 | 1% | 0.004† |

*p-values for student's t-test or chi square test of significant differences between treatment groups.
†Statistically significant difference in descriptive parameter across treatment arms at $\alpha = 0.10$.
FISH: fluorescence in situ hybridization; ISS: international staging system; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PN: peripheral neuropathy; RST: risk-stratified treatment; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

tariffs from the Dutch Healthcare Insurance Board [25]. The cost of SCT was taken from a multicenter study conducted in the Netherlands [26]. The cost of FISH testing for two markers (i.e., t(4;14) and del(17p)) was based on the tariff for a cytogenetic FISH panel established by the Netherlands Healthcare Authority [27]. Since the gene signature was not yet available for purchase at the time of analysis, a price of 1500 euros was assumed based on expert opinion.

The incidence of peripheral neuropathy (PN) was significant different between treatment arms and was hence made to vary according to regimen. The cost of treatment of PN was calculated based on the frequency and duration of treatment with the onset of PN corresponding to the average number of days since start of protocol observed per treatment arm. Patients with grade 3 or 4 PN were assumed to be treated daily with pregabalin for 600mg/day based on communication with a treating neurologist. The treatment duration was assumed 6 months until improvement of symptoms, which was based on the median time to improvement/resolution reported in APEX trial [28]. Grade 4 patients were assumed to require 6 months of revalidation therapy at an hourly rate and cost of a specialist visit based on the Dutch manual for costing in economic evaluations [29].

An average cost for treatment of relapsed/refractory disease was applied monthly until death following progression, which was taken from a cost study based on 139 Dutch patients treated for relapsed/refractory disease in the era of novel agents [22].

***Quality-adjusted life-years***

To calculate quality-adjusted life-years (QALYs), utility values were applied to patients in the progression-free (0.81) and progression (0.65) health states using estimates derived from the study by van Agthoven et al. (2004) [30].

All patients with grade 3 or 4 PN were assumed to experience a reduction in quality-of-life since the NCI CTCAE assessment criteria characterize grade 3 PN to be associated with severe pain interfering in daily life activities and grade 4 to be further disabling and leading to life-threatening consequences. A review of the literature identified the results of a study by McDermott et al. (2006) which reported a mean EQ-5D health state valuation for patients with severe neuropathic pain equal to 0.16 [31], which was applied for a duration of 6 months to patients experiencing grade 3 or 4 PN.

***Discounting***

In accordance with Dutch guidelines, the costs and effectiveness were discounted at a yearly rate of 4% and 1.5%, respectively [32].

**Scenario analysis and assessment of model uncertainty**

A scenario analysis was performed to assess the potential cost-effectiveness of a RST-3 approach where high-risk patients were assumed to receive experimental RIC allo-SCT by sibling donor. The impact of RIC allo-SCT was assumed to result in 30% of patients experiencing a 100% risk of TRM within one year following receipt of allo-SCT and 30% of patients experiencing cure with the remaining 40% experiencing a median PFS of 2 years. Patients experiencing cure were assumed at risk of death due to other causes based on the age-adjusted mortality rate for the Dutch population [33].

Univariate sensitivity analyses were performed to assess the impact of parameters expected to be influential on the incremental cost-effectiveness ratio (ICER).

| Gold standard ⬇ | Gene signature positive | Gene signature negative | Gene signature unknown | Totals: |
|---|---|---|---|---|
| High risk - Any of the following: ISS III, t(4;14), del(17p), or elevated creatinine | 45 | 68 | 163 | 276 |
| Standard risk - None of the following: ISS III, t(4;14), del(17p), or elevated creatinine | 9 | 96 | 136 | 241 |
| Unknown - no confirmed high risk with missing data on ISS III, t(4;14), del(17p), or elevated creatinine | 9 | 60 | 229 | 298 |
| Totals: | 63 | 224 | 528 | 815 |

**Figure 3.** Classification and reclassification of patients according to high-risk definitions applied in strategies RST-1, RST-2, RST-3.
Net reclassification rate: 9 + 9 = 18/815 = 2%;
Diagnostic performance of the gene signature: 40% sensitivity (45/45+68); 91% specificity (96/9+96).

## RESULTS

### Patient groups and health outcomes

Table 1 shows that the age at start of treatment and creatinine levels were similar between both treatment arms. The groups differed with respect to ISS stage with VAD-T patients more often presenting with advanced ISS stage compared to PAD-B patients (26% vs 20%; p=0.04). Consequently, more patients in the VAD-T group (37% vs 31%; p=0.07) were detected as high-risk by the gold standard. Despite missing data, the groups were also similar in terms of a positive test for t(4;14), deletion 17p or the high-risk gene signature. The groups differed regarding incidence of PN, with PAD-B patients having an increased incidence of grade 3 (22% vs 12%; p<0.001) or 4 (4% vs 1%; p=0.004) compared to VAD-T patients.

Figure 3 describes the reclassification of all 815 patients when incorporating the high-risk gene classifier into the gold standard method for detecting high-risk. In total, 9 unknown-risk and 9 standard-risk patients were reclassified as high-risk using the high-risk gene signature, which corresponds to a reclassification rate of 2.2% (18/815). The diagnostic performance of the high-risk gene classifier, when excluding patients with unknown-risk according to both methods, corresponds to a sensitivity and specificity of 40% (45/45+68) and 91% (96/9+96), respectively.

Table 2 provides the median PFS and OS after progression in months per patient group per treatment arm. Median PFS for PAD-B was significantly improved compared to the VAD-T treatment arm. PAD-B was also superior to VAD-T when stratifying patients into the high-risk group across all RST alternatives. In standard-risk patients, however, there was no difference in median PFS between the two treatment regimens, thus justifying the assumption that standard-risk patients receive VAD-T while high-risk patients received PAD-B. Unknown-risk patients were also assumed to receive PAD-B on account of the superior PFS compared to VAD-T in this group.

OS after progression, however, was similar in the PAD-B versus VAD-T treatment arm, likely due to the rescue effect of subsequent therapy for patients treated initially with VAD-T. Similar findings were also seen when stratifying by standard-risk and unknown-risk patients. For high-risk patients, median OS after progression favored PAD-B.

### Costs

UT was more costly (€194,613) compared to RST-1, RST-2 and RST-3 (range: €189,689-192,771) (Table 3). Cost-savings with all RST alternatives were found for induction and maintenance since standard-risk patients received the less costly induction regimen VAD

and thalidomide maintenance. Cost-savings were also seen for treatment of PN since VAD-T has a lower incidence compared to PAD-B. On the other hand, the costs of therapy for relapsed/refractory disease were higher for RST compared to UT due to improved OS after progression in standard-risk patients receiving VAD-T compared to PAD-B.

A comparison between RST alternatives shows little variation in the discounted total lifetime costs with a small difference mainly attributable to the costs of molecular testing.

**Table 2.** Kaplan Meier survival estimates of progression free and overall survival after progression according to patient group and HOVON 65/GMMG-HD4 treatment arm

| | Progression free survival [median (95% CI)] | | Overall survival after progression [median (95% CI)] | |
|---|---|---|---|---|
| **Strategy/patient group** | **PAD-B** | **VAD-T** | **PAD-B** | **VAD-T** |
| **UT** | | | | |
| Average all patients | 34.9 (30.6 - 38.6)† | 29.1 (25.5 - 32.4)† | 30.5 (23.2- 36.9) | 31.3 (22.6 - 37.0) |
| **RST** | | | | |
| **High-risk** | | | | |
| RST-1: Gold standard only | 27.4 (22.4 - 34.9)† | 19.9 (16.8 - 24.1)† | 20.4 (17.4 - 38.7)† | 16.4 (9.6 - 21.6)† |
| RST-2: Gene signature only | 17.0 (10.8 - 22.4)† | 11.6 (4.0 - 15.2)† | 14.1 (2.8 - 35.4) | 6.7 (3.4 - 13.2) |
| RST-3: Gold standard or gene signature | 26.0 (21.5 - 30.8)† | 19.9 (16.8 - 24.1)† | 20.4 (17.4 - 35.4)† | 16.4 (9.7 - 22.6)† |
| **Standard-risk** | | | | |
| RST-1: Gold standard only | 35.1 (30.4 - 40.7) | 37.0 (32.1 - 46.0) | 35.4 (30.3 - NR) | 37.4 (26.4 - 39.3) |
| RST-2: Gene signature only | 34.9 (30.1 - 39.6) | 36.5 (28.8 - 41.4) | 34.2 (20.0 - NR) | 42.7 (32.4 - NR) |
| RST-3: Gold standard or gene signature | 35.5 (30.6 - 40.8) | 41.1 (33.2 - 50.7) | 36.6 (30.3 - NR) | 37.4 (27.2 - 39.3) |
| **Unknown-risk** | | | | |
| RST-1: Gold standard only | 38.2 (32.6 - 44.4)† | 32.0 (25.1 - 37.7)† | 27.3 (20.1 - NR)† | 42.7 (32.9 - NR)† |
| RST-2: Gene signature only | 37.8 (31.5 - 41.5)† | 30.5 (26.7 - 35.2)† | 30.5 (22.0 - 38.7) | 32.9 (24.6 - 39.0) |
| RST-3: Gold standard or gene signature | 39.6 (34.7 - 45.1)† | 32.3 (25.5 - 37.7)† | 28.7 (20.1 - NR)† | 42.7 (32.9 - NR)† |

†Statistically significant difference in outcome within patient group across treatment arms at $\alpha = 0.10$ based on Wilcoxon statistic. CI: confidence interval; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; RST: risk-stratified treatment; SE: standard error; UT: uniform treatment; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

### Cost-effectiveness

At a higher total cost and lower health benefits (6.441 LYs or 4.385 QALYs), UT was the least cost-effective option at €30,215/LY or €44,382/QALY compared to all RST alterna-

**Table 3.** Modeled estimates of lifetime direct medical costs per treatment strategy

| | Induction | CAD/SCT/ HDM | Mainte-nance | Grade 3 PN | Grade 4 PN | Therapy for RRMM | Mo-lecular testing | Total costs (discounted) |
|---|---|---|---|---|---|---|---|---|
| **UT** | | | | | | | | |
| Bortezomib for all patients | € 11,737 | € 41,093 | € 29,734 | € 252 | € 711 | € 148,386 | € - | € 194,613 |
| **RST** | | | | | | | | |
| **RST-1: Gold standard** | | | | | | | | |
| average for all patients | € 8,359 | € 41,153 | € 20,715 | € 216 | € 548 | € 156,768 | € 868 | € 189,689 |
| High-risk - bortezomib | € 11,701 | € 40,582 | € 25,807 | € 247 | € 706 | € 97,500 | € 868 | € 156,201 |
| Standard-risk - no bortezomib | € 523 | € 41,347 | € 1,603 | € 134 | € 171 | € 188,407 | € 868 | € 187,102 |
| Unknown-risk - bortezomib | € 11,759 | € 41,455 | € 31,432 | € 255 | € 715 | € 179,308 | € 868 | € 218,217 |
| **RST-2: Gene signature only** | | | | | | | | |
| average for all patients | € 8,476 | € 41,172 | € 21,222 | € 217 | € 554 | € 158,973 | € 1,500 | € 192,545 |
| High-risk - bortezomib | € 11,508 | € 38,331 | € 12,236 | € 225 | € 681 | € 55,047 | € 1,500 | € 111,499 |
| Standard-risk - no bortezomib | € 523 | € 41,205 | € 1,572 | € 132 | € 169 | € 176,285 | € 1,500 | € 179,342 |
| Unknown-risk - bortezomib | € 11,764 | € 41,439 | € 30,979 | € 255 | € 715 | € 161,475 | € 1,500 | € 206,543 |
| **RST-3: Gold standard or gene signature** | | | | | | | | |
| average for all patients | € 8,437 | € 41,182 | € 20,765 | € 217.05 | € 552 | € 158,874 | € 2,368 | € 192,771 |
| High-risk - bortezomib | € 11,699 | € 40,502 | € 25,390 | € 246 | € 705 | € 98,246 | € 2,368 | € 157,924 |
| Standard-risk - no bortezomib | € 523 | € 41,543 | € 1,634 | € 135 | € 172 | € 200,728 | € 2,368 | € 197,274 |
| Unknown-risk - bortezomib | € 11,763 | € 41,505 | € 31,714 | € 255 | € 715 | € 180,010 | € 2,368 | € 220,333 |

CAD: cyclophosphamide adriamycin dexamethasone; HDM: high dose melphalan; PN: peripheral neuropathy; RRMM: relapsed refractory multiple myeloma; RST: risk-stratified treatment; SCT: stem cell transplantation.

tives. The average cost-savings for RST ranged from €1,842 to €4,924 for health gains in the range of 0.007 to 0.059 LYs or 0.009 to 0.040 QALYs, depending on the detection method for high-risk. The ICERs for all RST alternatives were considered dominant to UT as they offered a 'win-win' situation of health gains at reduced costs.

**Table 4.** Modeled estimates of the cost-effectiveness of stratified treatment versus uniform treatment

| | | Costs (€) | | Effectiveness (LYs) | | | Effectiveness (QALYs) | | |
|---|---|---|---|---|---|---|---|---|---|
| | % patients | Total costs | Incremental costs* | Total LYs | Incremental LYs * | ICER (LY) | Total QALYs | Incremental QALYs* | ICER (QALY) |
| **UT** | | | | | | | | | |
| Bortezomib for all patients | 100% | € 194,613 | | 6.441 | | | 4.385 | | |
| **RST** | | | | | | | | | |
| **RST-1: Stratified using gold standard** | | | | | | | | | |
| average for all patients | 100% | € 189,689 | € -4,924 | 6.449 | 0.007 | Dominant | 4.394 | 0.009 | Dominant |
| High-risk - bortezomib | 30.9% | € 156,201 | | 4.734 | | | 3.285 | | |
| Standard-risk - no bortezomib | 30.1% | € 187,102 | | 7.015 | | | 4.788 | | |
| Unknown-risk - bortezomib | 39.0% | € 218,217 | | 7.370 | | | 4.969 | | |
| **RST-2: Stratified using gene signature** | | | | | | | | | |
| average for all patients | 100% | € 192,545 | € -2,068 | 6.500 | 0.059 | Dominant | 4.425 | 0.040 | Dominant |
| High-risk - bortezomib | 6.4% | € 111,499 | | 2.487 | | | 1.723 | | |
| Standard-risk - no bortezomib | 29.1% | € 179,342 | | 6.614 | | | 4.534 | | |
| Unknown-risk - bortezomib | 64.5% | € 206,543 | | 6.846 | | | 4.645 | | |
| **RST-3: Gold standard or gene signature** | | | | | | | | | |
| average for all patients | 100% | € 192,771 | € -1,842 | 6.493 | 0.052 | Dominant | 4.419 | 0.034 | Dominant |
| High-risk - bortezomib | 33.3% | € 157,924 | | 4.676 | | | 3.240 | | |
| Standard-risk - no bortezomib | 29.4% | € 197,274 | | 7.349 | | | 4.995 | | |
| Unknown-risk - bortezomib | 37.3% | € 220,333 | | 7.440 | | | 5.018 | | |

*Calculated as RST minus UT. ICER: incremental cost-effectiveness ratio; LY: life-year; RIC allo-SCT: reduced-intensity conditioning allogeneic stem cell transplant; RST: risk-stratified treatment; QALY: quality-adjusted life-year; QoL: quality-of-life; UT: uniform treatment.

The cost-effectiveness ratio (CER) of the three RST scenarios was similar while their therapeutic and economic value differed. The RST-1 alternative offers the least improvement in health while at the greatest cost-savings at a cost-effectiveness ratio (CER) of €29,413/LY or €43,170/QALY. RST-2 and RST-3 offered greater health gains despite lower cost-savings compared to RST-1. RST-2 was less costly and offered greater health gain compared to RST-3, thus representing the more economical option with a CER of €29,622/LY or 43,512/QALY versus €29,689/LY or €43,623/QALY, respectively.

### Scenario analysis and parameter uncertainty

A hypothetical RST-3 scenario where all high-risk patients were assumed to receive experimental RIC allo-SCT was found to increase costs by €2,567 and improve health by 0.40 LYs and 0.30 QALYs compared to UT (Table 5). Survival-related parameters and the utility of the progressive health state were influential on the cost-effectiveness of RST, with lower values for these parameters leading to health loss with RST compared to UT. Incremental costs were influenced further by the cost of bortezomib demonstrating that at lower prices for bortezomib it is more cost-effective to treat all patients with bortezomib rather than spend the extra money to test all patients.

**Table 5.** Scenario analysis to estimate the effectiveness and cost-effectiveness of experimental therapy for high-risk patients

| | | Costs (€) | | Effectiveness (LYs) | | | Effectiveness (QALYs) | | |
|---|---|---|---|---|---|---|---|---|---|
| *Scenario analysis: High-risk receive RIC allo-SCT* | | **Total costs** | **Incremental costs*** | **Total LYs** | **Incremental LYs *** | **ICER (LY)** | **Total QALYs** | **Incremental QALYs*** | **ICER (QALY)** |
| **RST-3: Gold standard or gene signature** | | | | | | | | | |
| average for all patients | 100% | € 197,180 | € 2,567 | 6.841 | 0.400 | € 6,425 | 4.686 | 0.300 | € 8,545 |
| High-risk - bortezomib plus RIC allo-SCT | 33.3% | € 171,164 | | 5.721 | | | 4.041 | | |
| Cure: Background mortality only | 30% | € 158,913 | | 12.989 | | | 9.297 | | |
| TRM: Death within one year following transplant procedure | 30% | € 120,259 | | 0.987 | | | 0.686 | | |
| Median PFS of 2 years | 40% | € 218,531 | | 3.821 | | | 2.616 | | |
| **†Standard-risk - no bortezomib** | **29.4%** | € 197,274 | | 7.349 | | | 4.995 | | |
| **†Unknown-risk - bortezomib** | **37.3%** | € 220,333 | | 7.440 | | | 5.018 | | |

*Calculated as RST minus UT. †Modeled estimates of costs and effectiveness for these patients were unchanged since treatment was assumed to remain the same as in the base case. ICER: incremental cost-effectiveness ratio; LY: life-year; RIC allo-SCT: reduced-intensity conditioning allogeneic stem cell transplant; TRM: treatment-related mortality; QALY: quality-adjusted life-year; QoL: quality-of-life.

**Table 6.** Results of the univariate sensitivity analysis assessing the impact of influential parameters

| | | | RST-1 vs UT | | RST-2 vs UT | | RST-3 vs UT | |
|---|---|---|---|---|---|---|---|---|
| **Parameter/Strategy** | **Lower value** | **Upper value** | **Lower ICER (€/QALY)** | **Upper ICER (€/QALY)** | **Lower ICER (€/QALY)** | **Upper ICER (€/QALY)** | **Lower ICER (€/QALY)** | **Upper ICER (€/QALY)** |
| Median PFS | | | | | | | | |
| RST-1 | 19.4 | 62.9 | UT more effective | Dominant | | | | |
| RST-2 | 16.9 | 50.0 | | | UT more effective | Dominant | | |
| RST-3 | 19.7 | 63.4 | | | | | UT more effective | Dominant |
| Median OS after progression | | | | | | | | |
| RST-1 | 11.9 | 102.8 | UT more effective | € 64,144 | | | | |
| RST-2 | 10.9 | 81.0 | | | UT more effective | € 64,338 | | |
| RST-3 | 12.4 | 105.0 | | | | | UT more effective | € 65,077 |
| Utility progressive-free health state | 0.74 | 0.88 | Dominant | Dominant | Dominant | Dominant | Dominant | Dominant |
| Utility progression health state | 0.57 | 0.73 | UT more effective | Dominant | Dominant | Dominant | Dominant | Dominant |
| Utility grade 3 or 4 peripheral neuropathy | 0.10 | 0.22 | Dominant | Dominant | Dominant | Dominant | Dominant | Dominant |
| Cost of bortezomib per 3.5 mg vial | € 477 | € 1,432 | € 131,013 | Dominant | € 92,629 | Dominant | € 124,168 | Dominant |
| Cost of thalidomide per 50 mg | € 2 | € 5 | Dominant | Dominant | Dominant | Dominant | Dominant | Dominant |
| Cost of FISH testing | € 434 | € 1,301 | Dominant | Dominant | NA | NA | Dominant | Dominant |
| Cost of gene signature testing | € 750 | € 2,250 | NA | NA | Dominant | Dominant | Dominant | Dominant |

Parameters were varied according to the 95% CI when available and otherwise ± 50%. Shaded cells indicate influential parameter values. FISH: fluorescence in situ hybridization; ICER: incremental cost-effectiveness ratio; NA: not applicable; OS: overall survival; PFS: progression-free survival; QALY: quality-adjusted life-year; RST: risk-stratified treatment; UT: uniform treatment.

## DISCUSSION

We report here the first study to show using standard decision analytic modeling techniques the potential for both the therapeutic and economic benefits of RST approaches in MM that incorporates molecular markers of adverse prognosis. Decision analytical modeling was useful for assessing the potential clinical utility and budget impact of possible scenarios for RST management as well as the influence of the uncertainty surrounding the therapeutic evidence. As expected, the evidence regarding the survival outcomes for the treatment options in each risk group was found to be influential. Currently, it is only believed to be useful to stratify high-risk patients, particularly ultra-high-risk, to bortezomib, while the impact of RST on all other patients is not yet developed. To resolve the uncertainty about the therapeutic impact of RST approaches, the next step is to update the model with extended follow-up and additional patients once available.

The question remains which of the three methods for detecting high-risk disease is preferred since all three RST approaches incorporating molecular markers were found to be cost-effective compared to UT. It is not likely for all three scenarios to be feasible across all care settings or even appropriate for every treatment goal. Their feasibility requires the technical capability and an adequate amount of bone marrow cells to perform molecular testing. For settings that do have the capability, the detection methods that maximize health benefits for all patient groups are those that incorporate FISH with a high-risk gene signature. Despite a low sensitivity compared to FISH, the high-risk gene signature could have value when used alone as it did provide additional independent prognostic value as well as potential for cost-savings. The choice between the less sensitive detection method based only on the gene signature (RST-2) and that with greater sensitivity which incorporates additional markers (RST-3) will depend on the cutoff level imposed for the definition of high-risk prognosis which is actually dependent on the treatment goal at hand. Since a lower cutoff based on a median of 2 years will detect the most severe form of high-risk, an RST approach based on the gene signature alone would be preferred for treatment decisions with the goal of treating the most severe patients with an experimental treatment option that offers the change of cure despite a high-risk of TRM. A higher threshold, on the other hand, will identify additional patients with a less severe prognosis who have shown to demonstrate significant improvements with bortezomib. Nonetheless, our results demonstrate that RST management incorporating a high-risk gene signature to a current prognostic classification system can be a cost-saving method to identify additional patients of adverse prognosis that are not picked up by existing techniques.

Currently, bortezomib-based regimens as initial therapy is the preferred option for most patients regardless of prognostic risk [34, 35]. Only for high risk patients is stratification to alternative treatment options suggested [36], such as experimental regimens used alone or in combination with bortezomib and lenalidomide, more intensive bortezomib dose, or allo-SCT. However, there is currently no evidence for improved prognosis in high-risk patients receiving these alternatives. Taking into account the small number of high-risk patients, developers and key opinion leaders may face uncertainty regarding whether it is worthwhile to perform clinical efficacy studies for these alternatives, which are not only time-consuming and costly but also have a low chance of a significant finding. Using standard decision analytic modeling techniques, it has been shown that early economic evaluations assessing the potential therapeutic and economic impact of experimental treatment strategies allows decisions about product development to be based on the knowledge that is available prior to the clinical research phase [37]. By applying the same method, the findings shown here for the RST-3 scenario analysis demonstrate the circumstances that a strategy incorporating allo-SCT for high-risk patients must fulfill to improve prognosis and be economically favorable. The usefulness of developing such a model is that it can be adapted to evaluate the outcomes achieved with the alternatives mentioned above.

It is nowadays considered mandatory to perform cytogenetic testing of newly diagnosed patients [10]. However, our results demonstrate that a large, influential group of unknown-risk patients could not be classified due to missing data on cytogenetics. Surprisingly, this was observed within a trial setting where more stringent protocols exist suggesting that missing cytogenetic is surely to be an even greater issue in daily practice. The problem with missing data on prognostic factors is that it is not clear whether its occurrence is random. A randomly missing test could be associated with differences in treatment center protocol or technical issues with testing. If random, it is expected that the outcomes for such patients would be similar to that with the uniform approach. However, this group consistently demonstrated the most favorable prognosis suggesting that the missing data was not truly random and instead represents a more homogenous group. Given that there is potential for therapeutic and economic benefits to molecular testing, future research is needed to gain a better understanding of the barriers to the use of cytogenetic testing in daily practice. Insight into the attitude of practitioners toward molecular techniques may pinpoint strategies to improving their adoption in daily practice and consequently the quality and costs of care.

There are a few issues concerning the generalizability of the results. First, alternate molecular markers for high-risk could have been use in the detection method. It is well known that the definition of high-risk varies across studies, with details on cytogenetic

factors not consistently available. For example, recent evidence shows that t(4;14) may not be an independent marker for high-risk but instead indicative of an intermediate risk group consisting of heterogeneous patients[8, 38]. This marker was however retained on account of the numerous studies demonstrating bortezomib to overcome some high-risk features of t(4;14), indicating these patients would not be candidates for non-bortezomib-based regimens. Additional cytogenetic markers for adverse prognosis, such as t(14;16) and t(14;20)[39], are also candidate markers but were unfortunately not available. Inclusion of these markers could further improve the sensitivity of the gold standard assumed in our analysis, which presents an interesting scenario for further research. Second, the trial patients from whom the outcomes were based may not be representative of daily practice patients in terms of prognosis and/or comorbidities. The advantage of a decision analytic model is that it can easily be adapted with observational data to assess whether the outcomes seen here are realistic to the everyday daily practice patient. Third, preferences for particular regimens may differ across settings as a consequence of cost. For example, thalidomide is not available as generic in the US and has been shown in some instances to be more costly to bortezomib [40]. Given that the cost of bortezomib was influential in our analysis, an RST approach for bortezomib-based versus thalidomide-based regimens may not be financially attractive in the US compared to UT. Finally, the alternative induction regimen VAD may not be a relevant therapy option for some settings. At the time the trial was conducted, VAD was considered a relevant alternative, while many now consider thalidomide induction (TAD) the relevant alternative. However, we believe VAD to remain relevant given that its main active therapy is dexamethasone which continues to be commonly used as monotherapy within and outside a trial setting.

Some limitations to the methods also deserve mentioning. First, the diagnostic performance observed for the high-risk gene signature used in our analyses may be considered overly optimistic since the same patients from which it was developed were used to assess its performance and hence cost-effectiveness [41]. The results could also be susceptible to a phenomenon known as the spectrum effect which leads to a falsely high test performance when the study population used to evaluate the performance of a test does not adequately represent all subgroups with the target population [42]. Consequently, the diagnostic performance and hence cost-effectiveness of the gene signature could differ depending on the study population, particularly when the genetic determinants of high-risk disease differ in other patient populations. Future research efforts should aim to externally validate the findings of our study using an independent dataset. Second, the utility values to compute QALYs were based on conventional as opposed to targeted therapies. The choice of utility values was simply based on available estimates in the literature. The lack thereof is a consequence of the fact that preference-

based measures, such as the EuroQol five-dimensional (EQ-5D) questionnaire, are not included in many cancer trials but instead non-preference-based measures to assess patient-reported outcomes. Efforts to generate updated utility values from non-preference-based measures for multiple myeloma are currently underway using the method of mapping [43, 44].

To conclude, RST approaches using molecular markers for bortezomib-based regimens versus alternative regimens in newly diagnosed transplant-eligible MM patients have the potential to be more effective and less costly compared to UT. A few issues remain, however, that need to be addressed by future research to properly assess the value of RST in MM: 1) consensus on the definition of high-risk and the relevant treatment options; 2) the barriers to adoption of molecular testing in daily practice; and 3) model validation in independent datasets. A re-assessment of the effectiveness and cost-effectiveness of RST is needed as additional evidence regarding the clinical utility of markers of adverse prognosis in MM comes available.

## ACKNOWLEDGEMENTS

The authors would like to thank Joost Jorgen for his insight into the treatment of chemotherapy-induced peripheral neuropathy, Marian Stevens-Kroef for explaining the technical aspects of molecular diagnostics, Berna Beverloo for sharing her knowledge about the feasibility and reimbursement of molecular diagnostics in the Netherlands, Rowan Kuiper and Erik van Beers for explaining the technical aspects of the high-risk gene signature applied in this study, and Joost van Esser, MD and Mark-David Levin, MD for providing insight into the current care of multiple myeloma patients in the Netherlands and the feasibility of risk-stratified treatment.



## REFERENCES

1. Palumbo A, Anderson K. Multiple myeloma. N Engl J Med 2011; 364(11): 1046-60.
2. Avet-Loiseau H. Ultra high-risk myeloma. Hematology Am Soc Hematol Educ Program 2010; 2010: 489-93.
3. Greipp PR, San Miguel J, Durie BG, et al. International staging system for multiple myeloma. J Clin Oncol 2005; 23(15): 3412-20.
4. Munshi NC, Anderson KC, Bergsagel PL, et al. Consensus recommendations for risk stratification in multiple myeloma: report of the International Myeloma Workshop Consensus Panel 2. Blood 2011; 117(18): 4696-700.
5. Richardson PG, Laubach J, Mitsiades CS, et al. Managing multiple myeloma: the emerging role of novel therapies and adapting combination treatment for higher risk settings. Br J Haematol 2011.

6. Decaux O, Lode L, Magrangeas F, et al. Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myelome. J Clin Oncol 2008; 26(29): 4798-805.
7. Kuiper R, Broyl A, de Knegt Y, et al. A gene expression signature for high-risk multiple myeloma. Leukemia 2012; 26(11): 2406-13.
8. Shaughnessy JD, Jr., Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. Blood 2007; 109(6): 2276-84.
9. Dispenzieri A, Rajkumar SV, Gertz MA, et al. Treatment of newly diagnosed multiple myeloma based on Mayo Stratification of Myeloma and Risk-adapted Therapy (mSMART): consensus statement. Mayo Clin Proc 2007; 82(3): 323-41.
10. Fonseca R, Bergsagel PL, Drach J, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia 2009; 23(12): 2210-21.
11. Neben K, Lokhorst HM, Jauch A, et al. Administration of bortezomib before and after autologous stem cell transplantation improves outcome in multiple myeloma patients with deletion 17p. Blood 2012; 119(4): 940-8.
12. Rajkumar SV, Gahrton G, Bergsagel PL. Approach to the treatment of multiple myeloma: a clash of philosophies. Blood 2011; 118(12): 3205-11.
13. Mateos MV, Richardson PG, Schlag R, et al. Bortezomib plus melphalan and prednisone compared with melphalan and prednisone in previously untreated multiple myeloma: updated follow-up and impact of subsequent therapy in the phase III VISTA trial. J Clin Oncol 2010; 28(13): 2259-66.
14. San Miguel JF, Schlag R, Khuageva NK, et al. Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. N Engl J Med 2008; 359(9): 906-17.
15. Avet-Loiseau H, Leleu X, Roussel M, et al. Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). J Clin Oncol 2010; 28(30): 4630-4.
16. Cavo M, Tacchetti P, Patriarca F, et al. Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: a randomised phase 3 study. Lancet 2010; 376(9758): 2075-85.
17. Sonneveld P, Schmidt-Wolf IG, van der Holt B, et al. Bortezomib induction and maintenance treatment in patients with newly diagnosed multiple myeloma: results of the randomized phase III HOVON-65/ GMMG-HD4 trial. J Clin Oncol 2012; 30(24): 2946-55.
18. Morgan GJ, Gregory WM, Davies FE, et al. The role of maintenance thalidomide therapy in multiple myeloma: MRC Myeloma IX results and meta-analysis. Blood 2012; 119(1): 7-15.
19. Lokhorst H, Einsele H, Vesole D, et al. International Myeloma Working Group consensus statement regarding the current status of allogeneic stem-cell transplantation for multiple myeloma. J Clin Oncol 2010; 28(29): 4521-30.
20. Schilling G, Hansen T, Shimoni A, et al. Impact of genetic abnormalities on survival after allogeneic hematopoietic stem cell transplantation in multiple myeloma. Leukemia 2008; 22(6): 1250-5.
21. 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. J Clin Pharm Ther 2011; 36(1): 19-26.
22. Gaultney JG, Franken MG, Tan SS, et al. Real-world health care costs of relapsed/refractory multiple myeloma during the era of novel cancer agents. J Clin Pharm Ther 2013; 38(1): 41-7.

23. Ghatnekar O, Alvegard 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. Clin Ther 2008; 30(9): 1704-13.
24. Picot J, Cooper K, Bryant J, Clegg AJ. The clinical effectiveness and cost-effectiveness of bortezomib and thalidomide in combination regimens with an alkylating agent and a corticosteroid for the first-line treatment of multiple myeloma: a systematic review and economic evaluation. Health Technol Assess 2011; 15(41): 1-204.
25. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Farmacotherapeutisch Kompas. [cited 2013 January 30]; Available from: http://www.fk.cvz.nl/
26. Blommestein HM, Verelst SG, Huijgens PC, et al. Real-world costs of autologous and allogeneic stem cell transplantations for haematological diseases: a multicentre study. Ann Hematol 2012.
27. Dutch Healthcare Authority [Nederlandse Zorgautoriteit (NZa)]. DBC Healthcare Product Tariff Application [DBC Zorgproducten Tariefapplicatie] [cited 2013 January 30]; Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/Welkom.aspx
28. Richardson PG, Sonneveld P, Schuster MW, et al. Reversibility of symptomatic peripheral neuropathy with bortezomib in the phase III APEX trial in relapsed multiple myeloma: impact of a dose-modification guideline. Br J Haematol 2009; 144(6): 895-903.
29. Hakkart-van Roijen L, Tan SS, Bouwmans-Frijters CAM. Dutch manual for costing: Methods and reference prices for economic evaluations in healthcare, updated version [Handleiding voor kostenonderzoek: Methoden en referentieprijzen voor economische evaluaties in de gezondheidszorg, geactualiseerde versie]. 2010.
30. van Agthoven M, Segeren CM, Buijt I, et al. A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma; a prospective randomised phase III study. Eur J Cancer 2004; 40(8): 1159-69.
31. McDermott AM, Toelle TR, Rowbotham DJ, et al. The burden of neuropathic pain: results from a cross-sectional survey. Eur J Pain 2006; 10(2): 127-35.
32. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Guidelines for pharmacoeconomic research, updated version. Diemen, Netherlands: CVZ, 2006 [cited 2013 January 30]; Available from: http://www.cvz.nl/binaries/live/cvzinternet/hst_content/nl/documenten/rubriek+zorgpakket/cfh/richtlijnen+farmaco-economisch+onderzoek+-+engels.pdf
33. Association of Comprehensive Cancer Centres [Integral Kankercentrum Nederland]. Cancer figures [Cijfers over kanker]. http://cijfersoverkanker.nl/
34. Rajkumar SV. Multiple myeloma: 2013 update on diagnosis, risk-stratification, and management. Am J Hematol 2013; 88(3): 225-35.
35. Zweegman S, Lokhorst HM, Levin MD, et al. Richtlijnen behandeling multipel myeloom 2012. Nederlands Tijdschrift voor Hematologie 2012; 9(8): 300-20.
36. Rajkumar SV. Treatment of multiple myeloma. Nat Rev Clin Oncol 2011; 8(8): 479-91.
37. Vallejo-Torres L, Steuten LM, Buxton MJ, et al. Integrating health economics modeling in the product development cycle of medical devices: a Bayesian approach. Int J Technol Assess Health Care 2008; 24(4): 459-64.
38. Moreau P, Attal M, Garban F, et al. Heterogeneity of t(4;14) in multiple myeloma. Long-term follow-up of 100 cases treated with tandem transplantation in IFM99 trials. Leukemia 2007; 21(9): 2020-4.
39. Bergsagel PL, Mateos MV, Gutierrez NC, et al. Improving overall survival and overcoming adverse prognosis in the treatment of cytogenetically high-risk multiple myeloma. Blood 2012.

40. Teitelbaum A, Ba-Mancini A, Huang H, Henk HJ. Health care costs and resource utilization, including patient burden, associated with novel-agent-based treatment versus other therapies for multiple myeloma: findings using real-world claims data. Oncologist 2013; 18(1): 37-45.
41. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012; 98(9): 691-8.
42. Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. Fam Pract 2008; 25(5): 390-6.
43. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and mapped health state utility values for multiple myeloma cancer. Value Health 2012; 15(8): 1059-68.
44. Versteegh MM, Leunis A, Luime JJ, et al. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. Med Decis Making 2012; 32(4): 554-68.

## SUPPLEMENTAL DATA

**Table S1.** Regression coefficients to model survival in the uniform treatment strategy

| | coefficient | se | p-value |
|---|---|---|---|
| **Regression coefficients for Weibull PFS** | | | |
| constant | 3.553 | 0.111 | <0.001 |
| PAD-B (versus VAD-T) | 0.1712 | 0.0713 | 0.0163 |
| shape | 1.1811 | 0.0428 | |
| **Regression coefficients for exponential OS model after progression** | | | |
| constant | 3.7071 | 0.2143 | <0.001 |
| PAD-B (versus VAD-T) | 0.0411 | 0.1375 | 0.765 |
| shape | 1.0000 | | |
| **Regression coefficients for logistic model to estimate probability of first event being progression (versus death)** | | | |
| constant | 0.1377 | 0.4706 | 0.7699 |
| PAD-B (versus VAD-T) | -0.3104 | 0.3006 | 0.3019 |
| time to event (months) | 0.1481 | 0.0184 | <0.001 |

OS: overall survival; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PFS: progression-free survival; se: standard error; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

**Table S2.** Regression coefficients to model survival in risk-stratified treatment-1 strategy

| | coefficient | se | p-value |
|---|---|---|---|
| **Regression coefficients for Weibull PFS** | | | |
| constant | 3.6853 | 0.1222 | <0.001 |
| PAD-B (versus VAD-T) | 0.1593 | 0.0705 | 0.0238 |
| high-risk (versus unknown-risk) | -0.3614 | 0.0825 | <0.001 |
| standard-risk (versus unknown-risk) | 0.0028 | 0.0895 | 0.9746 |
| shape | 1.1971 | 0.0433 | |

**Table S2.** (Continued)

| | coefficient | se | p-value |
|---|---|---|---|
| **Regression coefficients for exponential OS model after progression** | | | |
| constant | 3.9988 | 0.2439 | <0.001 |
| PAD-B (versus VAD-T) | 0.0155 | 0.1377 | 0.9104 |
| high-risk (versus unknown-risk) | -0.7039 | 0.1559 | <0.001 |
| standard-risk (versus unknown-risk) | 0.0458 | 0.1934 | 0.8128 |
| shape | 1 | | |
| **Regression coefficients for logistic model to estimate probability of first event being progression (versus death)** | | | |
| constant | 0.3591 | 0.5361 | 0.503 |
| PAD-B (versus VAD-T) | -0.3256 | 0.3016 | 0.2803 |
| high-risk (versus unknown-risk) | -0.3079 | 0.3447 | 0.3717 |
| standard-risk (versus unknown-risk) | -0.1798 | 0.4256 | 0.6726 |
| time to event (months) | 0.1469 | 0.0186 | <0.001 |

OS: overall survival; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PFS: progression-free survival; se: standard error; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

**Table S3.** Regression coefficients to model survival in risk-stratified treatment-2 strategy

| | coefficient | se | p-value |
|---|---|---|---|
| **Regression coefficients for Weibull PFS** | | | |
| constant | 3.6058 | 0.1095 | <0.001 |
| PAD-B (versus VAD-T) | 0.1728 | 0.0684 | 0.0115 |
| high-risk (versus unknown-risk) | -1.0381 | 0.1152 | <0.001 |
| standard-risk (versus unknown-risk) | 0.0169 | 0.0787 | 0.8304 |
| shape | 1.2302 | 0.0443 | |
| **Regression coefficients for exponential OS model after progression** | | | |
| constant | 3.7341 | 0.2175 | <0.001 |
| PAD-B (versus VAD-T) | 0.0704 | 0.1381 | 0.6102 |
| high-risk (versus unknown-risk) | -1.0142 | 0.1888 | <0.001 |
| standard-risk (versus unknown-risk) | 0.1748 | 0.1731 | 0.3124 |
| shape | 1 | | |
| **Regression coefficients for logistic model to estimate probability of first event being progression (versus death)** | | | |
| constant | 0.262 | 0.4926 | 0.5948 |
| PAD-B (versus VAD-T) | -0.2864 | 0.3034 | 0.3452 |
| high-risk (versus unknown-risk) | -0.3154 | 0.393 | 0.4223 |
| standard-risk (versus unknown-risk) | -0.3606 | 0.3557 | 0.3108 |
| time to event (months) | 0.1472 | 0.0186 | <0.001 |

OS: overall survival; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PFS: progression-free survival; se: standard error; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

**Table S4.** Regression coefficients to model survival in risk-stratified treatment-3 strategy

| | coefficient | se | p-value |
|---|---|---|---|
| **Regression coefficients for Weibull PFS** | | | |
| constant | 3.6993 | 0.1217 | <0.001 |
| PAD-B (versus VAD-T) | 0.1607 | 0.0701 | 0.0219 |
| high-risk (versus unknown-risk) | -0.4043 | 0.0814 | <0.001 |
| standard-risk (versus unknown-risk) | 0.0220 | 0.0913 | 0.8092 |
| shape | 1.2026 | 0.0434 | |
| **Regression coefficients for exponential OS model after progression** | | | |
| **constant** | 4.0000 | 0.2357 | <0.001 |
| PAD-B (versus VAD-T) | 0.0182 | 0.1377 | 0.8905 |
| high-risk (versus unknown-risk) | -0.6953 | 0.1554 | <0.001 |
| standard-risk (versus unknown-risk) | 0.1180 | 0.2027 | 0.5605 |
| shape | 1.0000 | | |
| **Regression coefficients for logistic model to estimate probability of first event being progression (versus death)** | | | |
| constant | 0.3889 | 0.5353 | 0.4675 |
| PAD-B (versus VAD-T) | -0.328 | 0.3018 | 0.2771 |
| high-risk (versus unknown-risk) | -0.3735 | 0.3434 | 0.2768 |
| standard-risk (versus unknown-risk) | 0.0042 | 0.4586 | 0.9927 |
| time to event (months) | 0.1454 | 0.0185 | <0.001 |

OS: overall survival; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PFS: progression-free survival; se: standard error; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

**Table S5.** Input parameters for resource use, quality of life utilities and discounting

| Input parameters | Estimate | Source | Reference |
|---|---|---|---|
| **Frequency and duration of resource use** | | | |
| Duration of induction (no. cycles) | 3 | Sonneveld et al. (2012) | [17] |
| Proportion of patients receiving auto-SCT | 88% | Sonneveld et al. (2012) | [17] |
| Cycle when CAD/auto- SCT/HDM is given | 6 | Sonneveld et al. (2012) | [17] |
| Proportion of patients receiving 1 cycle HDM | 52% | Sonneveld et al. (2012) | [17] |
| Proportion of patients receiving 2 cycles HDM | 34% | Sonneveld et al. (2012) | [17] |
| Proportion of patients receiving allo-SCT | 8% | Sonneveld et al. (2012) | [17] |
| Cycle when allo-SCT is given | 9 | Sonneveld et al. (2012) | [17] |
| Proportion of patients with grade 3 PN | | Sonneveld et al. (2012) | [17] |
| PAD-B treatment arm | 22% | Sonneveld et al. (2012) | [17] |
| VAD-T treatment arm | 12% | Sonneveld et al. (2012) | [17] |

**Table S5.** (Continued)

| Input parameters | Estimate | Source | Reference |
|---|---|---|---|
| Proportion of patients with grade 4 PN | | | |
| PAD-B treatment arm | 4% | Sonneveld et al. (2012) | [17] |
| VAD-T treatment arm | 1% | Sonneveld et al. (2012) | [17] |
| Duration of grade 3 PN (total no. of cycles) | 6 | Richardson et al. (2009) | [28] |
| Duration of grade 4 PN (total no. of cycles) | 6 | Richardson et al. (2009) | [28] |
| Incident cycle of PN grade 3 | | Sonneveld et al. (2012) | [17] |
| PAD-B treatment arm | 4 | Sonneveld et al. (2012) | [17] |
| VAD-T treatment arm | 12 | Sonneveld et al. (2012) | [17] |
| Incident cycle of PN grade 4 | | Sonneveld et al. (2012) | [17] |
| PAD-B treatment arm | 1 | Sonneveld et al. (2012) | [17] |
| VAD-T treatment arm | 9 | Sonneveld et al. (2012) | [17] |
| Daily dose of pregabalin for grade 3 or 4 PN (mg) | 600 | Based on expert clinical opinion | |
| Number of revalidation visits per cycle | 22 | Based on expert clinical opinion | |
| Number of specialist visits for PN per month | 2 | Based on expert clinical opinion | |
| **Health state utilities** | | | |
| Progression-free health state | 0.81 | van Agthoven et al. (2004) | [30] |
| Progression health state | 0.65 | van Agthoven et al. (2004) | [30] |
| Grade 3 or 4 PN | 0.16 | McDermott et al.(2006) | [31] |
| **Discount rate** | | | |
| Costs | 4% | Dutch Healthcare Insurance Board (CVZ) | [32] |
| Effects | 1.5% | Dutch Healthcare Insurance Board (CVZ) | [32] |

CAD: cyclophosphamide adriamycin dexamethasone; HDM: high dose melphalan; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PN: peripheral neuropathy; SCT: stem cell transplantation; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

**Table S6.** Unit prices

| Unit prices | Unit | Unit price (Jan 2012) | Source | Reference |
|---|---|---|---|---|
| **Anticancer agents** | | | | |
| Vincristine | 1mg | € 9.74 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Doxorubicin | 1 mg | € 1.48 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Dexamethasone | 1 mg | € 0.06 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Thalidomide | 50 mg | € 3.50 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Bortezomib | 3.5 mg | € 955 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Cyclophosphamide | 1 mg | € 0.04 | Dutch Healthcare Insurance Board (CVZ) | [25] |

**Table S6.** (Continued)

| Unit prices | Unit | Unit price (Jan 2012) | Source | Reference |
|---|---|---|---|---|
| **Anticancer agents** | | | | |
| Melphalan | 1 mg | € 3.49 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Etoposide | 1 mg | € 1.14 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Cisplatin | 1mg | € 0.50 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Prednisone | 1 mg | € 0.01 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| **Pre-transplantation phase** | | | | |
| G-CSF (filgrastim) | 30x10^6 E/ML | € 103.01 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Auto-SCT pre-transplantation | 1 unit | € 11,937 | Blommestein et al. (2012) | [26] |
| RIC Allo-SCT pre-transplantation | 1 unit | € 31,480 | Blommestein et al. (2012) | [26] |
| **Transplantation and follow-up phase** | | | | |
| Auto-SCT pre-transplantation | 1 unit | € 21,123 | Blommestein et al. (2012) | [26] |
| RIC Allo-SCT pre-transplantation | 1 unit | € 24,894 | Blommestein et al. (2012) | [26] |
| **Post-transplantation 1-year follow-up phase** | | | | |
| Auto-SCT pre-transplantation | 1 unit | € 12,610 | Blommestein et al. (2012) | [26] |
| RIC Allo-SCT pre-transplantation | 1 unit | € 45,549 | Blommestein et al. (2012) | [26] |
| **Peripheral neuropathy** | | | | |
| Pregabalin | 1 mg | € 0.01 | Dutch Healthcare Insurance Board (CVZ) | [25] |
| Revalidation therapy | 1 hour | € 110 | Hakkart van-Rooijen et al. (2010) | [29] |
| Specialist visit | 1 unit | € 72 | Hakkart van-Rooijen et al. (2010) | [29] |
| | 1 month | € 3,981 | Gaultney et al. (2012) | [22] |
| **Relapsed/refractory therapy** | | | | |
| **Molecular testing** | | | | |
| FISH test | per test | € 868 | Dutch Healthcare Authority (NZa) | [27] |
| High-risk gene signature | per test | € 1,500 | Based on expert opinion | |

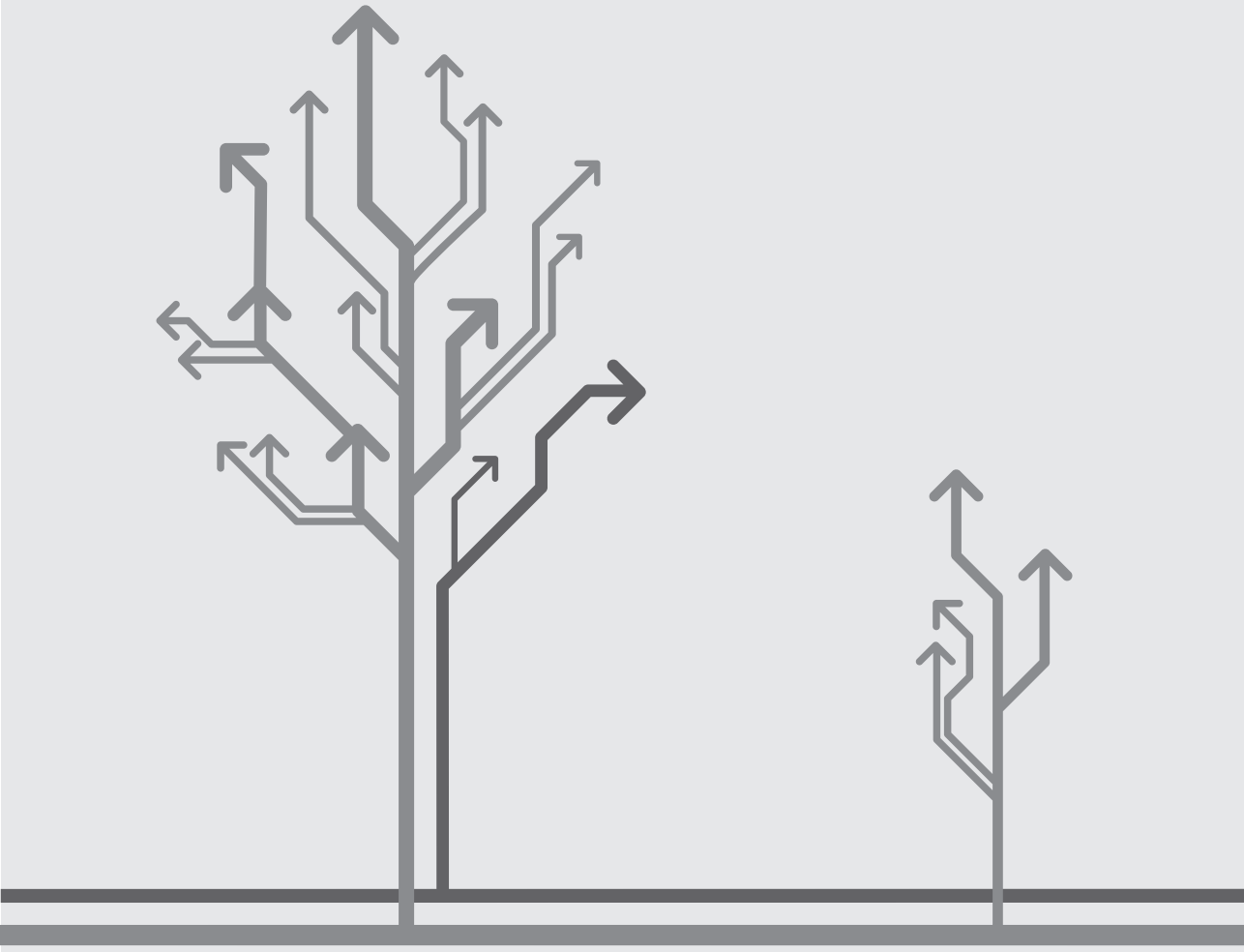
CAD: cyclophosphamide adriamycin dexamethasone; FISH: fluorescence in situ hybridization; G-CSF: granulocyte-colony stimulating factor; HDM: high dose melphalan; PAD-B: bortezomib adriamycin dexamethasone as induction and bortezomib maintenance; PN: peripheral neuropathy; RIC: reduced-intensity conditioning; SCT: stem cell transplantation; VAD-T: vincristine adriamycin dexamethasone as induction and thalidomide maintenance.

# Chapter 8.

## Application of cost-effectiveness analysis to demonstrate the potential value of companion diagnostics in chronic myeloid leukemia

Jennifer G. Gaultney, Enrique Sanhueza, Jeroen J. Janssen, William K. Redekop, Carin A. Uyl-de Groot

*Published in: Pharmacogenomics. 2011 Mar;12(3):411-21.*



## ABSTRACT

A cost-effectiveness analysis was performed to assess the potential value of companion diagnostics in supporting treatment decisions for dasatinib and nilotinib in chronic myeloid leukemia. A decision model was developed, and model inputs were taken from the literature and publicly available sources. The perspective of the healthcare sector in the Netherlands was used. Sensitivity and scenario analyses were performed to assess uncertainty in the results. The companion diagnostic could improve health and reduce costs though the estimates are uncertain due to limited evidence for comparative effectiveness between dasatinib and nilotinib. The results were sensitive to the cost of treatment, utility of progression and progression-free survival. This case demonstrates the use of cost-effectiveness analysis at an early stage of health technology assessment to generate economic evidence for their use in treatment decisions and to support decision-making for their development.

## INTRODUCTION

Recent advances in oncology treatment have resulted in a new treatment paradigm based on agents that specifically target biomarkers. Optimal use of *targeted* agents can be determined by companion diagnostics, either genomic- or tumor-based, that provide information about treatment efficacy at a functional level of the target for a particular tumor. Their application may lead to improved guidance of drug discovery at its earliest stages and direct individualized care of cancer patients [1]. The introduction of such tests into clinical practice has been lagging due to a variety of barriers such as the complexity of the technology and various clinical, economic and organizational barriers [2, 3].

### Cost-effectiveness of companion diagnostics

Cost-effectiveness analysis (CEA) is a method for comparing the relative costs and benefits of competing alternatives under conditions of limited resources. In the healthcare sector, an incremental cost-effectiveness ratio is calculated whereby a new treatment strategy is compared with current practice[4]. CEA has played a crucial role within the healthcare framework to inform decisions at various levels. Proponents believe that it has a similar role in assessing the clinical value of personalized medicine [5]. Testing before treating may be economically viable if the savings gained by avoiding ineffective treatment and adverse events are greater than the costs of testing. One well-known study reported favorable cost-effectiveness results for a companion diagnostic in breast cancer [6]. However, systematic reviews of CEAs in companion diagnostics have highlighted the scarcity of such studies and the need for higher quality economic evidence supporting their use [7, 8].

### Application of companion diagnostics in CML

Chronic myeloid leukemia (CML) is a myeloproliferative disorder characterized by a chromosomal translocation that gives rise to a gene that encodes a deregulated tyrosine kinase [9]. An effective tyrosine kinase inhibitor (TKI), imatinib, has demonstrated significant activity in CML. This novel therapy has few side effects, a favorable record of efficacy [10], and was found to be cost-effective [11-16]. Nonetheless, resistance to imatinib is common with up to 80% becoming resistant after five years, depending on disease phase [10, 17, 18]. The only clinically significant predictor of resistance to TKI therapy is the presence of T315I mutation [19].

Newer, alternative TKIs (nilotinib and dasatinib) are used to treat patients resistant to both standard-dose (400 mg daily) and high-dose (800 mg daily) imatinib. In the absence of a prospective study and a CEA comparing the two, it is difficult to make a

general recommendation for one over the other. Poor quality evidence demonstrating the cost-effectiveness of dasatinib and nilotib has resulted in some decision-makers, such as the National Institute for Health and Clinical Effectiveness (NICE), to claim that dasatinib and nilotinib cannot be recommended for use [20]. The choice between two alternative TKIs for which little comparative effectiveness evidence is available presents an opportunity for applying personalized medicine.

As improved outcomes in CML have been shown to be associated with a patient's response to therapy [21, 22], many practitioners are using fluorescence in situ hybridization (FISH) testing for response monitoring after start of treatment [23]. Prediction of an individual patient's response profile to therapy prior to therapy with a companion diagnostic may improve outcomes in current care by application of personalized medicine.

We conducted a CEA to assess the potential cost-savings and health gains of treating according to the results of a companion diagnostic in CML. The companion diagnostic described in this report is based on kinase activity profiling technology using a diagnostic microarray platform for predicting response profile to therapy prior to start of therapy. It has been applied in several therapy response prediction studies, including CML [24-26].

## METHODS

### Analytical approach

A decision model was developed for a baseline population of CML patients in the chronic phase who were eligible for second-line therapy with TKIs, failed to respond to high dose imatinib, and who lack the T315I mutation. The perspective of the healthcare sector in the Netherlands was used. Direct medical costs consumed by the hospital and/or healthcare insurer relevant to the comparator strategies were identified and valued. A two-year time horizon was chosen because the median follow-up available in the literature at the time of conducting the present CEA for patients undergoing second-line therapy with TKIs was limited [27, 28]. Costs and effects were discounted at a rate of 4% and 1.5%, respectively, in abidance with the Dutch guidelines for cost-effectiveness analyses[29].

#### *Model structure*

Figure 1 presents the structure of the model comparing the following treatment strategies: 1) *no testing strategy*, which represents current care and consists of second-line treatment with dasatinib during the first year and a switch to nilotinib during the second

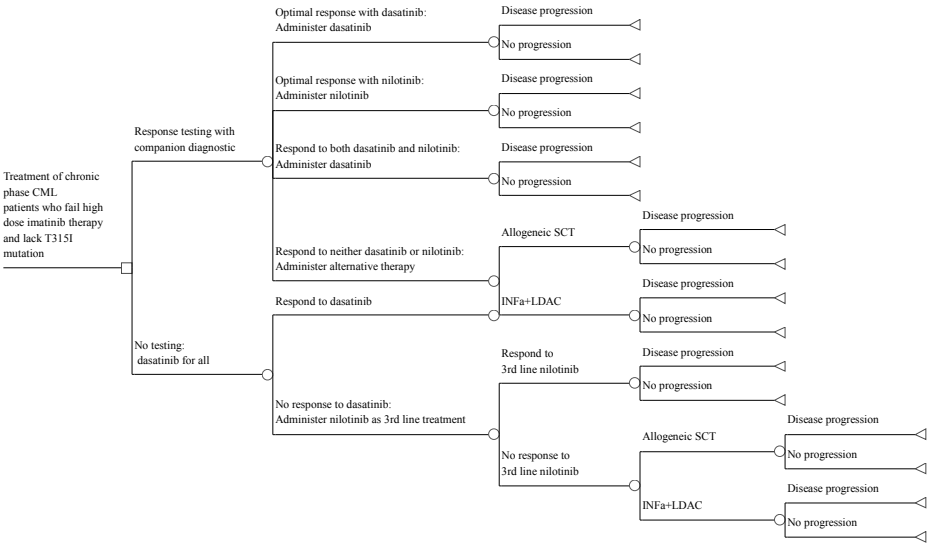



**Figure 1.** Structure of the decision model comparing the companion diagnostic strategy to the no testing strategy in second-line treatment for CML.
SCT: stem cell transplant; IFNa+LDAC: interferon-alpha plus low-dose arabinosylcytosine.

year if failing to respond; and 2) *companion diagnostic strategy*, where treatment decisions were made on the basis of the patient's response profile as depicted by biomarker-based testing.

The no testing strategy was based on the treatment recommendations for the target patient population as described by the Dutch Handbook for treatment of hematological disorders [30]. It was assumed that dasatinib is first administered to patients who fail first-line imatinib therapy followed by treatment with nilotinib after failing dasatinib. Patients not responding to dasatinib within the first 12 months were assumed to receive nilotinib for the second 12-month period.

In the second strategy, patients were classified into one of the following categories according to the predicted response by the companion diagnostic: 1) optimal responder to dasatinib, 2) optimal responder to nilotinib, 3) optimal responder to both and 4) optimal responder to neither. Each patient was treated according to the results of the test assuming the performance of the test achieves a sensitivity and specificity approximating 100%. If a patient was predicted to respond to both dasatinib and nilotinib equally, the patient received dasatinib. Alternative treatments in CML, such as allogeneic stem cell transplant (allo-SCT) if eligible and/or interferon-alpha plus low-dose arabinosylcytosine (IFNa+LDAC) were administered to patients identified as nonresponders to both TKIs.

**Table 1.** List of model input parameters used in the analysis.

| Input parameter | Treatment and Disease Phase | Base case estimate | Range† | Source | Reference |
|---|---|---|---|---|---|
| **Proportion of responders** | Dasatinib in second-line | 52% | 0.42 - 0.63 | Kantarjian and Pasquini et al. 2007 | [27] |
| | Nilotinib in second-line | 48% | 0.42 - 0.54 | Kantarjian and Giles et al. 2007 | [28] |
| | Dasatinib in third-line | 26% | 0.08 - 0.44 | Quintas-Cardamas et al. 2007 | [32] |
| | Nilotinib in third-line | 31% | 0.16 - 0.49 | Giles et al. 2006 | [31] |
| | Optimal responder to D, CP | 14% | NA‡ | See Figure 2. | |
| | Optimal responder to N, CP | 15% | NA‡ | See Figure 2. | |
| | Optimal responder to either D or N, CP | 35% | NA‡ | See Figure 2. | |
| | Optimal responder to neither D or N, CP | 36% | NA‡ | Assumed | |
| **Progression-free survival** | First year second-line dasatinib or nilotinib, responders | 98% | 0.94 - 1.0 | Kantarjian and Pasquini et al. 2007 | [27] |
| | First year second-line dasatinib, nonresponders | 93% | 0.86 - 1.0 | Kantarjian and Pasquini et al. 2007 | [27] |
| | Second year second-line dasatinib or nilotinib, responders | 97% | 0.93 - 1.0 | Tam et al. 2008 | [33] |
| | Second year third-line nilotinib, responders and nonresponders | 73% | 0.56 - 0.9 | Giles et al. 2006 | [31] |
| | First and second year allogeneic SCT | 82% | 0.80-0.84 | Gratwohl et al 2006 | [34] |
| | Incidence of progression year 1 | 9% | | Gratwohl et al 2006 | [34] |
| | Incidence of progression year 2 | 10% | | Gratwohl et al 2006 | [34] |
| | First and second year INFα+LDAC | 90% | 0.87-0.93 | O'Brien et al. 2003 | [35] |
| | Incidence of progression year 1 | 7% | | O'Brien et al. 2003 | [35] |
| | Incidence of progression year 2 | 3% | | O'Brien et al. 2003 | [35] |
| **Quality of life** | Utility in the chronic phase | 85% | 0.84 - 0.86 | Reed et al. 2003 | [14] |
| | Utility in progression from chronic phase | 60% | 0.45 - 0.75 | Reed et al. 2003 | [14] |

**Table 1.** (Continued)

| Input parameter | Treatment and Disease Phase | Base case estimate | Range† | Source | Reference |
|---|---|---|---|---|---|
| **Probability of receiving alternative treatments among nonresponders** | Allogeneic SCT | 50% | 0.4-0.6 | Based on expert clinical opinion | |
| | INFα+LDAC | 50% | 0.4-0.6 | Based on expert clinical opinion | |
| **Costs of TKI treatment** | 1 year of dasatinib treatment | € 48,321 | 33,825 - 62,817 | Dutch Healthcare Insurance Board (CVZ) | [36] |
| | 1 year of nilotinib treatment | € 50,738 | 35,516 - 65,959 | Dutch Healthcare Insurance Board (CVZ) | [36] |
| **Costs of treating non-responders** | Allogeneic SCT | € 92,056 | 64,439 - 119,673 | Dutch Healthcare Authority (NZa) | [37] |
| | Year of INFα+LDAC | € 19,758 | 13,831 - 25,686 | Groot et al. 2003 | [12] |
| **Costs of diagnostic testing** | FISH testing, per unit cost including specialist fee | € 313 | 0 - 626 | Dutch Healthcare Authority (NZa) | [37] |
| | Companion diagnostic | € 3,036 | 0 - 6,072 | Ross et al. 2008 | [38 |
| **Median time-to-response** | Responders to TKIs | 6 months | 1 - 12 | Kantarjian and Pasquini et al. 2007 | [27] |
| **Discount rate** | Effectiveness | 1.5% | 0 - 0.035 | Dutch Healthcare Insurance Board (CVZ) | [29] |
| | Costs | 4% | 0 - 0.035 | Dutch Healthcare Insurance Board (CVZ) | [29] |

*The 95% CI was selected for the range varied in the sensitivity analysis. For estimates where this was not available, the 95% CI was calculated using the standard error. †Assessment of these estimates in the sensitivity analysis was conducted by varying the estimates from which they were derived. FISH: fluorescence in situ hybridization; TKI: tyrosine kinase inhibitor; NA: not applicable.

### *Model inputs*

Table 1 describes all model input parameters used in the analysis. Clinical trial results were used to estimate the proportion of patients who will and will not respond to dasatinib and nilotinib (Figure 2). To derive the response rate of 15% for patients responding only to nilotinib, the nonresponse rate of 48% reported for second-line dasatinib by Kantarjian et al. (2007) was multiplied by the third-line response rate of 31% for nilotinib patients who failed second-line dasatinib reported by Giles *et al.* (2006) [27, 31]. To derive the response rate of 14% for patients responding only to dasatinib, the nonresponse rate of 52% reported for second-line nilotinib by Kantarjian et al. (2007) was multiplied by the third-line response rate of 26% for dasatinib patients who failed second-line nilotinib reported by Quintas-Cardamas et al. (2007) [28, 32]. The estimate

of 36% for the proportion of patients not responding to TKI therapy was based on an average of the estimated proportion not responding to dasatinib or nilotinib in second-line and third-line (i.e., 38.5% and 33%, respectively). Figure 3 illustrates the proportions responding to dasatinib, nilotinib or both, with the remaining proportion of patients (i.e., 35%) responding to neither drug.

Progression-free life-years (PFLYs) were calculated using PFS estimates from the literature, which were applied for the proportion of responders and nonresponders separately when available. Yearly PFS for second- and third-line TKI treatments was available from

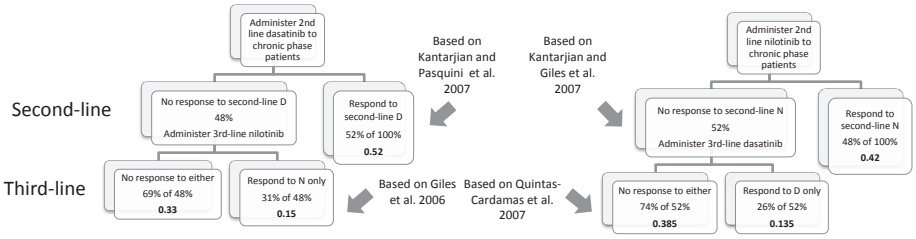



**Figure 2.** Estimates available from a literature review to calculate the proportion of responders and nonresponders to dasatinib and nilotinib.
†Based on Kantarjian *et al*. [27]. ‡Based on Giles *et al.* [31]. §Based on Kantarjian *et al.* [28]. ¶ Based on Quintas-Cardama *et al.* [32]. D: dasatinib; N: nilotinib

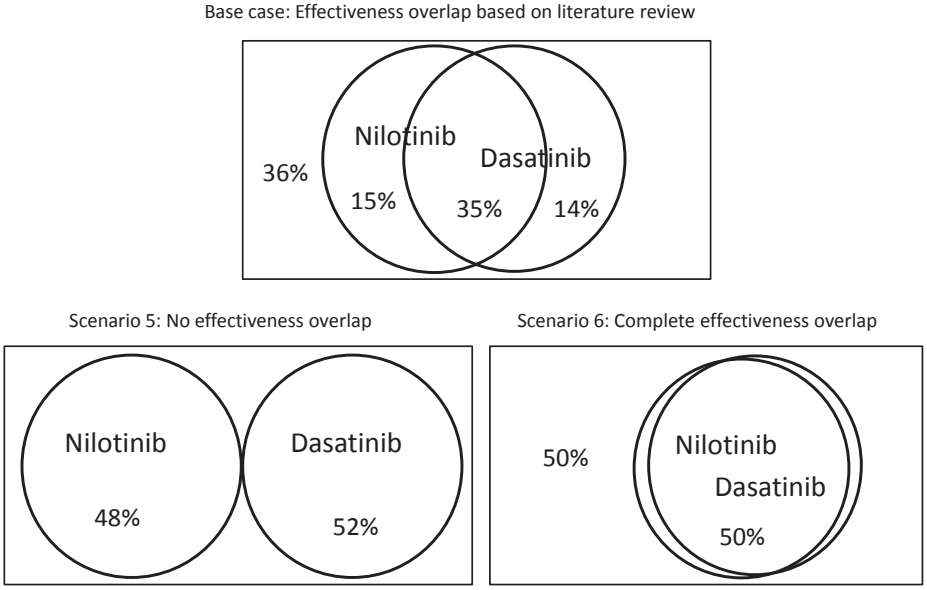



**Figure 3.** Comparative effectiveness assumed in the base case analysis and scenario analyses 5 and 6.

clinical trials [27, 31, 33]. The two-year PFS estimate for patients receiving IFNa+LDAC or allo-SCT were taken from separate sources in the literature [34, 35].

Quality-adjusted life-years (QALYs) were also reported. Utility measures for CML only differ when a patient progresses to the next phase and were available from a phase 3 study assessing imatinib [14]. A utility measure of 0.854 for the chronic phase was applied to the proportion of the progression-free population, and a utility of 0.595 for being in the subsequent disease phase was applied to the proportion progressing.

Unit costs for dasatinib and nilotinib were taken from the Dutch Healthcare Insurance Board of the Netherlands [36]. The cost of allo-SCT and FISH testing were taken from the Dutch Healthcare Authority [37]. The cost of the companion diagnostic was assumed to be €3,026 based on the cost estimate for a companion diagnostic for breast cancer [38]. Costs of adverse events were not included in the study since the safety profiles do not generally differ between dasatinib and nilotinib [39]. Cost of FISH testing among responders was applied monthly for six months, which was based on the median time until dasatinib failure reported by in the literature [27], and then every three months the remainder of the year. FISH testing among nonresponders was assumed monthly for the entire year. Unit prices were from the year 2008 and inflated to represent the cost year 2009 [40].

**Assessment of model uncertainty**

Univariate sensitivity analyses were conducted to assess the impact of model inputs on the incremental estimates. Ranges for probability and utility estimates were based on the 95% confidence interval. Clinical opinion was elicited for a range of possible estimates for the proportion of patients eligible for allo-SCT in second-line. Costs of treatment were varied between plus and minus thirty percent. The cost of FISH and companion diagnostic testing were varied between zero and twice the base case estimate. Time-to-response was varied between one and 12 months. Two-way sensitivity analyses were conducted for the discount rates for costs and effectiveness.

Six scenarios were designed and compared with the base case to quantify the impact of key assumptions on the incremental estimates. Since the base case analysis assumed that patients who progressed lived progression-free up until the end the year, four scenario analyses were designed to compare different median TTP for responders and nonresponders to second-line therapy in the second year, assuming an estimate for TTP of 11 months in the first year. The combinations of median TTP reflected in the four scenarios are based on the possibility that there is either no relationship between TTP and response status or a response advantage of longer TTP compared to nonresponders.

Scenarios five and six assessed the impact of the two most extreme examples of comparative efficiency between dasatinib and nilotinib, i.e., no effectiveness overlap and a complete effectiveness overlap (Figure 3). Scenario five assumed that 48% only respond to nilotinib and 52% only respond to dasatinib, as reported in second-line clinical trials [27, 41]. Scenario six assumed that approximately half of all patients will respond to either dasatinib or nilotinib based on the average of these estimates.

**Table 2.** Incremental costs and effectiveness results for the base case and scenario analyses.

| | | Comparator strategy | | | |
|---|---|---|---|---|---|
| | | No testing | Companion diagnostic | | |
| | Outcome | Estimate | Estimate | Incremental | ICER |
| **Base case scenario** | Costs | € 101,500 | € 89,000 | -€ 12,500 | |
| | PFLYs | 1.74 | 1.84 | 0.09 | Dominant |
| | QALYs | 1.61 | 1.63 | 0.02 | Dominant |
| **Time to progression† scenarios** | | | | | |
| Scenario 1 | Costs | € 110,400 | € 94,400 | -€ 16,000 | |
| Responders: 1 month | PFLYs | 1.77 | 1.85 | 0.08 | Dominant |
| Nonresponders: 1 month | QALYs | 1.65 | 1.67 | 0.02 | Dominant |
| Scenario 2 | Costs | € 102,800 | € 95,000 | -€ 7,800 | |
| Responders: 11 months | PFLYs | 1.78 | 1.86 | 0.09 | Dominant |
| Nonresponders: 1 months | QALYs | 1.65 | 1.67 | 0.02 | Dominant |
| Scenario 3 | Costs | € 99,000 | € 95,500 | -€ 3,500 | |
| Responders: 11 months | PFLYs | 1.83 | 1.87 | 0.05 | Dominant |
| Nonresponders: 6 months | QALYs | 1.66 | 1.68 | 0.01 | Dominant |
| Scenario 4 | Costs | € 95,200 | € 95,900 | € 700 | |
| Responders: 11 months | PFLYs | 1.88 | 1.88 | 0.01 | € 128,474 |
| Nonresponders: 11 months | QALYs | 1.68 | 1.68 | 0.00 | € 496,038 |
| **Comparative effectiveness scenarios** | | | | | |
| Scenario 5 | Costs | € 101,300 | € 98,900 | -€ 2,400 | |
| No effectiveness overlap: | PFLYs | 1.70 | 1.85 | 0.15 | Dominant |
| 52% respond to D and 48% respond to N | QALYs | 1.56 | 1.60 | 0.04 | Dominant |
| Scenario 6 | Costs | € 57,700 | € 46,400 | -€ 11,300 | |
| Complete effectiveness overlap: | PFLYs | 1.70 | 1.77 | 0.07 | Dominant |
| 50% respond to both and 50% respond to neither | QALYs | 1.56 | 1.58 | 0.02 | Dominant |

All costs are rounded to the nearest 100 euros. *Varied for 2nd line and set equal to 11 months in 1st line for both responders and non-responders with either dasatinib (D) or nilotinib (N). PFLYs: progression-free life-years; QALYs: quality-adjusted life-years; ICER: incremental cost-effectiveness ratio; ICERs shaded in grey represent scenarios of dominance for the companion diagnostic scenario.

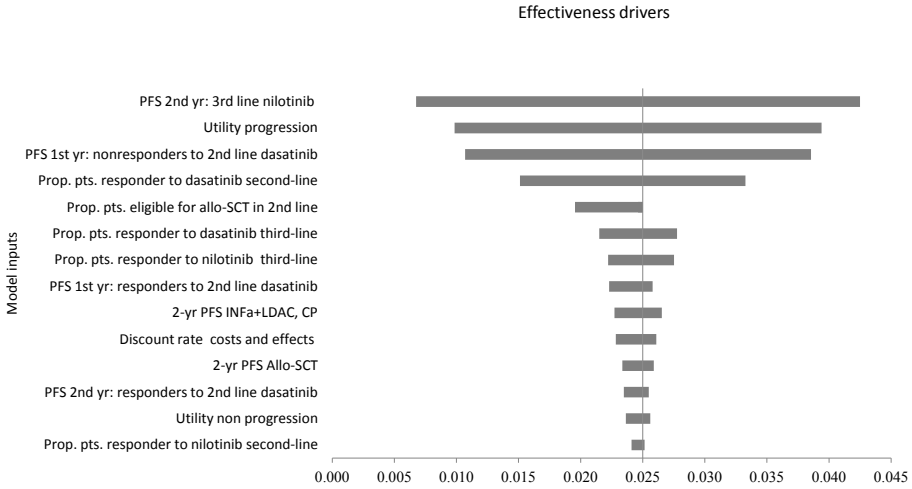



**Figure 4.** Tornado diagram depicting effectiveness drivers among the model inputs.
allo-SCT: allogeneic stem cell transplant; CP: chronic phase; IFNa+LDAC: interferon-alpha plus low-dose arabinosylcytosine; PFS: progression-free survival; Prop. pts: proportion patients; yr: year; QALYS: quality-adjusted life-years.

## RESULTS

Results are presented in Table 2. In the base case, the companion diagnostic strategy dominates the no testing strategy as it improves health outcomes and reduces costs. At a cost of €89,000, the companion diagnostic strategy results in 1.84 PFLYs or 1.63 QALYs, which provides greater health gains for the cost compared to 1.74 PFLYs or 1.61 QALYs at a cost of €101,500 with the no testing strategy.

Figures 4 and 5 illustrate the main effectiveness and cost drivers, respectively, among the model inputs. The incremental effectiveness estimate was most sensitive to the PFS estimate for patients treated with third-line nilotinib, the utility of progressive disease and the PFS for nonresponders to second-line dasatinib therapy. The incremental cost estimate was most sensitive to the cost of dasatinib treatment and the PFS estimate for patients treated with third-line nilotinib.

Scenarios one through four show that health gains and cost-savings for the companion diagnostic strategy are maximized when TTP is short for all patients in year 2 (Table 2). Scenarios two through four suggest that TTP for nonresponders may indeed impact the incremental costs and effectiveness of the companion diagnostic strategy. Scenarios five and six show that the larger the effectiveness overlap, the greater the cost-savings since the cheaper alternative of the two, dasatinib, can be administered (Table 2). Alternatively, health outcomes are maximized with no effectiveness overlap.

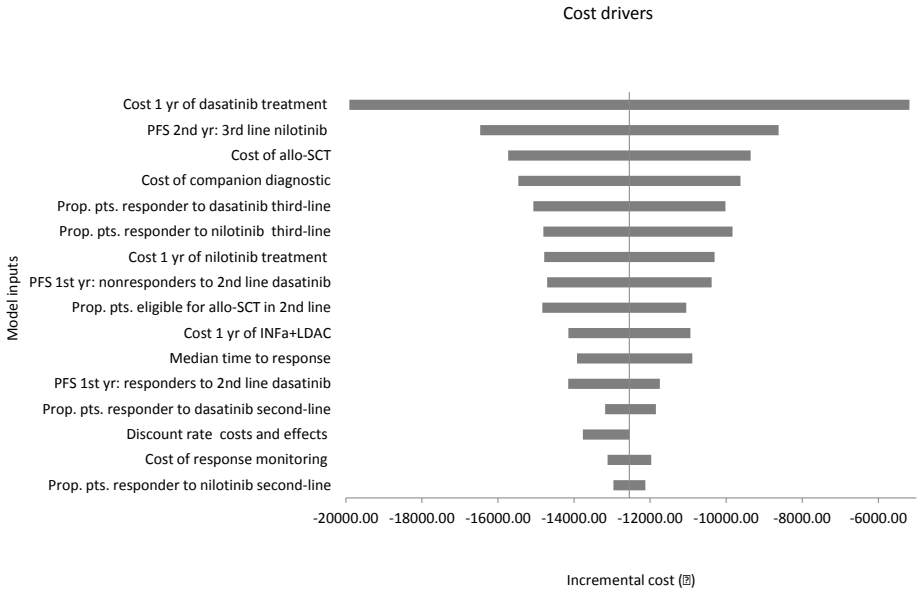



**Figure 5.** Tornado diagram depicting cost drivers among the model inputs.
allo-SCT: allogeneic stem cell transplant; IFNa+LDAC: interferon-alpha plus low-dose arabinosylcytosine; yr: year.

## DISCUSSION

A companion diagnostic strategy in CML offers the potential to improve both the effectiveness and costs of second-line treatment at a time horizon of 2 years. Patients are treated with the most effective TKI or diverted to more effective alternative treatments at an earlier moment in treatment than currently implemented in usual care.

The value of companion diagnostics in CML is still uncertain due to the limited clinical evidence for second-line TKIs. The value is strongly influenced by two factors: the difference in health outcomes between responders and nonresponders and the correlation between responsiveness to dasatinib and responsiveness to nilotinib. The clinical utility of this companion diagnostic is dependent on not only the effectiveness of targeted therapies but also their comparative effectiveness. This is a crucial level of uncertainty since the societal benefit of personalized medicine has been found to largely depend on the quality of the evidence demonstrating its clinical utility [42-44].

The sensitivity and specificity assumed in the analysis are not based on that of another test but assumed due to lack of evidence. Imperfect sensitivity and specificity would obviously result in false negatives (i.e., patients missing out on the benefits of TKI treat-

ment) and false positives (i.e., patients receiving TKI despite no benefits and unnecessary costs). The main aim in this early phase of development was to explore the potential of a companion diagnostic under ideal diagnostic performance. The rationale was that if a companion diagnostic is not attractive at that level of performance, then it will not be viewed as a test with much potential value. Therefore, in this phase, we did not examine the impact of sensitivity and specificity in the analysis. This can, however, be included in the model and we agree this is a necessary next step once more information is available regarding both the performance and the impact of false results.

The calculation of the test price as a pre hoc objective was not included in the analysis because cost-effectiveness is not considered when determining price. The price should instead reflect its market value. Because the market value is still unknown, the price assumed was chosen to represent the upper range of possible prices for currently available companion diagnostics. If a more conservative estimate were assumed, the cost-savings with the companion diagnostic strategy would improve at most by a factor of €3,000. Therefore, the price of the companion diagnostic can range between €0 and €15,500 to retain its cost-savings assuming all patients are tested. Sensitivity analysis to assess the impact of price at an early stage of technology development may be useful if an estimate of the price is available though uncertainty regarding the population size for which its true market value is based.

Further research should focus on obtaining better estimates of PFS and TTP by response status to demonstrate robust incremental estimates. To speculate whether the trend of cost-savings and health gains will continue beyond two years, extending the analysis to consider a life-time horizon will provide more valuable information. It is also worth exploring whether an adjustment for quality-of-life differences among responders and nonresponders is necessary and if the average QALYs gained actually differ by response status.

Future changes to the approved placement of TKIs within the treatment paradigm of CML are also relevant for future analyses. For example, recent evidence lead the United States Federal Drug Administration (FDA) to approve nilotinib and dasatinib for first-line therapy [45, 46]. Future analyses from the Dutch healthcare perspective may require an assessment of the cost-effectiveness of companion diagnostics for CML in first-line. As imatinib is expected to be the first of the three TKIs to go off patent, the impact of cheaper generics on the cost-effectiveness of a companion diagnostic in first-line CML treatment should also be considered. Evaluation of these issues may be necessary for all companion diagnostics. Further, whether or not the results of the test will impact clinical decision-making is an additional concern for companion diagnostics. We recommend

that further assessment of the potential value of companion diagnostics in CML include the results of a survey which examines the probability that physicians will treat according to the results of a companion diagnostic.

Despite these limitations, economic models evaluating all emerging diagnostics are helpful for examining whether systemic changes in healthcare are required for their adoption [47, 48]. They can be used at any phase in development to analyze the circumstances a personalized treatment strategy based on companion diagnostic testing must fulfill to be favorable. Their results will stimulate key stakeholders to discuss the circumstances under which a particular companion diagnostic makes sense in economic terms. Furthermore, they provide the kind of evidence required by payers, such as comparative information on new tests versus usual care and the impact of their predictive value [49].

## CONCLUSION

We present the use of CEA in assessing the potential value for companion diagnostics in CML. Such methods can be used to complement other standard methods applied at early stages of technology development for decision-making about investment in research and development. Given the barriers of physician awareness and information available for the clinical benefit of companion diagnostics, the results of this analysis can also be useful for communicating the potential value of this technology to various stakeholders.

## FUTURE PERSPECTIVE

Cost-effectiveness analyses may be required in future reimbursement decisions for companion diagnostics [43]. Incentives from regulating bodies will likely increase the number of cost-effectiveness analyses conducted in companion diagnostics. The promise of funding specific to the health technology assessment of diagnostics in the UK and increased federal support in the US for comparative effectiveness research may result in higher standards for quantity and quality of their clinical and economic evidence [50, 51]. The European Commission has also identified cost-effectiveness research among the gaps in current evidence needed to ensure effective policy on personalized medicine [52]. As the knowledge base within the field of biomarkers expands, the economic assessment of companion diagnostics may become more relevant but may require more sophisticated modeling methods.

## ACKNOWLEDGEMENTS

The authors would like to thank Rinie van Beuningen for his helpful comments while performing this study.

## REFERENCES

1. Papadopoulos N, Kinzler KW, Vogelstein B. The role of companion diagnostics in the development and use of mutation-targeted cancer therapies. Nat Biotechnol 2006; 24(8): 985-95.
2. Deverka PA, Doksum T, Carlson RJ. Integrating molecular medicine into the US health-care system: opportunities, barriers, and policy challenges. Clin Pharmacol Ther 2007; 82(4): 427-34.
3. Gurwitz D, Zika E, Hopkins MM, et al. Pharmacogenetics in Europe: barriers and opportunities. Public Health Genomics 2009; 12(3): 134-41.
4. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. Third ed. Oxford: Oxford Medical Publications; 2005.
5. Deverka PA, Vernon J, McLeod HL. Economic opportunities and challenges for pharmacogenomics. Annu Rev Pharmacol Toxicol 2010; 50: 423-37.
6. Hornberger J, Cosler LE, Lyman GH. Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. Am J Manag Care 2005; 11(5): 313-24.
7. Phillips KA, Van Bebber SL. A systematic review of cost-effectiveness analyses of pharmacogenomic interventions. Pharmacogenomics 2004; 5(8): 1139-49.
8. Vegter S, Boersma C, Rozenbaum M, et al. Pharmacoeconomic evaluations of pharmacogenetic and genomic screening programmes: a systematic review on content and adherence to guidelines. Pharmacoeconomics 2008; 26(7): 569-87.
9. Deininger MWN, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. Blood 2000; 96(10): 3343-56.
10. Druker BJ, Guilhot F, O'Brien SG, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med 2006; 355(23): 2408-17.
11. Gordois A, Scuffham P, Warren E, Ward S. Cost-utility analysis of imatinib mesilate for the treatment of advanced stage chronic myeloid leukaemia. Br J Cancer 2003; 89(4): 634-40.
12. Groot MT, Ossenkoppele GJ, Kramer MH, et al. [Drug treatment of chronic myeloid leukemia: a cost effectiveness analysis of first- and second-line treatment ]. Ned Tijdschr Geneeskd 2003; 147(30): 1460-5.
13. Dalziel K, Round A, Stein K, et al. Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis. Health Technol Assess 2004; 8(28): iii, 1-120.
14. Reed SD, Anstrom KJ, Ludmer JA, et al. Cost-effectiveness of imatinib versus interferon-alpha plus low-dose cytarabine for patients with newly diagnosed chronic-phase chronic myeloid leukemia. Cancer 2004; 101(11): 2574-83.
15. Warren E, Ward S, Gordois A, Scuffham P. Cost-utility analysis of imatinib mesylate for the treatment of chronic myelogenous leukemia in the chronic phase. Clin Ther 2004; 26(11): 1924-33.
16. Skrepnek GH, Ballard EE. Cost-efficacy of imatinib versus allogeneic bone marrow transplantation with a matched unrelated donor in the treatment of chronic myelogenous leukemia: a decision-

analytic approach. Pharmacotherapy 2005; 25(3): 325-34.
17. O'Dwyer ME, Mauro MJ, Blasdel C, et al. Clonal evolution and lack of cytogenetic response are adverse prognostic factors for hematologic relapse of chronic phase CML patients treated with imatinib mesylate. Blood 2004; 103(2): 451-5.
18. Soverini S, Martinelli G, Rosti G, et al. ABL mutations in late chronic phase chronic myeloid leukemia patients with up-front cytogenetic resistance to imatinib are associated with a greater likelihood of progression to blast crisis and shorter survival: a study by the GIMEMA Working Party on Chronic Myeloid Leukemia. J Clin Oncol 2005; 23(18): 4100-9.
19. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 2001; 293(5531): 876-80.
20. National Institute for Health and Clinical Evidence (NICE). Evidence to support chronic myeloid leukaemia drugs "very poor". . [cited 18 April 2010]; Available from: www.nice.org.uk/newsroom/news/evidencetosupportchronicmyeloidleukaemiadrugs.jsp
21. Kantarjian HM, Talpaz M, Giles F, et al. New insights into the pathophysiology of chronic myeloid leukemia and imatinib resistance. Ann Intern Med 2006; 145(12): 913-23.
22. Kantarjian HM, Smith TL, O'Brien S, et al. Prolonged survival in chronic myelogenous leukemia after cytogenetic response to interferon-alpha therapy. The Leukemia Service. Ann Intern Med 1995; 122(4): 254-61.
23. Kantarjian H, Schiffer C, Jones D, Cortes J. Monitoring the response and course of chronic myeloid leukemia in the modern era of BCR-ABL tyrosine kinase inhibitors: practical advice on the use and interpretation of monitoring methods. Blood 2008; 111(4): 1774-80.
24. Boender P, Janssen J, Smit L. Application of Kinase Activity Profiles to Predict Upcoming TKI Resistance In CML-Patients abstract 31077 at the American Society of Hematology 2010.
25. Folkvord S, Flatmark K, Dueland S, et al. Prediction of response to preoperative chemoradiotherapy in rectal cancer by multiplex kinase activity profiling. Int J Radiat Oncol Biol Phys 2010; 78(2): 555-62.
26. Hilhorst R, Schaake E, van Pel R, et al. Application of kinase activity profiles to predict response to erlotinib in a neoadjuvant setting in early stage non-small cell lung cancer (NSCLC). J Clin Oncol (Meeting Abstracts) 2010; 28(15_suppl): 10566-.
27. Kantarjian H, Pasquini R, Hamerschlak N, et al. Dasatinib or high-dose imatinib for chronic-phase chronic myeloid leukemia after failure of first-line imatinib: a randomized phase 2 trial. Blood 2007; 109(12): 5143-50.
28. Kantarjian HM, Giles F, Gattermann N, et al. Nilotinib (formerly AMN107), a highly selective BCR-ABL tyrosine kinase inhibitor, is effective in patients with Philadelphia chromosome-positive chronic myelogenous leukemia in chronic phase following imatinib resistance and intolerance. Blood 2007; 110(10): 3540-6.
29. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Guidelines for Pharmacoeconomic Research [Richtlijnen voor farmaco-economisch onderzoek, geactualiseerde versie] Diemen: CVZ, March 2006 [cited 2010 December 1]; Available from: http://www.cvz.nl/binaries/live/cvzinternet/hst_content/nl/documenten/rubriek+zorgpakket/cfh/richtlijnen+farmaco-economisch+onderzoek.pdf
30. Lowenberg B, Ossenkoppele GJ, de Witte T, Boogaerts MA. Handboek: Hematologie. Utrecht: De Tijdstroom; 2008.
31. Giles F, le Coutre P, Bhalla K, et al. A Phase II Study of Nilotinib, a Novel Tyrosine Kinase Inhibitor Administered to Patients with Imatinib Resistant or Intolerant Chronic Myelogenous Leukemia (CML) in Chronic Phase (CP), Accelerated Phase (AP) or Blast Crisis (BC) Who Have Also Failed

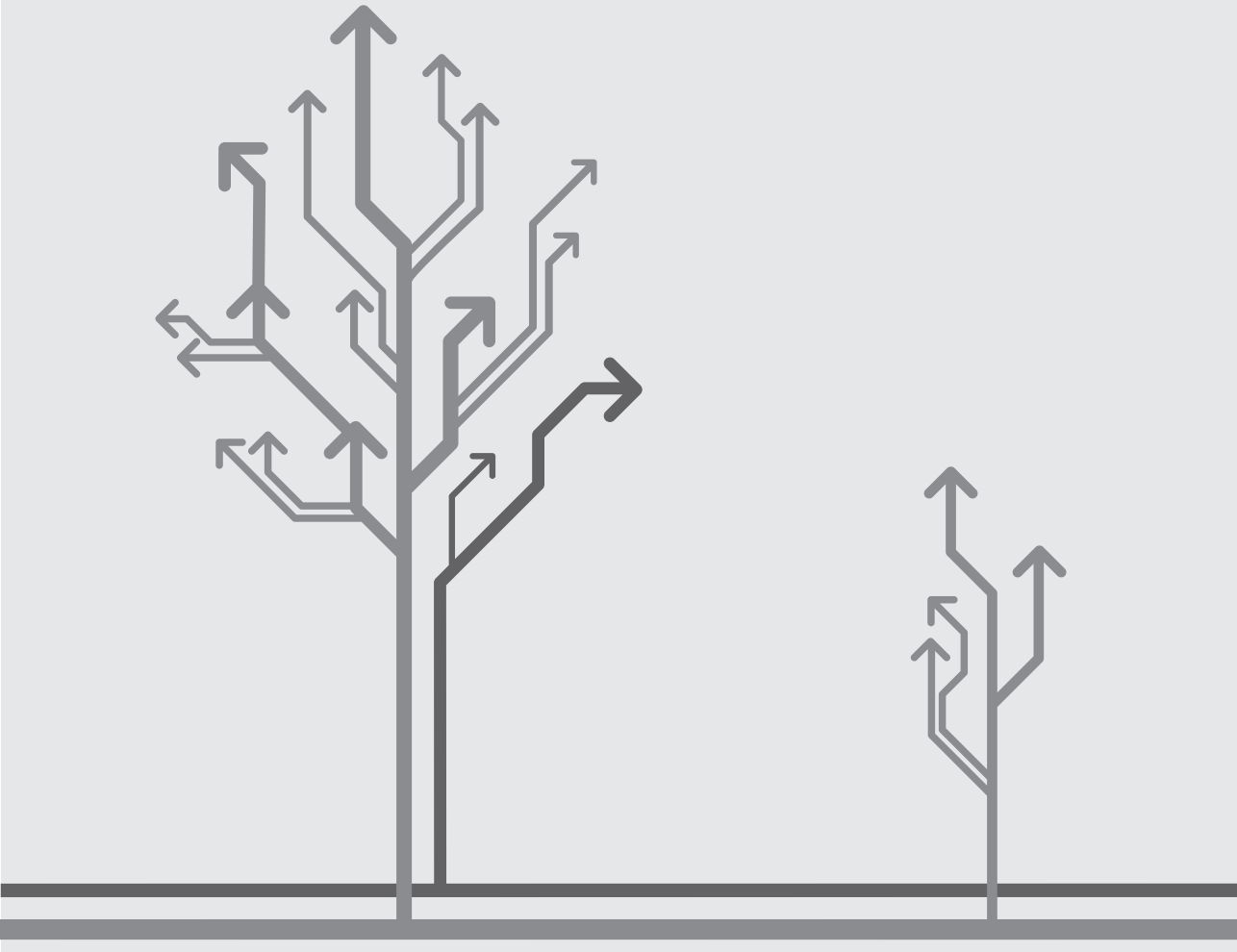
Dasatinib Therapy. ASH Annual Meeting Abstracts 2006; 108(11): 2170-.
32. Quintas-Cardama A, Kantarjian H, Jones D, et al. Dasatinib (BMS-354825) is active in Philadelphia chromosome-positive chronic myelogenous leukemia after imatinib and nilotinib (AMN107) therapy failure. Blood 2007; 109(2): 497-9.
33. Tam CS, Kantarjian H, Garcia-Manero G, et al. Failure to achieve a major cytogenetic response by 12 months defines inadequate response in patients receiving nilotinib or dasatinib as second or subsequent line therapy for chronic myeloid leukemia. Blood 2008; 112(3): 516-8.
34. Gratwohl A, Brand R, Apperley J, et al. Allogeneic hematopoietic stem cell transplantation for chronic myeloid leukemia in Europe 2006: transplant activity, long-term data and current results. An analysis by the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation (EBMT). Haematologica 2006; 91(4): 513-21.
35. O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med 2003; 348(11): 994-1004.
36. Dutch Health Care Insurance Board [College voor Zorgverzekeringen (CVZ)]. Farmacotherapeutisch Kompas. http://www.fk.cvz.nl/
37. Dutch Healthcare Authority [Nederlandse Zorgautoriteit (NZA)]. DBC healthcare product tariff application [DBC zorgproducten tariefapplicatie]. http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/Welkom.aspx
38. Ross JS, Hatzis C, Symmans WF, et al. Commercialized multigene predictors of clinical outcome for breast cancer. Oncologist 2008; 13(5): 477-93.
39. Rix U, Hantschel O, Durnberger G, et al. Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. Blood 2007; 110(12): 4055-63.
40. Central Statistical Office of the Netherlands [Centraal Bureau voor de Statistiek (CBS)]. Statline. http://statline.cbs.nl
41. Kantarjian H, Giles F, Wunderle L, et al. Nilotinib in imatinib-resistant CML and Philadelphia chromosome-positive ALL. N Engl J Med 2006; 354(24): 2542-51.
42. Flowers CR, Veenstra D. The role of cost-effectiveness analysis in the era of pharmacogenomics. Pharmacoeconomics 2004; 22(8): 481-93.
43. Deverka PA. Pharmacogenomics, evidence, and the role of payers. Public Health Genomics 2009; 12(3): 149-57.
44. Garber AM, Tunis SR. Does comparative-effectiveness research threaten personalized medicine? N Engl J Med 2009; 360(19): 1925-7.
45. Saglio G, Kim DW, Issaragrisil S, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. N Engl J Med 2010; 362(24): 2251-9.
46. Kantarjian H, Shah NP, Hochhaus A, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med 2010; 362(24): 2260-70.
47. Stallings SC, Huse D, Finkelstein SN, et al. A framework to evaluate the economic impact of pharmacogenomics. Pharmacogenomics 2006; 7(6): 853-62.
48. Vallejo-Torres L, Steuten LM, Buxton MJ, et al. Integrating health economics modeling in the product development cycle of medical devices: a Bayesian approach. Int J Technol Assess Health Care 2008; 24(4): 459-64.
49. Epstein RS, Frueh FW, Geren D, et al. Payer perspectives on pharmacogenomics testing and drug development. Pharmacogenomics 2009; 10(1): 149-51.
50. UK Secretary of State for Health. Government Response to the House of Lords: Science and Technol-

ogy Committee Inquiry into Genomic Medicine. December 2009 [cited 2010 December 1]; Available from: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_110005.pdf

51. US Department of Health and Human Services. Federal Coordinating Council for Comparative Effectiveness Research: Report to the President and Congress. June 2009 [cited 2010 December 1]; Available from: www.hhs.gov/recovery/programs/cer/cerannualrpt.pdf
52. Gaisser S, Vignola-Gagne E, Husing B, et al. EU policies in personalized medicine-related technologies. Personalized Medicine 2009; 6(1): 93-102.

# Chapter 9

## Discussion



## INTRODUCTION

For the decades prior to the start of the research conducted in this thesis, multiple myeloma (MM) was a disease treated primarily with conventional therapies. Chapters 2 and 3 demonstrated that the quantity and quality of existing economic evaluations in MM were limited despite the introduction of a number of novel therapies that have contributed to an extended survival at significantly higher costs. To fill the gaps in knowledge about the health economics of MM care, this thesis addressed a number of questions regarding the health benefits and costs of current care.

The following section of this thesis first discusses the feasibility of using outcomes research to assess the appropriate use, effectiveness and cost-effectiveness of novel therapies for MM in Dutch daily practice. The extent to which a risk-stratified treatment approach could improve patient prognosis and the cost-effectiveness of current MM care is then discussed followed by a reflection on the feasibility and usefulness of early-stage economic evaluations for hypothetical treatment approaches. Finally, some observations are then discussed regarding the extent to which there are similar challenges to performing early-stage economic evaluations and outcomes research in MM and some recommendations for initiatives to meet the objectives of both are then provided.

## FEASIBILITY OF OUTCOMES RESEARCH IN MULTIPLE MYELOMA

Part 2 of this thesis demonstrated that outcomes research did not clarify all questions about the value of the novel therapy bortezomib (Velcade®) for MM in Dutch daily practice. Observational data was found to be useful for assessing a novel drug's appropriate use (i.e., dosages and cycles), effectiveness in terms of overall survival, and cost-effectiveness in initially eligible daily practice patients. Its usefulness in decision-making regarding relative value compared to standard care was however limited.

### Findings and implications

#### *Real-world appropriate use, effectiveness and cost-effectiveness of bortezomib*

Using a medical chart review, it was feasible to assess the appropriate use of bortezomib in daily practice although data on effectiveness measures was limited (Chapter 4). Regarding appropriate use, detailed data was available to determine the dosages, total number of cycles and treatment combinations that occurred in daily practice. It was not evident in observational data however why one drug was chosen over the other(s). This is an important flaw of retrospective observational research in general. As for effective-

ness measures, it was feasible to assess overall survival, while progression-free survival and response rates were difficult to measure on account of commonly missing data. Treatment sequence was an influential factor in assessing overall survival due to the possibility of salvage with novel therapies following conventional therapies. Data was also often missing on safety outcomes and some important prognostic factors.

The real-world costs of relapsed/refractory treatment varied according to prognostic factors, treatment line and treatment-related resource use (Chapter 5). It was concluded therefore that cost comparisons between particular treatments regimens be differentiated according to (or corrected for) prognostic group and treatment sequence as opposed to the average lifetime costs per average patient. In general, bortezomib regimens were more costly compared to thalidomide (Thalidomid®) and lenalidomide (Revlimid®) regimens across all relapsed/refractory treatment lines. Hospital costs were especially influential during bortezomib-based regimens, demonstrating that its intravenous administration method further increased its costs burden compared to that for orally administered drugs, i.e., thalidomide and lenalidomide. This finding underscores the importance of including treatment-related costs in economic evaluations, as the inclusion of these costs leads to underestimation of total costs and budget impact and thus overestimation of cost-effectiveness.

In Chapter 6, it was shown that bortezomib treatment in Dutch daily practice achieved a mean health gain in overall survival of 29.5 months or 2.46 life-years (LYs) at a total mean cost of €81,626, resulting in a cost-effectiveness ratio (CER) in the relapsed/refractory setting of €33,181 per LY. In comparison to the results of the dossier submitted by the manufacturer, which reported a total health gain of 35.6 months or 2.92 LYs and costs of €45,720 at a CER of €15,446/LY [1], the findings from outcomes research show that the initial effectiveness and cost-effectiveness of bortezomib in Dutch daily practice was overestimated. The discrepancy is most likely attributable to the less favorable prognosis of daily practice patients compared to the pivotal trial patients and the fact that the manufacturer's analysis assumed 40% of patients would save costs of hospital visits due to receipt of bortezomib as outpatient. The CER in daily practice was further shown to be most favorable when administered at first relapse (total mean costs: €53,726; mean OS: 1.56 LYs; CER: €34,440/LY) versus second (total mean costs: €95962; mean OS: 2.58 LYs; CER: €37,194/LY) and later (total mean costs: €97,937; mean OS: 2.63 LYs; CER: €37,238/LY). It is therefore recommended that bortezomib be used as early as possible in the relapsed/refractory setting to maximize its cost-effectiveness.

The external validity of these effectiveness and cost-effectiveness estimates were however limited to few patients and primarily those with unfavorable prognosis. The

cause for this limitation was attributed to the dynamics in MM care affecting diffusion of innovation. During the years of data collection, it was found that there were indications of underutilization and residual regional differences in the first two years following its introduction into Dutch daily practice suggestive of accessibility issues [2]. Since bortezomib administration does not require specialised skills or hospital facilities, it was speculated that underutilization was the result of variation in prescription behavior and patient referral. Expertise was further speculated to be a possible reason for referring patients to other more experienced centers. In consideration of the findings along with that of our outcomes research study, it is concluded that decision-making incorporating inferences about effectiveness of novel drugs in MM requires careful consideration of the limitations to performing outcomes research during the drug's uptake period in daily practice.

To summarize, outcomes research into the appropriate use, effectiveness, and cost-effectiveness of bortezomib in advanced MM generated useful evidence on the dynamics of care in daily practice and the limitations to observational data. It was feasible to address *who* received bortezomib and *how* bortezomib was administered in daily practice. For reimbursement decisions, it was found to be useful for obtaining a CER based on real-world data. It should be emphasized however that the real-world CER for a new drug in MM may have limited external validity and should strictly be interpreted according to clinically significant prognostic subgroup and treatment sequence (when applicable) rather than as representative of the average daily practice patient.

***Real-world incremental cost-effectiveness of bortezomib***

Outcomes research was however not the preferred method to answer questions about the incremental effectiveness of bortezomib in daily practice compared to standard care (Chapter 6). With outcomes research it was not feasible to understand *why* bortezomib was administered versus alternative therapy. Consequently, it was not feasible to correct for residual confounding by indication, and thus no appropriate comparator group could be found. Hence, uncertainty remains regarding ICER for bortezomib compared to alternative therapies available in daily practice. In light of these findings, it is not recommended that decision-makers in reimbursement utilize outcomes research alone to make inferences about the relative effectiveness and cost-effectiveness of new drugs in daily practice for MM patients.

**Study limitations and recommendations for future research**

An important shortcoming to any medical chart review is the lack of documentation regarding treatment decision-making (i.e., prescribing behavior). Without such information, confounding by indication may persist. Other options for data collection could be

explored, such as a pragmatic (practical) clinical trial which are specifically designed to answer policy questions such as the studies in this thesis [3] or a population-based registry which includes all patients treated in daily practice [4]. Regarding the former, the advantages of its prospective, experimental nature are accompanied by generalizability and enhancement of compliance in daily practice [5]. However, there are drawbacks to practical clinical trials since they require significant expansion of infrastructure, health care provider support for more reliance on high-quality evidence, and a substantial increase in public and private funding [3]. Registry studies also have drawbacks, as they are not capable of resolving the issue of missing and low quality data since the data is derived from the same source as a medical chart review. Early results from the Population-based HAematological Registry for Observational Studies (PHAROS), which in the meantime was implemented to monitor the care of all Dutch hematological patients [6], show that the estimation of an incremental cost-effectiveness ratio (ICER) was similarly challenged by low patient numbers due to similar issues of treatment variation and heterogeneity [7]. Regardless of data collection method, it is important to emphasise that there is a high need to improve reporting of clinical data, particularly reasons for prescribing behavior. This requires that health care providers support the aims of outcomes research and are involved in their implementation.

To arrive at a valid ICER using outcomes research, further research is needed into the feasibility and methodology for merging the complementary aspects of randomized controlled trials (RCT) and observational data. Such research could pave the way for future evaluations of new drugs in MM that integrate observational data within the original RCT-based model used for reimbursement decision-making. The development of a model based on the original RCT trial results that is updated according to the heterogeneity observed in daily practice retains the advantages of randomization with the added benefit of generalizability. In this respect, RCT data and observational data would both be viewed as vital, complementary evidence necessary to informed decision-making about relative gains in value for money.

One essential question that remains to be addressed is whether the evidence from outcomes research sufficiently resolves decision maker uncertainty regarding value for money. It was not possible to answer this question in the present thesis due to the pilot nature of the outcomes research study of bortezomib. A number of other drugs are currently reimbursed under the coverage with evidence development policy, including bortezomib in the treatment of newly diagnosed MM patients[8]. While decisions are still awaited for the majority of these drugs, the outcomes of recent reassessment suggest that outcomes research after four years alone is not sufficient for resolving uncertainty regarding value for money. Namely, the reassessments for omalizumab (Xolair®) for

severe asthma (ICER: €35,905 per quality-adjusted life-year (QALY)) [9] and ranibizumab (Lucentis®) for macular degeneration (ICER: of €53,453/QALY) [10] have revealed insufficient data to perform a reassessment after four years of outcomes research. It is questionable therefore whether outcomes research was a worthwhile use of resources for assessing uncertainty regarding value for money at reimbursement.

Recent efforts to apply the techniques of value of information (VOI) of analysis in the context of conditional reimbursement could be applied in the future at the initial reimbursement decision to prioritize outcomes research from the start [11]. Using the techniques of VOI analysis, specific types of evidence can be prioritized at the start of the outcomes research in terms of their contribution in resolving uncertainty about a new drug's real-world effectiveness and cost-effectiveness. Based on such prioritization, an outcomes research study could be designed to collect data only on critical parameters shown to be feasible with observational data rather than collecting data on all parameters, as was the case in this thesis.

Another option is to instead implement other coverage schemes such as financial- and outcomes-based risk-sharing arrangements, similar to that implemented in England and Wales for bortezomib [12]. Future reimbursement policies in the Netherlands may indeed incorporate other options. For example, the Dutch minister has recently proposed the implementation of risk-sharing arrangements from 2013 onwards [13]. A unique arrangement from the Dutch context was recently advised by the Dutch Healthcare Insurance Board (CVZ) for omalizumab [9]. The advice to the Minister of Health was that a pay-for-performance scheme be instituted for a period of two years during which time outcomes research will continue to be conducted. No pay-for-performance scheme was suggested for ranibizumab, but instead it was recommended that reimbursement continue until 2015 during which time data would continue to be collected from daily practice [10]. After this time, the decision to reimburse ranibizumab will be revisited in consideration of all additional evidence, both from daily practice and clinical studies, which could result in a similar risk-sharing arrangement to that for omalizumab. As more drugs reach the reassessment phase, future research should examine the extent to which outcomes research resolves uncertainty about value for money.

## POTENTIAL THERAPEUTIC AND ECONOMIC VALUE OF RISK-STRATIFIED TREATMENT APPROACHES IN MULTIPLE MYELOMA

One of the objectives of part 3 was to apply the method of economic evaluation at the early-stage of product development to assess the potential value of risk-stratified treat-

ment approaches in MM. While the potential for therapeutic gains and cost-savings was demonstrated, there remained a large degree of uncertainty surrounding the results.

**Findings and implications**

Chapter 7 demonstrated that risk-stratified treatment approaches in MM incorporating existing and new molecular markers of adverse prognosis could improve health and save costs compared to a uniform treatment approach. Uncertainty surrounding the cost-effectiveness results remained, however, on account of the statistically non-significant confidence intervals for survival in some risk groups and clinically important uncertainty. The causes for non-significant findings included incomplete follow-up and low patient numbers with data on molecular markers. Despite the uncertainty, the usefulness of a modeling study was that it facilitated scenario analyses to assess the potential value of experimental treatment options.

Some unresolved issues remain with the current treatment paradigm for MM that impact the generalizability of these results. First, there is variation in the definition of high-risk prognosis and consequently the prognostic markers used to identify adverse prognosis. In other words, the prevalence of high-risk, and hence standard-risk, differs according to the definition of high-risk that is applied. This leads to variation in the diagnostic performance and hence cost-effectiveness of risk-stratified approaches across settings. Even when assuming a standardized definition for high-risk, treatment centers can further differ in their preferences for treatment of high and standard-risk patients. For example, it is widely accepted that high-risk patients detected by ISS stage, t(4;14) FISH or del(17p) FISH [14] should receive a bortezomib-based regimen [15]. However, there is no consensus on whether the treatment of standard-risk patients (i.e., patients not presenting with adverse prognostic factors) should be altered from the current standard of care, which is often bortezomib-based. This means that in settings where bortezomib remains the standard of care for all patient risk groups, the extra costs of testing required to implement a risk-stratified treatment approach may not be justified since the treatment decision will not be influenced by the test. On the other hand, a risk-stratified treatment approach for bortezomib would have more value in settings where standard-risk patients receive non-bortezomib-based regimens. It would also be valuable where an experimental therapy option exists which offers high-risk patients the chance for cure. With these issues in mind, the cost-effectiveness of a risk-stratified treatment approach in MM is setting-specific and should be assessed accordingly.

**Study limitations and recommendations for further research**

An influential limitation of the methods applied in Chapter 7 was that the observed associations between the high-risk gene signature and adverse prognosis were susceptible

to optimism bias since the data used in the model was taken from the same patients used to develop the test. Validation of the results using efficacy data from an independent source, such as a prospective stratified trial and/or daily practice, is needed. Data could be taken from, for example, the European Myeloma Network (EMN)-02 trial which enrolls newly diagnosed MM patients, which incorporates a protocol requiring testing using gene signature techniques for their validation in prediction of prognosis and resistance [16]. Moreover, the model results can also be validated using daily practice data available from the PHAROS registry [6].

Future research is needed to address the barriers to testing for molecular markers since a large number of patients had missing data on molecular markers. Missing tests were found despite the recommendation of key opinion leaders to perform testing in all patients at diagnosis[17]. It is known that the feasibility of molecular testing varies depending on the technical capability of the treatment center performing the test as well as biological feasibility (i.e., adequate bone marrow sample). If the missing data were indeed arising completely at random due to these factors, the prognosis of unknown-risk patients should have been similar to the uniform treatment approach, which reflected the average prognosis across all patients. However, the patients with missing data appeared to be a nonrandom group since the average prognosis for this group was the most favorable of all risk groups, even when compared to standard-risk patients. This suggests that there are other r easons explaining missing data for molecular merkers, which is likely to be a combination of factors[18-20]. Future studies aimed at identifying the factors affecting adoption of molecular testing specifically in the Netherlands are recommended so as to pinpoint strategies to improve their use in daily practice.

## METHODOLOGICAL CONSIDERATIONS FOR EARLY-STAGE ECONOMIC EVALUATIONS OF STRATIFIED TREATMENT

With the addition of Chapter 8 in Part 3, the objective was also to evaluate the feasibility and usefulness of early-stage economic evaluations of molecular markers in risk-stratified treatment approaches for hematological malignancies. Based on the findings of both studies, a number of methodological considerations to early-stage economic evaluations of risk-stratified medicine approaches in hematology can be formulated.

The main difference between the methods applied in Chapters 7 and 8 is the data source used to estimate effectiveness of risk-stratified treatment. The magnitude of uncertainty surrounding the effectiveness estimates would arguably be smaller in Chapter 7 compared to Chapter 8 since assumptions were made based on actual patient-level trial

data versus the literature. The uncertainty surrounding the results however remained an important issue in both studies. The reason for this was that robust evidence for the comparative effectiveness between novel therapies was difficult to find for subgroups and when available was often based on a small number of patients since data on prognostic markers were often missing. Although these early-stage economic evaluations could not provide absolute estimates for cost-effectiveness of risk-startified treatment, the development of an accurate, straightforward model was necessary for quantifying the effectiveness and cost-effectiveness of current care strategies. The current care model was important as it provided a platform for discussion among a variety of stakeholders on the needs in care and the potential clinical utility of risk-stratified treatment management in daily practice. Based on these conversations, scenario analyses were an important method to assess various assumptions about treatment implications of the new technology in terms of health impact and costs. Sensitivity analyses to assess the impact of parameter uncertainty were further helpful in identifying the influence of potentially modifiable characteristics on the probability that a risk-stratified treatment approach would be more effective and cost-effective than uniform treatment. To summarize, the primary objective of an early economic evaluation of risk-stratified treatment is to develop an accurate model of current care, and then only when feasible should it be extended to assess the potential impact of the new technology.

Since the decision to develop a molecular marker is often made under uncertainty about clinical utility, the evidence generated by economic evaluations is a useful method for developers of the test to demonstrate and monitor the uncertainty throughout the development process. The evidence can also be used to convince potential investors and collaborators to commence and/or continue support of clinical studies demonstrating clinical utility. The existence of a simple model during the preclinical phase is useful in this regard as it can be updated to generate more robust estimates of effectiveness and cost-effectiveness throughout the development phase. By updating the model at various Go/No Go decisions throughout the development phase, it can be assessed whether the uncertainty about effectiveness and cost-effectiveness is reduced as data on influential parameters becomes available. While this was not our aim, it should be feasible to apply a Bayesian approach, as demonstrated by others [21, 22], to assess the impact of additional data on the statistical uncertainty surrounding the results from the models in Chapters 7 and 8 of this thesis.

Early-stage economic evaluations can also provide useful evidence regarding the circumstances under which the molecular marker would be most promising in terms of meeting the needs of the end user. It is often the case that a number of potential uses for a new technology are considered at the preclinical development phase. Evidence

on the potential health and budget impact of the alternative can complement the criteria used by the developer to judge whether to pursue a particular indication for the test since these attributes are likely to be relevant to the end user. The techniques of multi-criteria decision analysis (MCDA), which are used to support decision-making that involves a number of competing attributes and multiple stakeholder perspectives, provide a formal way of valuing the alternative uses of the technology based on a set of attributes found to be relevant to the decision-maker. In the case of molecular tests, we will assume here that the user (i.e., clinician) of the test is the most relevant decision maker, although payers would be a relevant decision maker as well. For example where the user is the decision maker, Hilgerink et al. (2011) applied the technique of MCDA to increase the transparency of the treatment decision-making process incorporating medical devices [23]. To apply this technique, it is first important to rank the relevant attributes of the alternative uses of the marker in order of their importance to the end user. Possible attributes could include some measurement of need or disease burden, whether clinical utility has been established, diagnostic performance, effectiveness or cost. This can be done by applying the method of conjoint analysis to elicit attribute preferences from a survey of care providers [24]. A comparison of the alternative uses for the technology could then be performed based on the ranking of attributes by means of the method of analytical hierarchy process (AHP) [25]. Based on the results of the AHP analysis, the development of the test could then be steared toward the most promising indication that meets the needs and preferences of the user and thus ensuring its adoption in daily practice.

## RECOMMENDATIONS TO MEET THE OBJECTIVES OF BOTH EARLY-STAGE ECONOMIC EVALUATIONS AND OUTCOMES RESEARCH IN MULTIPLE MYELOMA

Despite methodological differences, outcomes research and early-stage economic evaluations in MM are complementary to one another. Both forms have the potential to address the needs of a variety of stakeholders, be they the developer, user (i.e., patient and/or care provider) or payer. Outcomes research during the diffusion phase of a new health technology is helpful in evaluating the effectiveness and cost-efectiveess of the current care process and to identify the needs for quality improvement. Using this knowledge about current care, early-stage economic evaluations can be then performed to guide the developer in developing a technology which aims to resolve the needs in care as identified by outcomes research and hence improves both patient prognosis and the efficiency of care. Likewise, the development an early economic evaluation that is based on real-world data will be useful for both the developer and the payer at the

market access stage. An existing model benefits the developer as it would avoid the time investment required to perform the analysis and it would provide the payer with more certain results about the effectiveness and cost-effectiveness of the technology in daily practice.

It was demonstrated in this thesis, however, that there are challenges to meeting the objectives of outcomes research and early economic evaluations in the context of MM. First, there was limited availability of robust, high quality data, with many types of data often missing (i.e., patient- and tumor-related prognostic factors, prescribing behavior, response and progression rates, patient-reported outcomes). One option to improve the quality of existing data is to emphasize to care providers the need to improve reporting of clinical data. Another option is to improve access to high quality data by setting up a novel data integration system that incorporates all clinical data and biomarker research in the Netherlands, similar to that implemented in other settings[26]. For example, a system at Duke University in the USA aggregates data from electronic patient records, tumor registry, billing, patient-reported outcome studies, and clinical trials into a single database[27]. The advantage of this system is that it integrates various data sources to improve transparency of the entire care process and align a number of objectives in health economics and outcomes research, including that of financial- and outcomes-based reimbursement schemes and the evaluation of treatment strategies incorporating biomarkers. For implementation of such an ambitious system in the Netherlands, this would require the combination of data from the two most important sources available in hematology: the PHAROS registry, which includes patient-reported outcomes and data from daily practice, and the clinical trial data collected by the Haemato Oncology Foundation for Adults in the Netherlands (HOVON) [28]. For the financial aspect, data from hospital billing records would also be required. One key issue with this option is the source of funding since these data sources are proprietary databases. The funding such a system could be organized by means of a multidisciplinary consortium since a number of stakeholders, including developers, care providers, payers, patients, and researchers, will benefit from high quality data. Obvious sources of funding would therefore include developers of health care technology in hematology, payers of health care and patient advocacy groups.

Another approach to improve the quality of evidence generated by outcomes research and early economic evaluations in MM could be to actually apply the Dutch coverage with evidence development policy to molecular tests. This was implemented by the Centers for Medicare and Medicaid Services for pharmacogenomic warfarin tests when used in trial settings [29]. Such an approach could enhance evidence generation for biomarker-based treatment strategies by encouraging their use in daily practice and

hence generate data on the potential impact of molecular testing in terms of their effectiveness and cost-effectiveness.

Drawing from the experience with outcomes research in this thesis, it is still possible however that these initiatives do not succeed in providing useful data about relative effectiveness and cost-effectiveness. The reason for this is the widespread treatment variation observed in Dutch daily practice, which makes it difficult to find comparable treatment groups to assess relative outcomes of current and future care strategies. This variation was even observed in the presence of established treatment guidelines [30]. The reasons for treatment variation are numerous, including differences in patient preferences, interpretation of the efficacy evidence and compliance to established guidelines, and policies affecting access to care. To help explain and resolve the causes for treatment variation, it would be worthwhile to conduct research into the barriers to following the treatment guidelines for both patients and care providers.

Ultimately the issues with data quality and treatment variation led to uncertainty in the results from economic evaluations of MM care. With this in mind, it would be helpful to develop a disease model for current care of MM in the Netherlands using daily practice data from the sources mentioned above (i.e., PHAROS, HOVON, and billing records). A disease model incorporating all sources of data available from RCTs and Dutch patients would generate more robust estimates of the effectiveness and cost-effectiveness of current, new and hypothetical treatment strategies in Dutch daily practice. Morevover, the disease model for MM could be used throughout the lifecycle of the technology by applying Bayesian updating and VOI analysis to help resolve the various types of uncertainty that exist about the effectiveness and cost-effectiveness of MM care.

## CONCLUDING REMARKS

The health economics of MM is complex owing primarily to the heterogeneity in patient prognosis. Future research in this area should move away from addressing efficiency-related questions in heterogeneous groups and instead focus on improving outcomes in more narrowly defined groups with homogenous prognosis, particularly by incorporating prognostic risk markers and treatment sequence. It should be emphasized that this is only possible with active interdisciplinary collaboration and the sharing of existing data among a variety of stakeholders.

Alongside providing much insight into the effectiveness and efficiency of MM care, this thesis also identified a number of methodological challenges and recommendations

for further research in conducting economic evaluations in MM. It is the hope rather that these challenges be seen as opportunities for furthering the aim to improve the prognosis of MM patients.

## REFERENCES


1. Dutch Healthcare Insurance Board [College voor Zorgverzekering (CVZ)]. Assessment of bortezomib (Velcade®) for the treatment of progressive multiple myeloma [Beoordeling bortezomib (Velcade®) voor de behandeling van progressief multipel myeloom]. Deimen, Netherlands: CVZ; September 2007.
2. Blommestein H, Franken M, Uyl-De Groot C. Accessibility and utilization of bortezomib in dutch daily practice. Internatioanl Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2010; Prague, Czech Republic: Value in Health.
3. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA : the journal of the American Medical Association. 2003;290(12):1624-32. Epub 2003/09/25.
4. Parkin DM. The evolution of the population-based cancer registry. Nat Rev Cancer. 2006;6(8):603-12.
5. Brass EP. The gap between clinical trials and clinical practice: the use of pragmatic clinical trials to inform regulatory decision making. Clinical pharmacology and therapeutics. 2010;87(3):351-5. Epub 2009/12/17.
6. Huijgens PC, Posthuma EFM, Coebergh JWW, Van de Poll-Franse LV, Uyl-de Groot CA, Sonneveld P. A population-based registry for hemato-oncology. [Een 'population-based' registry voor hemato-onologie.] Nederlands Tijdschrift voor Hematologie 2010;7:321-5.
7. Blommestein H, Franken M, Verelst S, Gaultney JG, Huijgens P, Sonneveld P, et al. Can a population-based patient registry improve the feasibility of outcomes research in multiple myeloma? Value in Health. 2012;15(7):A432.
8. Dutch Healthcare Insurance Board [College voor Zorgverzekering (CVZ)]. Conditionally Authorized Specialist Medicine, update October 2012 [Voorwaardelijk Toegelaten Specialistische Geneesmiddelen, geactualiseerd Oktober 2012]. Diemen, Netherlands: CVZ; 2012.
9. Dutch Healthcare Insurance Board [College voor Zorgverzekering (CVZ)]. Pay-for-performance agreement for omalizumab (Xolair®) for severe asthma [Pay-for-performance overeenkomst omalizumab (Xolair®) bij ernstige asthma]. Diemen, Netherlands: CVZ; August 2012.
10. Dutch Healthcare Insurance Board [College voor Zorgverzekering (CVZ)]. Reimbursement advice for ranibizumab (Lucentis®) for age-related macular degeneration [Pakketadvies ranibizumab (Lucentis®) bij leeftijdsgebonden maculadegeneratie]. Diemen, Netherlands: CVZ; August 2012.
11. Corro Ramos I, Rutten-van Molken MP, Al MJ. The Role of Value-of-Information Analysis in a Health Care Research Priority Setting: A Theoretical Case Study. Medical decision making : an international journal of the Society for Medical Decision Making. 2012. Epub 2013/01/01.
12. National Institute for Health and Clinical Excellence. Bortezomib monotherapy for relapsed multipe myeloma. London: National Institute for Health and Clinical Excellence; 2007.
13. Schippers EI. Proposals for provisional and conditional inclusion of new drugs in the drug reimbursement scheme [Voorstellen voor voorlopige en voorwaardelijke opname van nieuwe geneesmiddelen in het pakket]. Ministerial letter; 2011.

14. Munshi NC, Anderson KC, Bergsagel PL, Shaughnessy J, Palumbo A, Durie B, et al. Consensus recommendations for risk stratification in multiple myeloma: report of the International Myeloma Workshop Consensus Panel 2. Blood. 2011;117(18):4696-700. Epub 2011/02/05.
15. Bergsagel PL, Mateos MV, Gutierrez NC, Rajkumar SV, San Miguel JF. Improving overall survival and overcoming adverse prognosis in the treatment of cytogenetically high-risk multiple myeloma. Blood. 2012. Epub 2012/11/21.
16. European Myeloma Network and Nordic Myeloma Study Group. Prognostic Classification and Prediction of Clinical Resistence by Gene Signatures and Minimal Residual Disease in Multiple Myeloma - Validation of New Diagnostics in the EMN-02 Clinical Prospective Trial. February 2012. http://myeloma-europe.org.linux9.curanetserver.dk/download.php?f=05e3726e04bc75a78a84b5d8d3dd6a5c&target=0
17. Fonseca R, Bergsagel PL, Drach J, Shaughnessy J, Gutierrez N, Stewart AK, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK. 2009;23(12):2210-21. Epub 2009/10/03.
18. Frueh FW. Real-world clinical effectiveness, regulatory transparency and payer coverage: three ingredients for translating pharmacogenomics into clinical practice. Pharmacogenomics. 2010;11(5):657-60. Epub 2010/04/27.
19. Hresko A, Haga SB. Insurance coverage policies for personalized medicine. Journal of Personalized Medicine. 2012;2:201-16.
20. Deverka PA, Schully SD, Ishibe N, Carlson JJ, Freedman A, Goddard KA, et al. Stakeholder assessment of the evidence for cancer genomic tests: insights from three case studies. Genetics in medicine : official journal of the American College of Medical Genetics. 2012;14(7):656-62. Epub 2012/04/07.
21. Vallejo-Torres L, Steuten L, Parkinson B, Girling AJ, Buxton MJ. Integrating health economics into the product development cycle: a case study of absorbable pins for treating hallux valgus. Medical decision making : an international journal of the Society for Medical Decision Making. 2011;31(4):596-610. Epub 2010/12/04.
22. Vallejo-Torres L, Steuten LM, Buxton MJ, Girling AJ, Lilford RJ, Young T. Integrating health economics modeling in the product development cycle of medical devices: a Bayesian approach. International journal of technology assessment in health care. 2008;24(4):459-64. Epub 2008/10/03.
23. Hilgerink MP, Hummel MJ, Manohar S, Vaartjes SR, Ijzerman MJ. Assessment of the added value of the Twente Photoacoustic Mammoscope in breast cancer diagnosis. Med Devices (Auckl). 2011;4:107-15.
24. Herrmann A, Gustafssson A, F. H. Conjoint Analysis as an Instrument of Market Research Practice. Gustafssson A, Herrmann, A., Huber F., editor. Berlin: Springer; 2001.
25. Hummel MJ, van Rossum W, Verkerke GJ, Rakhorst G. Assessing medical technologies in development. A new paradigm of medical technology assessment. International journal of technology assessment in health care. 2000;16(4):1214-9. Epub 2001/01/13.
26. Hirsch BR, P. AA. Structure decision-making: using personalized medicine to improve the value of cancer care. Journal of Personalized Medicine. 2013;3:1-13.
27. Horvath MM, Winfield S, Evans S, Slopek S, Shang H, Ferranti J. The DEDUCE Guided Query tool: providing simplified access to clinical data for research and quality improvement. Journal of biomedical informatics. 2011;44(2):266-76. Epub 2010/12/07.
28. (HOVON) THOFfAitN. Available from: http://www.hovon.nl.
29. Center for Medicare and Medicaid Services. CMS Manual System: Pub 100-03 Medicare National

Coverage Determinations. December 2009 [cited 2012 4 February ]; Available from: http://www.cms.gov/regulations-and-guidance/guidance/transmittals/downloads/r111ncd.pdf.

30. Zweegman S, Lokhorst HM, Levin MD, de Waal E, Bos GMJ, Kersten MJ, et al. Richtlijnen behandeling multipel myeloom 2012. Nederlands Tijdschrift voor Hematologie. 2012;9(8):300-20.

# Chapter 10.

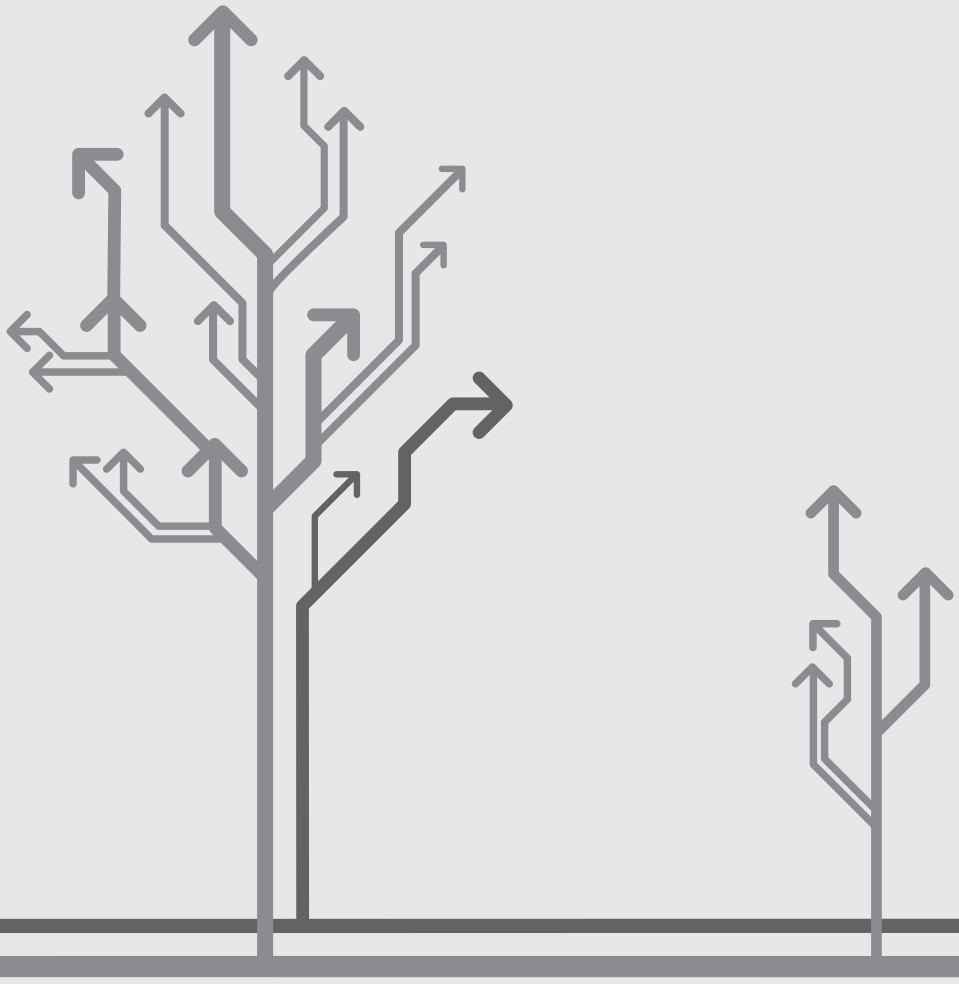
**Summary**

**Samenvatting**

**Acknowledgements**

**PhD Portfolio**

**Curriculum Vitae**



## SUMMARY

Multiple myeloma (MM) is a cancer of the plasma cells that accounts for approximately 1% of all cancers and 10% of hematological malignancies. The average age of onset is 65 years and the median survival is 4-7 years. As it is an incurable disease, patients eventually relapse from initial treatment consisting of one of more anti-cancer therapy regimens with or without stem cell transplant depending on eligibility. In the relapsed/refractory stage, patients are re-treated with a number of regimens upon signs of progressive disease until the patient eventually becomes resistant to therapy. Traditionally, MM was treated using conventional chemotherapy regimens. This changed with the introduction of the first targeted therapy thalidomide around the late 1990's followed by bortezomib and lenalidomide in the early to mid-2000's and more widespread use of stem cell transplantation. With manufacturers requesting higher unit prices for novel versus conventional therapies, the overall budget impact of the increased costs of treatment advances could be substantial. In response to uncertainty regarding the value of novel drugs, payers asked for additional information on the effectiveness and cost-effectiveness of expensive drugs in daily practice. Specifically in the Netherlands, uncertainty resulted in a coverage with evidence development policy instituted in 2008 requiring outcomes research into the appropriate use, effectiveness, and costs of expensive inpatient drugs. The drug bortezomib for the indication of relapsed/refractory MM would have qualified for such a policy. Since little knowledge was known about the feasibility and usefulness of outcomes research, a pilot outcomes research study for bortezomib was conducted.

To begin to understand the impact of treatment advances in the efficiency of MM care, it was important to first assess what was already known about the effectiveness and cost-effectiveness of the standard of care. Part 1 of this thesis assessed the state of knowledge about the efficiency of care in multiple myeloma and the quality of previous work in this area (Chapter 2), with a particular focus on novel therapies (Chapter 3). In Chapter 2, it was found that stem cell transplantation in combination with intensive chemotherapy has been demonstrated to be cost-effective, while evaluations have become less frequent in the last decade, especially for novel therapies despite their important contribution to improvements in outcomes. Furthermore the quality of the methodology applied and its documentation could be improved in many aspects. Few studies incorporated standard methods expected in high-quality economic evaluations, such as assessment of uncertainty in the results and discounting. Many existing studies relied on effectiveness inputs for the comparators that were based on non-experimental studies including different patient populations. Based on the scarcity of economic evaluation of novel therapies available in the literature, in Chapter 3 a more focused

review of the therapeutic and economic evidence available was performed specifically for novel therapies in MM. It was found that the therapeutic evidence indeed revealed improved efficacy and safety for novel therapies compared to conventional therapies. Patient-reported outcomes studies have demonstrated improvements in health-related quality of life during thalidomide maintenance therapy and relapsed/refractory patients responding to bortezomib. Evidence is still needed to demonstrate improvements during lenalidomide therapy. At the time the review presented in Chapter 3 was conducted, a total of three studies were found which assessed the cost-effectiveness of novel therapies; two of which compared bortezomib versus lenalidomide in relapsed/refractory patients by means of an indirect comparison using two separate trials. The assumptions for health benefits with novel therapies incorporated in these modeling studies were found to be influential and often overestimated the lifetime health gains. Consequently, it was concluded that existing modeling studies in MM have not been able to overcome the data limitations regarding comparative effectiveness given the conflicting results in studies based on efficacy from similar trials. It was also found that the utility values used to generate quality-adjusted life-years (QALYs) have consistently been based on an earlier study performed prior to the availability of novel therapies, suggesting a need for patient-reported outcomes studies of novel therapies.

Part 2 of this thesis examines the feasibility and implications of outcomes research in evaluating the appropriate use and effectiveness (Chapter 4), costs (Chapter 5), and incremental cost-effectiveness (Chapter 6) of bortezomib in relapsed/refractory MM patients. The results of this pilot study found that observational data was useful in assessing the appropriate use, overall survival, and costs of bortezomib in daily practice patients. However, when assessing the appropriate use and effectiveness of bortezomib in daily practice, as shown in Chapter 4, generalizability of the findings was limited to patients with unfavorable prognosis. The cohort of patients that were treated with bortezomib during its adoption phase in daily practice were of less favorable prognosis compared to those treated in the pivotal RCT used for reimbursement. Furthermore these patients were worse off compared to patients treated with bortezomib following the time period for conducting outcomes research. Chapter 5 showed that the real-world costs of relapsed/refractory treatment vary greatly with cost drivers including hospital admissions and the price of novel therapies. Increased costs were further shown to be associated with adverse prognostic factors and treatment-related side effects. Bortezomib-based regimens in particular were shown to be significantly more costly due to a higher unit price, more frequent hospital visits necessary for intravenous administration, and higher costs for concomitant medications. These results demonstrate that it is possible to develop evidence on the cost-effectiveness of a novel drug during the diffusion phase using outcomes research. However, as shown in Chapter 6, the incomparability of the

patients treated with bortezomib compared to controls treated with alternative regimens (i.e., confounding by indication) made it impossible to resolve the uncertainty regarding incremental cost-effectiveness. Based on the findings in Part 2, it was concluded that the dynamics in daily practice for MM during the diffusion period, i.e., inexperience with the drug and the already existing treatment variation, challenged the feasibility to identify a valid comparator and to generate incremental estimates. Decision-makers requesting outcomes research to make inferences about the relative effectiveness and cost-effectiveness of new drugs in daily practice for MM were warned of the influence of the diffusion period for new drug in daily practice. Instead, it was advised that decision-makers carefully consider whether outcomes research sufficiently decreases uncertainty or whether other options (e.g., financial- and outcomes-based risk-sharing arrangements) might be more appropriate to ensure sufficient value for money.

While experimental studies demonstrated improvements in survival during the novel agent era, the gains in health have not been widespread as some patients consistently have poor prognosis despite such progress. The heterogeneity in treatment effectiveness has led many to recommend a risk-stratified treatment (RST) approach, where patients are treated according to the presence of biomarkers shown to have independent associations with patient outcomes. Though yet to be demonstrated, a more personalized approach in MM could improve the effectiveness and cost-effectiveness of novel therapies compared to a uniform treatment approach. In light of this uncertainty, the objective of Part 3 was to apply the method of economic evaluation at the early-stage to assess the potential value of risk-stratified treatment (RST) approaches. The potential for health gains and cost-savings were demonstrated for adverse prognostic markers MM patients in Chapter 7, while uncertainty surrounding the results was an influential factor. Similar findings were found in Chapter 8 when applying the same method at an even earlier phase of technology development for hypothetical markers to predict upfront treatment response in chronic myeloid leukemia. Both studies highlighted a number of methodological considerations when performing early-stage economic evaluations of RST approaches in hematology. Namely, robust evidence for comparative effectiveness between novel therapies remains difficult to find and limited to subgroups with small patient numbers. Early-stage economic modeling was however useful since the development of a simple, straightforward model facilitated the quantification of the costs and effectiveness of current care strategies. It also provided a platform for discussion with stakeholders on the needs in care and the possible implications the technology could have on treatment decision making in daily practice. By incorporating possible care scenarios into the model, scenario analyses could be performed to assess various assumptions about effectiveness of the new technology and identify at what cost the new technology may negate its economic argument.

In the final chapter of this thesis (Chapter 9), a detailed discussion of the implications of these findings for policy makers, care providers, and researchers is provided. Some observations are then discussed regarding the extent to which there are similar challenges to performing early-stage economic evaluations and outcomes research in MM. This thesis concludes with recommendations for initiatives to meet the objectives of both early-stage and classical economic evaluations in MM.

## SAMENVATTING

Het multipel myeloom (MM) is een vorm van bloedkanker die zich op oudere leeftijd manifesteert. MM beslaat 10% van de vormen van bloedkanker en ongeveer 1% van alle soorten kanker. De gemiddelde leeftijd bij aanvang van de ziekte is 65 jaar en de gemiddelde overleving 4 tot 7 jaar. MM is een ongeneeslijke ziekte waardoor alle patiënten ondanks initiële behandeling na verloop van tijd een terugval zullen krijgen. De initiële behandeling bestaat uit een of meerdere anti-kankerkuren gevolgd door stamceltransplantatie indien de patiënt hiervoor in aanmerking komt. In de refractaire fase daarna worden patiënten behandeld zodra ze tekenen van voortgaande ziekte vertonen. Hiervoor zijn meerdere strategieën voorhanden totdat de patiënt uiteindelijk therapieresistent wordt. Aanvankelijk werd MM behandeld met conventionele chemotherapie. Dit is veranderd sinds de introductie van doelgerichte therapieën zoals thalidomide aan het eind van de vorige eeuw, gevolgd door bortezomib en lenalidomide aan het begin van de 21e eeuw. Daarnaast wordt door vooruitgang in de techniek vaker gebruik gemaakt van stamceltransplantatie. Deze vernieuwde therapieën hebben een hogere unit prijs dan conventionele chemotherapie en zorgen daardoor voor een flinke impact op budgetten. Als reactie hierop vragen financiers vaker om aanvullende informatie over de kosten en de effectiviteit van dure geneesmiddelen in de dagelijkse praktijk. Deze methode heeft als doel de onzekerheid betreffende de werkelijke, zogenaamde 'real world', waarde van deze middelen te verminderen. In Nederland specifiek heeft die onzekerheid geleid tot een beleid waarin vergoeding van dure intramurale geneesmiddelen wordt gekoppeld aan bewijsvoering van de kosteneffectiviteit. Sinds 2008 wordt er voor deze geneesmiddelen onderzoek verricht naar het gerechtvaardigd gebruik, de effectiviteit en de kosten in de dagelijkse praktijk. Het medicijn bortezomib zou in aanmerking zijn gekomen voor dit nieuwe beleid, echter de behandeling van recidief / refractair MM met bortezomib was ten tijde van de beleidswijziging reeds goedgekeurd. Omdat er destijds nog weinig bekend was over de uitvoerbaarheid en het nut van uitkomstenonderzoek in het verzamelen van bewijs van waarde van een medicijn is besloten tot het uitvoeren van de in dit proefschrift gepresenteerde pilot uitkomstenonderzoek naar bortezomib.

Voor een beter begrip van de impact van de vooruitgang in MM behandeling op de efficiëntie van zorg was het van belang om eerst de effectiviteit en kosteneffectiviteit van de huidige standaard van zorg in kaart te brengen. In Deel 1 van dit proefschrift beschrijven wij de huidige kennis van de efficiëntie van behandeling van het multipel myeloom en de kwaliteit van onderzoek tot dusver op dit gebied (Hoofdstuk 2), met in het bijzonder aandacht voor de doelgerichte therapieën (Hoofdstuk 3). In Hoofdstuk 2 bespreken wij meerdere studies die aantonen dat stamceltransplantatie in combinatie met intensieve chemotherapie kosteneffectief is. Het aantal studies was echter beperkt doordat derge-

lijke evaluaties in het laatste decennium minder frequent worden verricht, vooral voor de doelgerichte therapieën zijn er weinig van dergelijke studies ondanks hun belangrijke bijdrage aan de verbeterde prognose. Daarnaast zou de kwaliteit en documentatie van de gebruikte methodologie in deze studies op een aantal vlakken verbeterd kunnen worden. Slechts enkele studies maakten gebruik van standaard methoden die men mag verwachten van een kwalitatief goede economische evaluatie, zoals de beoordeling van onzekerheid in de resultaten en discontering. Daarnaast baseerden de meeste studies de effectiviteit van de comparator groepen op niet-experimentele studies met verschillende patiëntenpopulaties. Vanwege een schaarste aan literatuur over economische evaluaties van doelgerichte therapieën presenteren wij in Hoofdstuk 3 een overzicht van het therapeutische en economische bewijs beschikbaar over de efficiëntie van doelgerichte therapieën in MM. Wij stellen vast op basis van therapeutisch bewijs dat doelgerichte therapieën een betere werkzaamheid en veiligheid hebben ten opzichte van conventionele behandelingen. Uitkomsten onderzoek heeft laten zien dat zowel patiënten tijdens thalidomide onderhoudsbehandeling als recidief / refractaire patiënten reagerende op bortezomib, een betere gezondheidsgerelateerde kwaliteit van leven rapporteren. Er is meer bewijs nodig over het effect van lenalidomide behandeling op de kwaliteit van leven. Op het moment van schrijven van Hoofdstuk 3 werden in totaal drie studies gevonden met betrekking tot de kosteneffectiviteit van doelgerichte therapieën in MM. Twee van deze studie vergeleken bortezomib met lenalidomide bij recidief / refractaire patiënten door een indirecte vergelijking met behulp van data uit twee onafhankelijke experimentele studies. De gezondheidswinst van behandeling met bortezomib en lenalidomide die deze modeleer studies veronderstelden bleek invloedrijk en vaak een overschatting van de werkelijkheid.

Op basis van tegenstrijdige resultaten van modeleer studies die zijn gebaseerd op efficiëntie van vergelijkbare experimentele studies, concluderen wij dat bestaande modeleer studies in MM niet in staat zijn gebleken om de beperkingen in data van vergelijkende effectiviteit te overwinnen. Tot slot blijkt uit ons onderzoek dat de utiliteitswaarde die wordt gebruikt om de gezondheidsgerelateerde kwaliteit van leven gecorrigeerde levensjaren (QALY) te kunnen genereren gebaseerd is op onderzoek van vóór de beschikbaarheid van de doelgerichte therapieën. Hiermee tonen wij aan dat er een noodzaak is voor patiëntuitkomsten onderzoek van doelgerichte therapieën.

In Deel 2 van dit proefschrift onderzoeken wij de haalbaarheid en de implicaties van uitkomstenonderzoek in het evalueren van de effectiviteit en het juist gebruik (Hoofdstuk 4), de kosten (Hoofdstuk 5), en de relatieve kosteneffectiviteit (Hoofdstuk 6) van bortezomib bij recidief / refractaire MM patiënten. Uit ons onderzoek blijkt dat observationele gegevens bruikbaar zij voor het beoordelen van juist gebruik, de kosten en

de totale overleving van deze patiënten behandeld met bortezomib in de dagelijkse praktijk. Echter, de evaluatie van de effectiviteit en het juist gebruik van bortezomib in Hoofdstuk 4 laat zien dat de resultaten alleen generaliseerbaar zijn bij patiënten met een ongunstige prognose. Vooral de groep patiënten die werden behandeld in de eerste jaren nadat bortezomib op de markt beschikbaar was gekomen, hadden een ongunstigere prognose dan degenen die behandeld waren in de cruciale RCT die was gebruikt voor de goedkeuring van vergoeding. Bovendien waren deze patiënten slechter af in vergelijking met patiënten behandeld met bortezomib na de periode voor het uitvoeren van uitkomstenonderzoek. Hoofdstuk 5 laat zien dat de 'real-world' kosten van behandeling van recidief / refractaire MM sterk verschillend is en dat deze kosten vooral beïnvloed worden door ziekenhuisopnames en de prijs van doelgerichte therapieën. Hogere kosten werden daarnaast geassocieerd met ongunstige prognostische factoren en behandelingsgerelateerde bijwerkingen. Behandelstrategieën gebaseerd op bortezomib bleken aanzienlijk duurder te zijn door de hogere prijs, frequentere ziekenhuisbezoeken voor intraveneuze toediening van medicatie en hogere kosten voor comedicatie. Voorgaande hoofdstukken laten zien dat het mogelijk is om met uitkomstenonderzoek bewijs te verzamelen voor de kosteneffectiviteit van nieuwe doelgerichte therapieën, zoals bortezomib, in de eerste jaren na introductie op de markt. Tot slot laten wij in Hoofdstuk 6 zien dat het onmogelijk is om de onzekerheid over relatieve kosteneffectiviteit weg te nemen doordat patiënten behandeld met bortezomib niet vergelijkbaar zijn met controles die werden behandeld met alternatieve behandelingen (oftewel confounding by indication). Op basis van de bevindingen in Deel 2 van dit proefschrift concluderen wij dat de dynamiek in de dagelijkse praktijk tijdens de introductie fase van bortezomib voor MM, anders gezegd de onervarenheid met het medicijn en de reeds bestaande variatie van behandeling, de haalbaarheid beperkt om een valide controle te identificeren en om relatieve kosteneffectiviteitschattingen te maken. Beleidsmakers, die uitkomstenonderzoek verlangen om conclusies te trekken over de relatieve effectiviteit en kosteneffectiviteit van nieuwe doelgerichte geneesmiddelen voor MM, worden gewaarschuwd voor de invloed van de introductie fase van deze geneesmiddelen in de dagelijkse praktijk. In plaats daarvan adviseren wij beleidsmakers zorgvuldig af te wegen of uitkomstenonderzoek voldoende onzekerheid wegneemt of dat andere opties (zoals financieel- of resultaatgebaseerde regelingen voor risicodeling) geschikter zijn om voldoende waarde voor geld te garanderen van nieuwe doelgerichte geneesmiddelen.

Sinds het tijdperk van de doelgerichte therapieën is er een betere overleving maar deze is nog beperkt doordat sommige patiënten ondanks de medische vooruitgang een slechte prognose houden. De heterogene effectiviteit van behandeling zou ondervangen kunnen worden door het gebruik van een behandelschema naar risico, ook wel 'risk-stratified treatment' (RST) genaamd, gebaseerd op biomarkers die onafhankelijk

geassocieerd zijn met de uitkomsten van de verschillende patiëntgroepen. De waarde van RST moet nog worden bewezen maar een dergelijke individuele benadering van MM patiënten zou de effectiviteit en kosteneffectiviteit van nieuwe doelgerichte geneesmiddelen sterk kunnen verbeteren. In Deel 3 van dit proefschrift bestuderen wij deze RST benadering en passen wij de methode van vroeg stadium economische evaluatie toe om de potentiële waarde van de RST in te schatten. In Hoofdstuk 7 worden de potentiële therapeutische waarde en de kostenbesparing aangetoond voor prognostisch ongunstige markers in MM patiënten, de onzekerheid over de resultaten was daarbij een invloedrijke factor. Vergelijkbare bevindingen werden gevonden in Hoofdstuk 8 waarin wij dezelfde methode toepassen op een nog vroegere fase van technologische ontwikkeling van hypothetische markers die vooraf respons op de behandeling bij chronische myeloïde leukemie voorspellen. Deze twee studies benadrukken een aantal methodologische overwegingen die men in acht moet nemen bij de uitvoering van vroeg stadium economische evaluaties van RST benaderingen in de hematologie, vooral het vinden van robuust bewijs voor vergelijkbare effectiviteit tussen nieuwe doelgerichte therapieën blijkt moeizaam en beperkt tot subgroepen met kleine patiëntaantallen. Het ontwikkelen van een vroeg stadium eenvoudig economisch model was daarentegen wel nuttig voor het kwantificeren van de kosten en de effectiviteit van de huidige zorgstrategieën. Het bood eveneens een platform voor discussie met belanghebbenden over de behoeften in de zorg en de mogelijke implicaties die deze technologie zou kunnen hebben op de besluitvorming over behandeling in de dagelijkse praktijk. Door verschillende zorgpaden in het model te verwerken konden scenarioanalyses worden verricht waarmee de verschillende veronderstellingen over de klinische utiliteit van de nieuwe technologie konden worden onderzocht en waarmee bovendien bepaald kon worden bij welke prijs de nieuwe technologie niet langer rendabel is.

In het laatste Hoofdstuk 9 van dit proefschrift voeren wij een gedetailleerde discussie over de implicaties van deze bevindingen voor beleidsmakers, zorgverleners en onderzoekers. Op basis van onze bevindingen bespreken wij vervolgens in welke mate vergelijkbare uitdagingen kunnen worden verwacht in vroeg stadium economische evaluaties en uitkomstenonderzoek in MM. Wij besluiten dit proefschrift met een discussie over de aanbevelingen voor nieuwe initiatieven om de doelen te halen van zowel vroeg stadium als klassieke economische evaluaties in MM.

## ACKNOWLEDGEMENTS

There are a number of people that I must extend a thank-you to for their help along the way. To my dear promoters, Carin and Pieter, I thank you both for supporting me in this endeavor. You have given me a great deal of your time and I appreciate everything I learned by working with you. Carin, you were always the first to remind me to not lose sight of the patient and that therapeutic effects are most important. This is a lesson that all health economics should take to heart. Pieter, thank you for your patience with my endless questions about treatment decision making. Thank you also for challenging me as an epidemiologist turned health economist to know when it's important to choose between the two. As for my co-promoter, Ken, I don't know where to start and end since you have taught me so many things. I could fill pages of my thesis with my gratitude. Thank you first and foremost for being a mentor. That was something I have always needed but even more so during this experience. Thanks to your open door policy we have had many interesting discussions, both related and (especially) unrelated to work. Thank you for these inspiring moments. I would also like to thank the other members of the small and large committee for reviewing this thesis and for challenging me during my defense.

To the two special ladies that I worked with most closely during my years at iBMG, Chantal and Margreet, thank you for all the fun times as well as the more stressful moments we shared when working together. I hope we are able to maintain such a close working relationship even as we each move onto the next career phase. When are we going to have that party to celebrate the end of the pilot study? I guess January 17th is the second of three, and I'm very much looking forward to number three.

My dear friend and dearest colleague, Jacoline, we have come so far together. I am so thankful for your talent to eavesdrop and your love for gossip columns, because without it we might never have met. But most importantly, thank you for your support throughout the years.

To Annemieke, my officemate of 5 plus years, you have always listened to me ramble about anything and everything and you always followed it up with genuine sympathy and advice. Thanks to our half Dutch/half English conversations, you have played quite an important role in the progression of my Dutch. Thank you for everything.

To my other colleagues at iBMG, thank you all for being the most exclusive yet inclusive group of kind-hearted, fun loving, and seriously cool people. I don't think I will ever have the chance again to work with such a special group of colleagues. I will miss the many

fun times we shared. For those of you still engaged in battle, I wish you all the best and I can't wait to celebrate your moment with you. Thank you for sharing in mine.
As for the last of my colleagues, I must extend a huge thank you to Cobi. Thank you so much for your help in getting me through the final stages. You always seemed happy to help and your support made what could have been a stressful process much easier for me.

There are a few external people that I worked with along the way that I must acknowledge for their contribution to this thesis. Ronnie van der Holt, thank you for the time you invested in extracting the data used in Chapters 4 and 7. I also want to thank you for your helpful comments on the draft versions of these chapters and for the very helpful discussions we had about the value of outcomes research. Your unique perspective has taught me a great deal. Silvia Verelst, thank you for the long, helpful discussions about the real-world treatment of multiple myeloma patients in the Netherlands. And to Michel van Agthoven, thank you for your interest in my research and your support in getting it published.

To my family, you have all in your own separate way contributed to this achievement. Christy, Stephanie, Terri and Danny, each of you have played an important role in the process that has gotten me here, whether it was more recent or from the beginning. Thank you all for believing in me and never doubting for a second that I would get this far.

My dear Gran, I thank you especially for your emotional, moral and financial support throughout the many years I have spent studying. And since I know you will probably be the only person besides me who will ever read this book entirely from front to cover, I want to apologize in advance for any typos

My dear mother, this thesis is dedicated to your spirit of hard work and wonderfully stubborn commitment. I couldn't have done this without having you as an example. Thank you from the bottom of my heart for teaching me how important it is to always follow through with a promise.

My dear Emile, thank you for all of your emotional support along the way. Your tough love was exactly what I needed at the more difficult moments. Thank you for proving by example that it can be done.

To Isla, my sunshine, thank you for helping mommy in the final reach for the end. Hooray, we made it!

## PHD PORTFOLIO

| | |
|---|---|
| Phd student: | Jennifer Gaultney |
| Department: | Institute of Health Policy and Management |
| PhD period: | 2008 - 2013 |
| Promoters: | Prof. dr. Carin Uyl-de Groot |
| | Prof. dr. Pieter Sonneveld |
| Co-promoter: | Dr. William K. Redekop |

### Training

Decision analytic modeling for economic evaluation, University of York Centre for Health Economics, York, England, 2008

Bayesian statistics, Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, Netherlands, 2009

Advanced diagnostic research, Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, Netherlands, 2009

Genetic analysis in clinical research, Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, Netherlands, 2009

Academic writing in English, Erasmus University Rotterdam, 2010

### Teaching

Thesis coordinator, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University Rotterdam 2010-2012

Evidence-based medicine, continuing education program International Health and Policy Evaluation, Institute of Health Policy and Management, Erasmus University Rotterdam, 2010-2011

Health technology assessment, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University Rotterdam, 2010-2011

Health technology assessment minor, bachelor program Health Sciences, Policy and Management, Institute of Health Policy and Management, Erasmus University Rotterdam 2011-2012

Health economics, Netherlands Institute for Health Sciences summer program, Erasmus Medical Center, 2010-2012

Supervision and co-supervision of several theses, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University Rotterdam, 2010-2013

### Oral presentations

Economic evaluation of the multiple myeloma biochip, CTMM Biochip annual meeting, 2009-2013

Critical review of economic evaluations in Multiple Myeloma: An overview of the trends and quality in the methodology, GE-iMTA lunch seminar, 2010

Pilot outcomes research: effects and costs of bortezomib in relapsed or refractory multiple myeloma, BMG lunch seminar, 2010

Methods and challenges with conducting early economic evaluations of companion diagnostics: Experience based on NSCLC and CML, GE-iMTA CTMM meeting, 2011

Early-stage health technology assessment of pharmacogenomic testing in multiple myeloma, GE-iMTA lunch seminar, 2012

### Research posters

Universal Steps in Performing Early-stage Technology Assessment, Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress,Athens, Greece, 2008

Outcomes research of bortezomib indicated for multiple myeloma in the context of the Dutch expensive medicines reimbursement policy: threats to the internal validity of the incremental effectiveness estimate, Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress, Paris, France, 2009

Real-world cost-effectiveness of bortezomib in relapsed or refractory multiple myeloma in the Netherlands, Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress in Prague, Czech Republic, 2010

Application of cost-effectiveness analysis to demonstrate the potential value of companion diagnostics in chronic myeloid leukemia, Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress in Prague, Czech Republic, 2010

Real-world healthcare costs of relapsed/refractory multiple myeloma during the era of novel cancer agents, 16$^{th}$ annual meeting of the European Hematology Association, London, England, 2011

What is the role of early health technology assessment of biomarkers in the pre-clinical development phase? A reflection on lessons learned with multiple myeloma, Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress, Berlin, Germany, 2012

Turning the tables to address the real value of real-world observational studies of novel anti-cancer agents in multiple myeloma, Annual European International Society

for Pharamacoeconomics and Outcomes Research (ISPOR) Congress, Berlin, Germany, 2012

Early-stage economic evaluations of stratified medicine in multiple myeloma. Annual European International Society for Pharamacoeconomics and Outcomes Research (ISPOR) Congress, Dublin, Ireland, 2013

**Publications not included in this thesis**

O'Prinsen A, Gaultney J, Redekop KW. Universal steps for conducting early-stage medical technology assessment. ISPOR Connections. November/December 2009;15(6).

Franken MG, Gaultney JG, Blommestein HM, Tan SS, van Gils CWM, Huijgens PC, Sonneveld P, Redekop WK, Uyl-de Groot CA (2012). Pilot outcomes research: effects and costs of bortezomib in relapsed or refractory multiple myeloma , 382 pp., Erasmus Universiteit Rotterdam - GE-iMTA: Rotterdam.

Franken MG, Van Gils CWM, Gaultney JG, Delwel GO, Goettsch W, Huijgens PC, Steenhoek A, Punt CJA Koopman M, Redekop WK, Uyl-de Groot CA. Practical feasibility of outcomes research in oncology: lessons learned in assessing drug use and cost-effectiveness in the Netherlands. Eur J Cancer. 2013 Jan;49(1):8-16.

Gaultney JG, Uyl-de Groot CA. Efficient allocation of novel agents in multiple myeloma: a work in progress. Oncologist. 2013;18(1):5-7.

**Other**

International student ambassador, master program Health Economics, Policy and Law, 2009-2010

Peer reviewer for The Oncologist, 2012-2013

Peer reviewer for the Journal of Personalized Medicine, 2012-2013

Peer reviewer for Pharmacogenomics, 2013

## CURRICULUM VITAE

Jennifer Gail Gaultney was born on November 12th, 1982 in Atlanta, Georgia, USA. In 2001, she graduated from John McEachern High School with honors. She then attended the University of Georgia in Athens, Georgia, USA where she graduated cum laude with a bachelor's degree in nutrition science within the pre-medicine track in 2005.

She then decided to further her studies by focusing on public health. In 2007, she obtained a masters degree cum laude in public health with a concentration in epidemiology from the Rollins School of Public Health at Emory University in Atlanta, Georgia, USA. Her master's thesis focused on modifiable risk factors for enteral feeding support dependence in patients with head and neck cancer. During her time at Rollins, Jennifer was a research assistant at the Winship Cancer Institute of Emory University. She also served as an intern in strategic development at the International Affairs department of the American Cancer Society in Atlanta, Georgia, USA.

Combining both her interests in the growing healthcare expenditure dilemma and a desire to live abroad, in 2007 she enrolled in the Health Economics, Policy and Law master's program of the Erasmus University in Rotterdam, Netherlands. She graduated in 2008 after defending her master's thesis, which was based on a modeling study to assess the potential cost-effectiveness of a prognostic test being developed for prostate cancer. Since this time she has worked as a researcher in health economics at the institute for Medical Technology Assessment of the Institute of Health Policy and Management at the Erasmus University Rotterdam.

She currently lives in Amsterdam with her fiancé, Emile, and their daughter, Isla.