Evidence from clinical practice to support healthcare decision making

Evaluation of clinical and economic outcomes of new therapies for metastatic renal cell carcinoma

Evidence From Clinical Practice To
Support Healthcare Decision Making
Evaluation of clinical and economic outcomes
of new therapies for metastatic renal cell carcinoma

Bewijslast uit de klinische praktijk ter ondersteuning
van besluitvorming in de gezondheidszorg
De evaluatie van klinische en economische uitkomsten van nieuwe therapieën
voor gemetastaseerd niercelcarcinoom

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

donderdag 10 november 2016 om 11.30 uur

Saskia de Groot
geboren te Gouda

Erasmus University Rotterdam

Promotiecommissie

Promotoren
Prof.dr. C.A. Uyl-de Groot
Prof.dr. L.A.L.M. Kiemeney

Overige leden
Prof.dr. F.F.H. Rutten
Prof.dr. A. de Boer
Prof.dr. S. Osanto

Copromotor
Dr. W.K. Redekop

Contents

Chapter 1

General introduction

The prognosis of cancer patients has improved dramatically over the past years and this is reflected in a significant increase in life expectancy as well as important improvements in health-related quality of life. Factors that contributed to these improvements are the introduction of new drugs against cancer, but also advances in screening and diagnostic approaches and better surgical techniques.[1]

The costs of factors that contributed to the improvements in health outcomes are substantial; the economic burden of cancer was more than 126 billion Euros in the European Union (EU) in 2009, with healthcare costs accounting for 51 billion Euros[2] which accounted for 4% of the total healthcare expenditures in the EU. Of these, 13.5 billion Euros (27%) were spent on drugs.[2]

The enormous pressure of the costs of cancer drugs on healthcare budgets creates great tension between financial sustainability of healthcare systems and accessibility to (new) treatments. Healthcare authorities need to make choices between reimbursement of healthcare procedures, services and programs,[3] but these choices have to be made under considerable uncertainty about (long-term) costs and effects. As with other procedures, services and programs, the effectiveness and cost-effectiveness of new drugs are uncertain at the time they are introduced. As a consequence, conditional funding was introduced in the Netherlands in 2006. At that time, conditional funding implied the additional funding of innovative drugs (i.e. reimbursing hospitals [most of] the drug costs) for a predetermined period of time on the condition that data regarding uptake, use and outcomes of these drugs in clinical practice were to be collected. After this time period, these data were used to decide whether or not additional funding continued to exist. This policy applied to promising but expensive inpatient drugs, including two drugs for metastatic renal cell carcinoma (mRCC), that is, bevacizumab and temsirolimus, and ensured equal and undelayed access to these drugs. The collection and analysis of data about uptake, use, and outcomes of these new drugs for mRCC in clinical practice form the focus of the present thesis. This thesis also illustrates the possibilities and impossibilities (or limitations) of evaluating uptake, use, and outcomes of new drugs in clinical practice, and reflects on the difficulties of making decisions about future funding (based on evidence from clinical practice).

Before explaining the evolution of funding policies in the Netherlands and the introduction of the new drugs for mRCC in this context, the position of evidence on outcomes in clinical practice within a drug's life-cycle will be addressed using the framework by Jönsson et al (Figure 1.1).[4]

The position of evidence on outcomes in clinical practice

Generally, new drugs are introduced based on evidence from clinical trials to ensure that patients have access to safe and efficacious drugs. Clinical trials usually use strict in- and exclusion criteria, and as a consequence study populations are relatively homogeneous and often comprise relatively young and healthy patients. As a result, little is known about the effects of the drug in clinical practice comprising older patients and patients with comorbidities.

Besides evidence from clinical trials, some reimbursement authorities consider the costs of new drugs in their decision (or advice) to reimburse a new drug, in order to ensure short- and long-term financial sustainability of the healthcare system. Therefore, they require evidence of cost-effectiveness for certain types of drugs (e.g. expensive drugs). Cost-effectiveness analyses assess the additional effects of a new treatment against the additional costs. The outcomes of such analyses can be used to decide whether the additional benefits are worth the additional costs. However, initial estimates of cost-effectiveness, which are often based on clinical trials, are accompanied by many uncertainties. For example, little data are generally available on long-term costs and effects. Despite these uncertainties regarding effectiveness and cost-effectiveness, reimbursement authorities need to make decisions (or give advice) about reimbursement of new drugs.

Regardless of the decision about reimbursement, drugs will not necessarily be used widely in clinical practice. Several barriers, such as limited

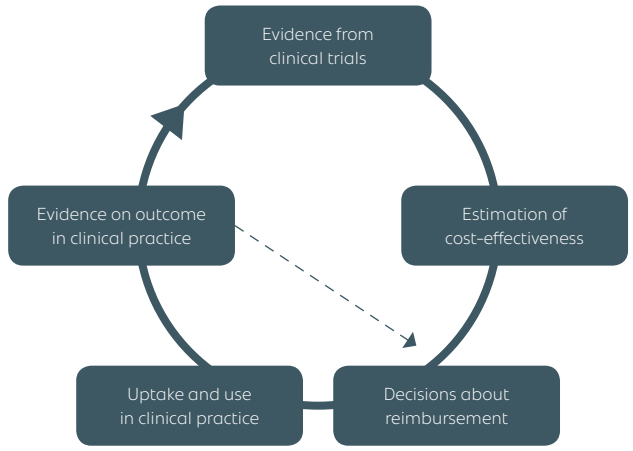


Figure 1.1 Adapted framework from Jönsson et al illustrating the position of evidence on outcomes in clinical practice[4]

hospital budgets, could hinder the uptake and use of a new drug in clinical practice. As a consequence, its value is not always exploited to its full potential. Furthermore, outcomes (i.e. effectiveness and cost-effectiveness) in clinical practice might be less favourable than expected, especially when new drugs are prescribed to a wider population than the population eligible to participate in the clinical trial.

In the Netherlands, information about uptake and use and evidence on outcomes in clinical practice were used to assess the real-world value of expensive drugs, after four years of conditional funding. After this time period, this information was used to decide whether or not additional funding continued to exist. This procedure has some similarity with the recently revised Cancer Drugs Fund in the United Kingdom (UK). New cancer drugs could be considered for funding within the new Cancer Drugs Fund for a predetermined time period if the drugs' effectiveness and cost-effectiveness are uncertain. During this time period, additional evidence are to be collected, which should help to determine whether the treatment should be accepted for routine use in the National Health Service (NHS).[5]

The evolution of funding policies in the Netherlands

Healthcare costs associated with cancer in the Netherlands doubled from 2.1 billion Euros in 2003 to 4.3 billion Euros in 2011.[6] In the same period, the costs of cancer drugs increased from 270 million Euros in 2003 to 670 million Euros in 2011.[7] From the 670 million Euros spent on drugs, 415 million Euros were spent on so-called expensive drugs. These costs further increased to 519 million Euros in 2013.

To keep a lid on healthcare costs, one of the proposed methods of the current Dutch cabinet (i.e. Rutte-Asscher cabinet installed in November 2012) was to introduce 'strict' package management, meaning the strict application of the reimbursement criteria, i.e. necessity, effectiveness, cost-effectiveness and feasibility, on both the treatments currently covered through the basic health insurance package as well as new treatments.[8] As a consequence, various measures were announced, including a wide implementation of conditional access and funding combined with risk-oriented package management (i.e. the option to carry out evaluations on subjects where risks have been identified regarding the package's affordability, accessibility or quality).

Policy for expensive inpatient drugs until 2012

Although conditional access and funding were announced in the coalition agreement in 2012, such a policy has existed in the Netherlands since 2006 to remove barriers for hospitals to prescribe expensive drugs.[9] This policy aimed to avoid unequal access (by removing financial barriers) and guarantee early access to expensive drugs through conditional funding. The policy applied to drugs with an added therapeutic value and an expected nationwide budget impact of at least 0.5% of the total drug costs of hospitals (i.e. an expected budget impact of at least 2.5 million Euros per year). For these drugs, hospitals received additional funding (i.e. 80% of the costs), besides their fixed budgets. Hospitals needed to pay the remaining 20% in order to discourage unnecessary care. In exchange, the collection of data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice was required. These data were intended to complement the findings from clinical trial(s), and to evaluate a drug's real-world value after three years (which became four years in 2010) of initial funding. This evaluation was used to decide about future funding. One of the criteria for a drug to remain eligible for additional funding was that the drug needed to be cost-effective.

The policy for expensive drugs was abolished on January 1st 2012. Since then, expensive drugs are defined as add-ons, i.e. additions to so-called DOT (i.e. Diagnosis Treatment Combination – Moving towards Transparency) healthcare products. Hospitals no longer receive a fixed percentage of the costs of expensive drugs, but instead have to negotiate with individual health insurers about prices and volumes.

Policy for expensive inpatient drugs since 2012

As of January 1st 2012, hospitals have been using a new system to claim their costs to health insurers, the so-called DOT system. A DOT healthcare product includes all healthcare activities provided as part of a particular treatment. For some products maximum prices are determined on a national level by the Dutch Healthcare Authority, but for most products hospitals and individual health insurers need to negotiate a price. Generally, the costs of inpatient drugs (i.e. drugs provided by a hospital pharmacist) are included in these DOT healthcare products.

Besides DOT healthcare products, so-called add-ons were created to prevent disturbances in the homogeneity of DOT healthcare products. Expensive drugs, i.e. drugs eligible for additional funding according to the former policy for expensive drugs are defined as add-ons.[10] Hospitals can claim these add-ons in addition to a DOT healthcare product, but they do not receive additional funding as with the former policy for expensive inpatient drugs. Instead, hospitals and health insurers need to negotiate a price.

New add-on applications are approved if the costs per patient per year are higher than 10,000 Euros.[11] However, since January 1st, 2015 a threshold no longer exists because several issues were observed including issues related

to access to the expensive drugs. Since then, representatives of both clinicians and health insurers provide advice about whether a drug is eligible to become an add-on.[12]

Despite these changes in the funding of expensive drugs, the Dutch National Health Care Institute can still advice to *conditionally* fund a new expensive drug. This means that the drug receives conditional access to the basic health insurance package on the condition that the manufacturer encourages and supports data collection regarding effectiveness and cost-effectiveness in clinical practice.[13] Based on these outcomes, the Dutch National Health Care Institute can advise to exclude a drug from the basic health insurance package.[14]

Policy for expensive outpatient drugs

In contrast to expensive inpatient drugs, no additional requirements, such as the collection of data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice were imposed on expensive *outpatient* drugs (i.e. drugs provided by an independent pharmacist or a general practitioner with a dispensary) until 2012. Expensive outpatient drugs were funded according to the framework of the Medicine Reimbursement System [in Dutch: Geneesmiddelenvergoedingssysteem, GVS].[15] This framework makes different demands on evidence depending on the drug's therapeutic value; if a drug's therapeutic value is similar to that of medicines that have already been added to the Medicine Reimbursement System, applications only require pharmacotherapeutic evidence (so-called List 1A). However, if a drug has an added therapeutic value, pharmacoeconomic evidence and an estimation of the budgetary impact is required, besides pharmacotherapeutic evidence (so-called List 1B). Reimbursement limits exist for medicines on List 1A, while such limits do not exist for medicines on List 1B.

Since no additional requirements were imposed on expensive outpatient drugs, the system did not guarantee efficient use of these drugs. Additionally, the system did not stimulate any price competition. Since 2012, the funding of expensive outpatient drugs has been changing; the funding of TNF alpha inhibitors, oral oncology drugs and growth hormones was transferred to the hospital budgets (although these drugs are used outside a hospital setting). Similar to expensive inpatient drugs, add-ons were created for these expensive outpatient drugs.

The costs of expensive in- and outpatient drugs are subject to the maximum budget growth of hospital care, which is agreed upon by the Ministry of Health, healthcare providers and health insurers.[16] Hospitals were allowed

to grow 2.5 percent in the period 2012–2015 (1 percent in 2016). This measure aimed to control costs and further improve the quality of hospital care.

The introduction of new treatments for metastatic renal cell carcinoma

Renal cell carcinoma (RCC) is the most common type of kidney cancer, i.e. RCC accounts for about 90% of kidney cancers.[17] With RCC, a tumour develops from cells from the kidney tubules. These tubules are responsible for reabsorption, i.e. the process by which solutes and water are removed from the tubular fluid and returned to the bloodstream. Other types of kidney cancer, such as transitional cell carcinoma, begin in the renal pelvis (responsible for urine collection). Kidney cancer is the seventh most common cancer in men and eleventh most common cancer in women.[18] In the Netherlands, there were about 2,250 new cases with kidney cancer and 930 deaths in 2014.[19]

Twenty percent to 30% of all RCC patients have metastatic disease at initial presentation, and 20% to 40% undergoing nephrectomy (i.e. surgical removal of a kidney) for clinically localised RCC will develop metastases.[20] Patients with metastatic RCC (mRCC) are classified into three risk groups based on the International Metastatic Renal-Cell Carcinoma Database Consortium (IMDC) prognostic model.[21] The median overall survival is 43.2 months for patients with a favourable risk, 22.5 months for patients with an intermediate risk and 7.8 months for patients with a poor risk.[22] Previously, the Memorial Sloan-Kettering Cancer Center model was used to classify patients into risk groups.[23]

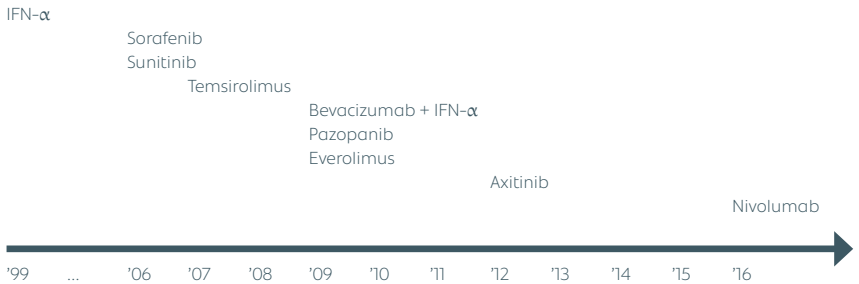
Until recently, treatment options for patients with mRCC were limited, because of its resistance to radiotherapy and chemotherapy.[24] As a consequence, interferon alfa (IFN-α) was the standard approach for many years, preceded by cytoreductive nephrectomy in patients with metastatic disease at initial presentation.

However, many new drugs have been introduced since 2006, starting with sorafenib and sunitinib (Figure 1.2). The better understanding of the molecular pathogenesis of RCC has led to the identification of targets for therapeutic interventions; as a result, these new therapies have been called targeted therapies.

Evidence of effectiveness from clinical trials

Targeted therapies have been introduced based on evidence from clinical trials that showed important improvement in health outcomes. For example, in one study, sunitinib increased median progression-free survival (PFS) from 5.0 months to 11.0 months compared to IFN-α.[25] Similarly, bevacizumab combined with IFN-α increased PFS compared to IFN-α monotherapy, and

Figure 1.2 The introduction (i.e. EMA approval) of targeted therapies for mRCC



Abbreviations: IFN-α, interferon alfa; EMA, European Medicines Agency.

pazopanib improved PFS compared to placebo, in patients with a favourable or intermediate risk.[26-28] For patients with a poor prognosis, temsirolimus improved PFS from 1.9 months to 3.8 months compared to IFN-α.[29] Besides these first-line therapies, several second-line therapies have been introduced since 2006.

The introduction of these drugs resulted in the updating of European guidelines for mRCC in 2009, including the recommendation of first-line therapy with sunitinib or bevacizumab (combined with IFN-α) for patients with a favourable or intermediate prognosis, and the recommendation to use temsirolimus for patients with a poor prognosis.[30] Second-line therapy with sorafenib has been recommended after cytokine therapy has been tried, and second-line everolimus has been recommended if patients had progressed on tyrosine kinase inhibitors (TKIs). In 2010, first- and second-line (post cytokines) pazopanib was added to the guideline.[31] Dutch guidelines were revised in 2010 to bring them in line with European guidelines.[32]

Evidence of cost-effectiveness

The improved health outcomes of patients with mRCC come at a price. Monthly costs of the targeted therapies for mRCC are higher than 3,000 Euros, whereas monthly costs of IFN-α amounted to about 800 Euros.

Little was known about the cost-effectiveness of the new drugs for mRCC in the Netherlands at the time they were introduced. Although the manufacturers of bevacizumab and temsirolimus were obliged to provide an indication of the cost-effectiveness of their drugs, the Dutch Committee of Pharmaceutical Aid [in Dutch, Commissie Farmaceutische Hulp] concluded that the cost-effectiveness of these drugs was insufficiently substantiated. Nevertheless, the inpatient drugs, bevacizumab and temsirolimus, were conditionally funded in the Netherlands; the condition being to collect data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice, in line with the policy for expensive drugs (both drugs are currently financed based on add-ons).

Also, the outpatient drugs for mRCC were reimbursed in the Netherlands, despite the lack of evidence of cost-effectiveness. This can be explained by the categories in which these drugs were placed. Some were placed on List 1A of the GVS, and therefore evidence of cost-effectiveness was not required. Others were placed on List 1B, but were exempted from pharmacoeconomic evidence, e.g. no such evidence was needed for everolimus, because everolimus was designated as an orphan drug and no alternative treatment was available.[33]

PERCEPTION registry

As noted above, one of the conditions for the funding of bevacizumab and temsirolimus for patients with mRCC was the collection of data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice. Therefore, a population-based registry (i.e. PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry) was created to evaluate the uptake and use of these drugs, as well as treatment outcomes in clinical practice (fourth and fifth element of the framework by Jönsson et al[4]). Since the registry was disease-oriented (instead of treatment-oriented), it did not only provide data on patients treated with bevacizumab and temsirolimus, but also on patients treated with other targeted therapies and on patients not treated with targeted therapy.

The PERCEPTION registry consisted of two parts; a retrospective study and a prospective study. Inclusion criteria for the retrospective study comprised a diagnosis of mRCC (i.e. metastases at initial presentation) between January 2008 and December 2010 of any histological subtype. In the prospective study, patients with RCC (all stages) of any histological subtype diagnosed from 2011 until June 30th 2013 were included.

Data on patient characteristics, treatment schemes and treatment endpoints (e.g. survival) were retrospectively collected from individual patient records. Furthermore, patients in the prospective study were asked to fill out questionnaires regarding health-related quality of life.

Thesis aims and outline

This thesis is structured as follows. The first part (i.e. chapters 2 to 5) focuses on treatment of patients with mRCC. In this part, the uptake and use of targeted therapies for mRCC in the Netherlands are evaluated, as well as treatment outcomes (i.e. survival, health-related quality of life and cost-effectiveness) in clinical practice, based on data from the PERCEPTION registry. This part will address the following aims:

1 To evaluate the uptake and use of targeted therapies for mRCC in the Netherlands and to study their effectiveness in terms of overall survival (Chapter 2).
2 To examine the factors associated with the prescription of targeted therapies in daily clinical practice (Chapter 2).
3 To evaluate the effect of cytoreductive nephrectomy on overall survival in primary mRCC patients treated with first-line sunitinib (Chapter 3).
4 To provide insight into the most important determinants of health-related quality of life (including progression of disease) of patients with mRCC (Chapter 4).
5 To estimate the real-world cost-effectiveness of several treatment strategies applied in patients with mRCC comprising one or more sequentially administered drugs (Chapter 5).

In the second part, practical recommendations will be provided regarding the collection and analyses of data about uptake, use and outcomes in clinical practice, in order to conduct outcomes research. This part will address the following aims:

6 To provide practical guidance in setting up patient registries for the collection of real-world data (Chapter 6).
7 To provide practical recommendations in constructing a discrete event simulation (DES) model to support cost-effectiveness analyses of treatment strategies spanning multiple treatment lines (Chapter 7).

Finally, chapter 8 provides a discussion of the findings and explores the implications and limitations of this thesis.

Note that chapters 2 to 7 are based on publications in, or intended for, international peer reviewed journals and can thus be read independently.

Chapter 2

Variation in use of targeted therapies for metastatic renal cell carcinoma: Results from a Dutch population-based registry

S. de Groot; S. Sleijfer; W.K. Redekop; E. Oosterwijk; J.B.A.G. Haanen; L.A.L.M. Kiemeney; C.A. Uyl-de Groot

BMC Cancer 2016;16:364.

Abstract

Objective: For patients with metastatic renal cell carcinoma (mRCC), targeted therapies have entered the market since 2006. The aims of this study were to evaluate the uptake and use of targeted therapies for mRCC in the Netherlands, examine factors associated with the prescription of targeted therapies in daily clinical practice and study their effectiveness in terms of overall survival (OS).

Methods: Two cohorts from PERCEPTION, a population-based registry of mRCC patients, were used: a 2008–2010 Cohort ($n=645$) and a 2011–2013 Cohort ($n=233$). Chi-squared tests for trend were used to study time trends in the use of targeted therapy. Patients were grouped based on the eligibility criteria of the SUTENT trial, the trial that led to sunitinib becoming standard of care, to investigate the use of targeted therapies amongst patients fulfilling those criteria. Multilevel logistic regression was used to identify patient subgroups that are less likely to receive targeted therapies.

Results: Approximately one-third of patients fulfilling SUTENT trial eligibility criteria did not receive any targeted therapy (29% in the 2008–2010 Cohort; 35% in the 2011–2013 Cohort). Patients aged 65+ years were less likely to receive targeted therapy in both cohorts and different risk groups (odds ratios range between 0.84–0.92); other factors like number of metastatic sites were of influence in some subgroups. Amongst treated patients, there was a decreasing trend in sunitinib use over time ($p=0.006$), and an increasing trend in pazopanib use ($p=0.001$).

Conclusions: Targeted therapies have largely replaced interferon-alfa as first-line standard of care. Nevertheless, many eligible patients in Dutch daily practice did not receive targeted therapies despite their ability to improve survival. Reasons for their apparent underutilisation should be examined more carefully.

Introduction

Kidney cancer accounts for about 3% of all cancers with an estimated incidence of 115,200 in Europe in 2012.[18] Renal cell carcinoma (RCC) represents 90% of all kidney cancers.[17] The prognosis is relatively good for patients with localised disease, which can be treated with surgery, but the prognosis of patients with advanced or metastatic disease (mRCC) is poor.[23]

Targeted therapies for mRCC have entered the market since 2006, sunitinib being the first. Sunitinib increased median progression-free survival (PFS) from five to 11 months,[25] and overall survival (OS) from 22 to 26 months compared to interferon-alfa (IFN-α) in mRCC patients with a clear-cell histology.[34] Subsequently, it became standard of care for patients with a good or intermediate prognosis according to the Memorial Sloan Kettering Cancer Center (MSKCC) risk score.[30] Recently, the effectiveness of sunitinib was demonstrated in a broader 'real-world' population.[35] Bevacizumab (in combination with IFN-α) and pazopanib were added to guidelines as first-line therapies for patients with a good or intermediate prognosis in 2009 and 2010, respectively.[30,31] For patients with a poor prognosis, temsirolimus was recommended[30] following the results of a multi-centre, phase III trial in mRCC patients without any restrictions in histologic type, showing an increase in OS from seven to 11 months compared to IFN-α.[29] Furthermore, a number of second-line therapies have been added to guidelines, such as sorafenib, everolimus and axitinib.[30,36]

Obviously, full and swift implementation of guidelines into clinical practice is essential to maximise the benefits of new therapies. However, the adoption of innovations in cancer care is generally quite heterogeneous, and differs between countries, and regions within countries.[1] A study by Jönsson et al showed widespread use of sunitinib in the eight of the countries they studied, despite small differences between countries.[4] Sorafenib was widely prescribed in France, while a very low uptake and use in the United Kingdom and the United States were found. Besides between-country variation, Jönsson et al found within-country variation in Sweden and suggested that more detailed information is needed on the use of first- and second-line therapies, to determine the extent of potential under- and overconsumption in different regions and different patient populations.[4]

The aims of this study were to evaluate the uptake and use of targeted therapies for mRCC in the Netherlands, examine factors associated with the prescription of targeted therapies in daily clinical practice and study their effectiveness in terms of OS.

Patients and methods

Study population

A population-based registry (entitled PERCEPTION) was created to include patients with mRCC. The PERCEPTION registry consisted of two parts; a retrospective study and a prospective study. In the retrospective study, eligible patients were selected from the Netherlands Cancer Registry (NCR), which maintains a cancer registration database of all cancer patients in the Netherlands. Inclusion criteria for the retrospective study comprised a diagnosis of mRCC (i.e. metastases at initial presentation) of any histological subtype. Patients diagnosed from January 2008 until December 2010 in 42 of 51 hospitals (both general and academic) in four regions, covering approximately half of the country, were included. All patients were followed for a minimum of three years or until death (2008–2010 Cohort).

The prospective study was designed differently in order to measure additional aspects of the disease, such as health-related quality of life (not reported in this study). In the prospective study, patients with RCC (all stages) of any histological subtype diagnosed from 2011 until June 30th 2013 in 25 of 32 hospitals (both general and academic) in three regions were included. In contrast to the 2008–2010 Cohort, this cohort also comprised patients with mRCC who were initially diagnosed with localised disease. Besides the NCR, the hospitals' financing systems were used to select eligible patients at an early phase (for quality of life measurements). All patients were followed until the end of 2013 or until death (2011–2013 Cohort).

Data collection

Data on baseline demographics, clinical and laboratory factors were retrospectively collected from individual patient records by using uniform case report forms to ensure consistent data collection. Furthermore, data on treatment schemes and treatment endpoints (e.g. survival) were collected. Laboratory factors, such as haemoglobin and corrected calcium levels, were standardised according to routinely used reference values. Data were collected by personnel of the NCR and data collection stopped at the end of 2013.

Statistical analyses

To study differences in the proportion of patients receiving targeted therapy per half a year chi-squared tests were used. Exact tests were used to study possible time trends in the use of different therapies amongst treated patients. Additionally, chi-squared tests for trend were conducted.

Then, the use of targeted therapies within risk groups was studied. Risk groups were created using a slightly modified version of the MSKCC risk score;[23] a time from initial diagnosis to metastatic diagnosis of less than one year was used as a risk factor instead of a time from initial diagnosis to initiation of treatment of less than one year, since many patients in the study population did not receive any targeted therapy, thereby making it impossible to calculate the time to treatment. Additionally, the WHO performance status was used instead of Karnofsky performance status.

Furthermore, patients were grouped based on the eligibility criteria of the SUTENT trial,[25] the trial that led to sunitinib becoming standard of care, to investigate the use of targeted therapies amongst patients fulfilling those criteria. Patients who had a clear-cell subtype, a WHO performance status of 0 or 1 and no brain metastases were classified as fulfilling the SUTENT trial eligibility criteria.

To identify patient subgroups that are less likely to receive targeted therapies in daily clinical practice among patients fulfilling SUTENT trial eligibility criteria, multilevel mixed-effects logistic regression was used to account for between-hospital variance. At the patient-level, patient and disease characteristics were taken into account including baseline demographics, clinical and laboratory factors.[37,38] Backward selection was used to select the covariates for the models; any non-significant covariates were excluded from the models one at a time.

OS was calculated from the start of therapy until death from any cause or the date of last follow-up, whichever came first, using the Kaplan-Meier method. For patients not receiving any targeted therapy, OS was calculated from the date of diagnosis.

Missing data regarding baseline characteristics were handled using multiple imputations by chained equations. This method generated imputations based on a set of imputation models, one for each variable with missing values.[39]

All analyses were performed separately for the 2008–2010 Cohort and the 2011–2013 Cohort, because of differences in inclusion criteria, patient selection and duration of follow-up. The significance level was set at $\alpha = 0.10$. Data analyses were conducted using STATA statistical analysis software (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP).

Results

Patient and disease characteristics of the 2008–2010 Cohort

714 patients newly diagnosed with mRCC between 2008 and 2010 were identified. Of these patients 69 were excluded (Figure S2.1), leaving 645 patients for data analysis. These patients were uniformly distributed across the three-year

period since 213 patients were diagnosed in 2008, 216 in 2009 and 216 in 2010. Median follow-up was 3.3 years (95% CI 3.2–3.6).

Table 2.1 shows the patient and disease characteristics for this cohort. Median age was 66 years (range 23–93) and the majority of patients was male (66%). The distribution of patients according to the MSKCC risk score showed a high proportion of patients (58%) with a poor prognosis (versus 42% with an intermediate prognosis). Since all patients in the 2008–2010 Cohort presented with metastatic disease, none of them had a favourable prognosis (i.e. time from initial diagnosis was less than one year). The supplementary material (Table S2.1) provides the observed patient and disease characteristics (without imputations).

Uptake of targeted therapies and their use in daily clinical practice (2008–2010 Cohort)

Table 2.2 shows the first-line therapies used in the 2008–2010 Cohort. 336/645 patients (52%) received a first-line therapy with the majority (282, 84%) treated with sunitinib. The distribution of patients across first-line therapies (per half-year period) is presented in Figure 2.1. There is evidence of a difference between the half-year periods in the proportion of patients receiving targeted therapy ($p=0.041$), but the chi-squared test for trend did not yield a significant result. Furthermore, no shift was found in the use of first-line therapies amongst treated patients.

Of the 336 patients receiving first-line therapy, 101 patients (30%) also received a second-line therapy, with everolimus being the most common (40%), followed by sorafenib (28%). There was an increasing trend in everolimus use over time ($p<0.001$) and a decreasing trend in sorafenib use ($p<0.001$); from 2010 onwards, everolimus largely replaced sorafenib.

Use of targeted therapies amongst patients with an intermediate prognosis (2008–2010 Cohort)

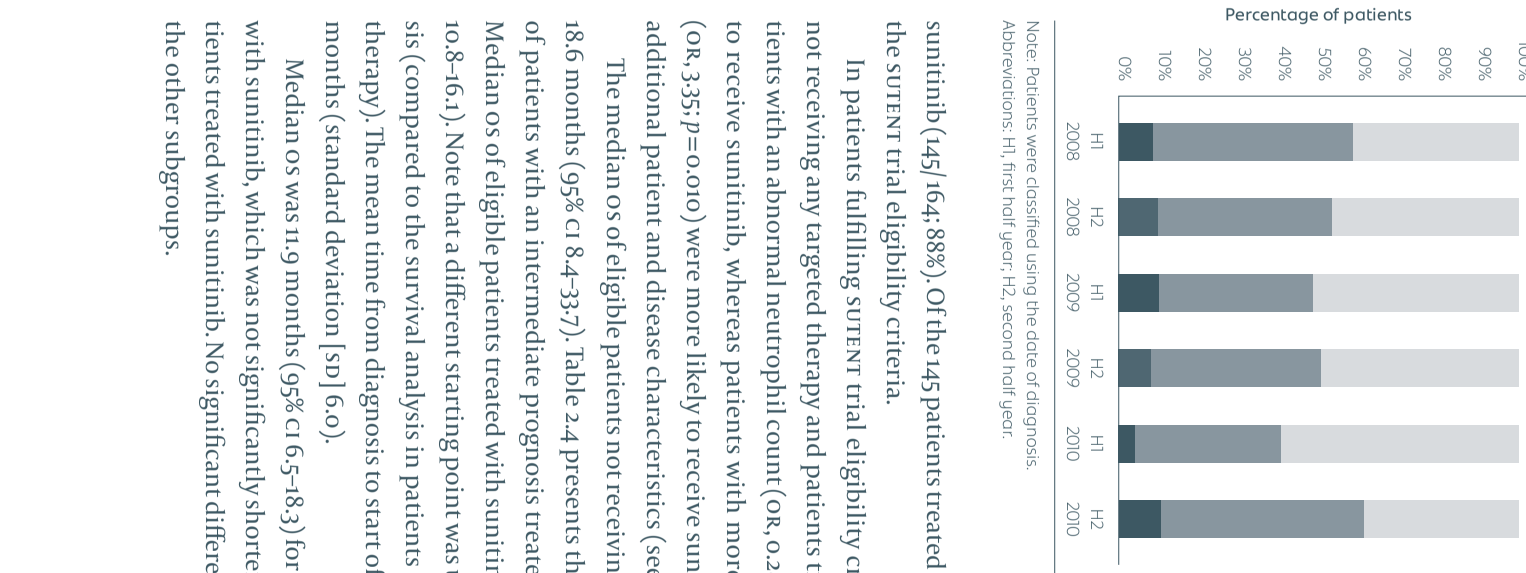
Forty-two percent (269/645) of the patients in the 2008–2010 Cohort had an intermediate prognosis. 105/269 patients (39%) received no targeted therapy. Some (n=15) of these patients received a metastasectomy (combined with a nephrectomy) with a possible curative intention, making systemic therapy redundant. 40 of the remaining 90 patients (44%) who were given neither targeted therapy nor a metastasectomy (combined with a nephrectomy) fulfilled the SUTENT trial eligibility criteria, indicating that they might have been eligible for treatment with sunitinib or another targeted therapy. 164/269 patients (61%) received a first-line treatment; the majority was treated with

Table 2.1 Patient and disease characteristics 2008–2010 Cohort and 2011–2013 Cohort

	2008–2010 Cohort: mRCC at the initial diagnosis (n=621)		2011–2013 Cohort: mRCC (n=221)	
Sex – n (%)				
Female	213	34%	60	27%
Male	408	66%	161	73%
Median age – yr (range)	66	23-93	66	27-93
Histology – n (%)				
Clear cell	354	57%	152	69%
Other *	267	43%	69	31%
WHO performance status – n (%)				
0-1	430	69%	178	81%
2-4	191	31%	42	19%
Site of metastasis – n (%)				
One	206	33%	87	39%
More than one	415	67%	134	61%
Liver metastasis – n (%)				
No	509	82%	175	79%
Yes	112	18%	46	21%
Lung metastasis – n (%)				
No	173	28%	74	33%
Yes	448	72%	147	67%
Bone metastasis – n (%)				
No	393	63%	158	71%
Yes	228	37%	63	29%
Brain metastasis – n (%)				
No	571	92%	200	90%
Yes	50	8%	16	7%
Haemoglobin – n (%)				
Normal	205	33%	85	38%
< LLN	416	67%	136	62%
Neutrophil count – n (%)				
Normal	383	62%	152	69%
> ULN	238	38%	69	31%
Platelet count – n (%)				
Normal	452	73%	159	72%
> ULN	169	27%	62	28%
Albumin – n (%)				
Normal	391	63%	130	59%
< LLN	230	37%	91	41%
Corrected serum calcium – n (%)				
Normal	421	68%	140	63%
> ULN	200	32%	81	37%
Alkaline phosphatase – n (%)				
Normal	432	70%	152	69%
> ULN	189	30%	69	31%
Lactate dehydrogenase – n (%)				
Normal	372	60%	179	81%
> 1.5 times ULN	249	40%	42	19%
Comorbidities – n (%)				
Zero or one	356	57%	151	68%
More than one	265	43%	67	30%
Time since RCC diagnosis				
> One year	NA	NA	16	7%
< One year	NA	NA	204	92%

Note: 24 patients in the 2008–2010 Cohort and 12 patients in the 2010–2013 Cohort were excluded from this table, since these patients received a metastasectomy (combined with a nephrectomy) with a possible curative intention, making systemic treatment redundant.
Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal; NA, not applicable.
* mRCC was clinically established without histopathological confirmation in 17% of patients and mRCC was classified as not otherwise specified without further subtyping in 13% of patients (Cohort 2008–2010). It is likely that a substantial proportion of these patients had a clear cell subtype.

Figure 2.1 Use of first-line drugs over time per half a year (2008–2010 Cohort)



Note: Patients were classified using the date of diagnosis.
Abbreviations: H1, first half year; H2, second half year.

sunitinib (145/164; 88%). Of the 145 patients treated with sunitinib, 102 fulfilled the SUTENT trial eligibility criteria.

In patients fulfilling SUTENT trial eligibility criteria (including patients not receiving any targeted therapy and patients treated with sunitinib), patients with an abnormal neutrophil count (OR, 0.28; $p=0.045$) were less likely to receive sunitinib, whereas patients with more than one metastatic site (OR, 3.35; $p=0.010$) were more likely to receive sunitinib after adjustment for additional patient and disease characteristics (see frequencies in Table 2.3).

The median OS of eligible patients not receiving any targeted therapy was 18.6 months (95% CI 8.4–33.7). Table 2.4 presents the median OS in subgroups of patients with an intermediate prognosis treated with first-line sunitinib. Median OS of eligible patients treated with sunitinib was 14.8 months (95% CI 10.8–16.1). Note that a different starting point was used for the survival analysis (compared to the survival analysis in patients not receiving any targeted therapy). The mean time from diagnosis to start of first-line sunitinib was 4.3 months (standard deviation [SD] 6.0).

Median OS was 11.9 months (95% CI 6.5–18.3) for ineligible patients treated with sunitinib, which was not significantly shorter than the OS of eligible patients treated with sunitinib. No significant differences were observed within the other subgroups.

Table 2.2 Treatment patterns 2008–2010 Cohort and 2011–2013 Cohort

	2008–2010 Cohort: mRCC at the initial diagnosis						2011–2013 Cohort: mRCC					
	All patients (n=645)		Intermediate prognosis (n=269)		Poor prognosis (n=376)		All patients (n=233)		Favourable/ intermediate prognosis (n=136)		Poor prognosis (n=97)	
No systemic therapy	309 (48%)		105 (39%)		204 (54%)		94 (40%)		52 (38%)		42 (43%)	
First-line therapy	336 (52%)	336 (100%)	164 (61%)	164 (100%)	172 (46%)	172 (100%)	139 (60%)	139 (100%)	84 (62%)	84 (100%)	55 (57%)	55 (100%)
Sunitinib		282 (84%)		145 (88%)		137 (80%)		110 (79%)		66 (79%)		44 (80%)
Temsirolimus		24 (7%)		5 (3%)		19 (11%)		3 (2%)		1 (1%)		2 (4%)
Sorafenib		11 (3%)		7 (4%)		4 (2%)		4 (3%)		3 (4%)		1 (2%)
Bevacizumab + IFN-α		6 (2%)		2 (1%)		4 (2%)		2 (1%)		1 (1%)		1 (2%)
Pazopanib		4 (2%)		4 (2%)		0 (0%)		11 (8%)		7 (8%)		4 (7%)
IFN-a		3 (1%)		0 (0%)		3 (2%)		1 (1%)		0 (0%)		1 (2%)
Everolimus		3 (1%)		1 (1%)		2 (1%)		2 (1%)		1 (1%)		1 (2%)
Pazopanib-everolimus		0 (0%)		0 (0%)		0 (0%)		3 (2%)		3 (4%)		0 (0%)
Other		3 (1%)		0 (0%)		3 (2%)		3 (2%)		2 (2%)		1 (2%)
Second-line therapy	101 (16%)	101 (100%)	57 (21%)	57 (100%)	44 (12%)	44 (100%)	37 (16%)	37 (100%)	25 (18%)	25 (100%)	12 (12%)	12 (100%)
Everolimus		40 (40%)		26 (46%)		14 (32%)		21 (57%)		12 (50%)		9 (75%)
Sorafenib		28 (28%)		15 (26%)		13 (30%)		5 (14%)		4 (16%)		1(8%)
Sunitinib		14 (14%)		8 (14%)		6 (14%)		1 (3%)		1 (4%)		0 (0%)
Temsirolimus		11 (11%)		4 (7%)		7 (16%)		3 (8%)		2 (8%)		1 (8%)
Pazopanib		4 (4%)		2 (4%)		2 (5%)		5 (14%)		4 (16%)		1(8%)
Bevacizumab + IFN-α		1 (1%)		1 (2%)		0 (0%)		2 (5%)		2 (8%)		0 (0%)
Other		3 (3%)		1 (2%)		2 (5%)		0 (0%)		0 (0%)		0 (0%)

Abbreviations: IFN-α, interferon alfa.

Table 2.3 Patient subgroups that are more or less likely to receive targeted therapy while fulfilling SUTENT trial eligibility criteria

	2008–2010 Cohort: mRCC at the initial diagnosis				2011–2013 Cohort: mRCC			
	Intermediate prognosis (n=142)		Poor prognosis (n=99)		Favourable/ intermediate prognosis (n=70)		Poor prognosis (n=39)	
	No targeted therapy (n=40)	Sunitinib (n=102)	No targeted therapy (n=29)	Sunitinib n=26	No targeted therapy n=25	Sunitinib (n=70)	No targeted therapy n=13	Sunitinib n=26
Sex – n (%)								
Female	NS	NS	NS	NS	2 (8%)	15 (33%)	NS	NS
Male	NS	NS	NS	NS	23 (92%)	30 (67%)	NS	NS
Median age – yr (range)	NS	NS	71 (43–84)	63 (42–79)	71 (44–79)	61 (39–79)	72 (57–82)	63 (42–79)
Site of metastasis – n (%)								
One	25 (62%)	38 (37%)	15 (53%)	20 (79%)	NS	NS	NS	NS
More than one	15 (38%)	64 (63%)	14 (47%)	6 (21%)	NS	NS	NS	NS
Neutrophil count – n (%)								
Normal	27 (68%)	87 (85%)	NS	NS	NS	NS	NS	NS
> ULN	13 (33%)	15 (15%)	NS	NS	NS	NS	NS	NS
Comorbidities								
Zero or one	NS	NS	15 (52%)	22 (85%)	NS	52 (74%)	NS	NS
More than one	NS	NS	14 (48%)	4 (15%)	NS	18 (26%)	NS	NS

Note: This table shows patient subgroups that are more or less likely to receive targeted therapy (i.e. first-line sunitinib) among patients fulfilling SUTENT trial eligibility criteria (according to the multi-level mixed-effects models). The multi-level models initially included all patient and disease characteristics as mentioned in Table 2.1 (besides hospital of diagnosis). Not significant (NS) means that this variable was not significantly associated to prescription of sunitinib at α=0.10 in a particular risk group/ cohort.
Abbreviations: NS, not significant; ULN, upper limit of normal.

Use of targeted therapies amongst patients with a poor prognosis (2008–2010 Cohort)

Fifty-eight percent (376/645) of the patients in the 2008–2010 Cohort, had a poor prognosis. 204/376 patients (54%) did not receive any targeted therapy. Of these patients, 9 patients received a metastasectomy (combined with a nephrectomy). 29 of the remaining 195 patients (15%) who were given neither targeted therapy nor a metastasectomy (combined with a nephrectomy) fulfilled the SUTENT trial eligibility criteria. 172/376 (46%) patients received a first-line treatment, which was mainly sunitinib (137/376; 80%). Of the 137 patients treated with sunitinib, 70 fulfilled the SUTENT trial eligibility criteria.

Amongst patients fulfilling SUTENT trial eligibility criteria, older patients (OR, 0.90; p=0.006) and patients with more than one comorbidity (OR, 0.26; p=0.090) were less likely to receive sunitinib, whereas patients with more than one metastatic site (OR, 5.38; p=0.034) were more likely to receive sunitinib (see frequencies in Table 2.3). Furthermore, a significant association was found between hospital of diagnosis and prescription of sunitinib (p=0.006).

Median OS of eligible patients not receiving any targeted therapy was 6.2 months (95% CI 1.7–9.9). Table 2.4 shows the median OS in subgroups of patients with a poor prognosis treated with first-line sunitinib. Median OS of eligible patients treated with sunitinib was 6.8 months (95% CI 5.3–10.7). The mean time from diagnosis to start of first-line sunitinib was 2.9 months (SD 5.5).

Median OS was significantly reduced in poor-prognosis patients treated with sunitinib but not fulfilling the SUTENT trial eligibility criteria (4.7 months, 95% CI 3.3–6.9). Additionally, OS was significantly reduced in patients with brain metastases and patients with a WHO performance status of 2–4.

Patient and disease characteristics of the 2011–2013 Cohort

The second cohort study included 791 patients with (m)RCC diagnosed between 2011 and 2013. Of these patients, 233 had metastatic disease; 75 in 2011, 102 in 2012 and 55 in 2013 (one unknown). Median follow-up of the patients with mRCC was 1.2 years (95% CI 1.1–1.4).

Table 2.1 shows the patient and disease characteristics of the patients with mRCC in this cohort. Median age was 66 years, and 73% (170/233) of the patients was men. Metastatic disease was present in 77% (179/233) of patients at the time of diagnosis, whereas 23% was initially diagnosed with localised disease. In this cohort, 4% of the patients with mRCC had a favourable prognosis, whereas 54% and 42% had an intermediate or poor prognosis, respectively.

Table 2.4 Overall survival in subgroups of patients treated with first-line sunitinib (Cohort 2008–2010 and Cohort 2011–2013)

		2008–2010 Cohort: mRCC at the initial diagnosis			2011–2013 Cohort: mRCC		
		n	Median OS in months (95% CI)	p-value	n	Median OS in months (95% CI)	p-value
All patients		282	9.1 (7.2-11.1)	–	109	10.1 (7.2-13.8)	–
Fulfilling SUTENT trial eligibility criteria	No	110	6.5 (4.9-8.9)		38	6.9 (3.4-10.9)	
	Yes	172	11.9 (8.8-14.6)	0.001	71	12.1 (8.9-NR)	0.007
Brain metastases	No	261	9.3 (7.6-11.9)		101	10.9 (7.8-18.0)	
	Yes	21	4.3 (2.1-11.5)	0.082	8	2.5 (0.8-7.5)	0.013
WHO performance status	0-1	248	10.3 (8.4-13.0)		100	11.3 (7.8-18.0)	
	2-4	34	3.3 (1.8-6.2)	<0.001	9	1.4 (0.6-7.5)	<0.001
Histology	Clear cell	204	10.0 (7.6-13.3)		81	10.6 (7.2-20.3)	
	Non-clear cell	78	6.9 (5.4-11.0)	0.081	28	10.0 (3.5-13.8)	0.333
Age	< 65 years	162	8.9 (6.5-10.8)		64	11.3 (7.2-20.3)	
	>= 65 years	120	10.0 (6.5-13.8)	0.837	45	10.0 (5.3-16.6)	0.429
Patients with an intermediate prognosis (or favourable prognosis)*		145	14.6 (11.5-16.0)	–	65	16.6 (10.1-NR)	–
Fulfilling SUTENT trial eligibility criteria	No	43	11.9 (6.5-18.3)		20	10.9 (2.7-NR)	
	Yes	102	14.8 (10.8-16.1)	0.290	45	18.0 (10.1-NR)	0.121
Brain metastases	No	136	14.6 (10.7-16.0)		61	16.6 (10.9-NR)	
	Yes	9	11.9 (4.3-29.3)	0.807	4	6.9 (2.5-NR)	0.228
WHO performance status	0-1	143	14.4 (10.8-16.0)		64	16.6 (10.1-NR)	
	2-4	2	-	0.230	1	-	0.247
Histology	Clear cell	111	14.8 (11.8-16.2)		49	18.0 (10.0-NR)	
	Non-clear cell	34	11.5 (6.3-17.7)	0.195	16	13.8 (2.7-NR)	0.314
Age	< 65 years	87	10.8 (7.2-15.7)		36	12.1 (7.2-NR)	
	>= 65 years	58	16.1 (12.4-18.8)	0.261	29	16.6 (8.5-NR)	0.716
Patients with a poor prognosis		137	6.1 (4.9-7.7)	–	44	6.5 (3.4-10.0)	–
Fulfilling SUTENT trial eligibility criteria	No	67	4.7 (3.3-6.9)		18	3.5 (1.3-7.8)	
	Yes	70	6.8 (5.3-10.7)	0.015	26	6.6 (3.8-NR)	0.072
Brain metastases	No	125	6.5 (5.3-8.4)		40	6.5 (3.8-10.1)	
	Yes	12	2.1 (0.7-4.2)	0.006	4	1.2 (0.8-NR)	0.013
WHO performance status	0-1	105	6.9 (5.3-9.8)		36	6.6 (3.8-10.1)	
	2-4	32	3.1 (1.4-5.5)	<0.001	8	1.2 (0.6-7.5)	0.009
Histology	Clear cell	93	6.1 (4.6-7.8)		32	6.5 (2.7-10.1)	
	Non-clear cell	44	5.7 (3.7-10.3)	0.659	12	4.1 (2.6-NR)	0.998
Age	< 65 years	75	6.9 (4.9- 9.8)		28	7.8 (3.8-13.7)	
	>= 65 years	62	5.4 (3.8-6.8)	0.404	16	3.2 (1.1-6.6)	0.026

Abbreviations: OS, overall survival; CI, confidence interval; NR, not reached.
* Since all patients in the 2008–2010 Cohort presented with metastatic disease, none of the patients had a favourable prognosis (i.e. time from initial RCC diagnosis was less than one year).

Uptake of targeted therapies and their use in daily clinical practice (2011–2013 Cohort)

Table 2.2 shows the first-line therapies used in the 2011–2013 Cohort. During the follow-up period, 139/233 (60%) patients received a first-line therapy; the majority (110, 79%) was treated with sunitinib. The distribution of patients across first-line therapies over time (half-year periods) is presented in Figure 2.2. There were no significant differences between the half-year periods in the proportion of patients receiving targeted therapies. However, amongst treated patients, there was a decreasing trend in sunitinib use over time (p=0.006) and an increasing trend in pazopanib use (p=0.001).

Thirty-seven patients also received a second-line therapy within the follow-up period. The majority was treated with everolimus (57%), but a decreasing trend in everolimus use over time was observed (p=0.002).

Use of targeted therapies amongst patients with a favourable or intermediate prognosis (2011–2013 Cohort)

136/233 patients (58%) had a favourable or intermediate prognosis. 52/136 patients (38%) did not receive any targeted therapy within the follow-up period. However, 12 of these 52 patients received a metastasectomy (combined with a nephrectomy). 25 of the remaining 40 patients (63%) who were given neither targeted therapy nor a metastasectomy (combined with a nephrectomy) fulfilled the SUTENT trial eligibility criteria. In addition, 45 of the 66 patients treated with sunitinib fulfilled the SUTENT trial eligibility criteria.

Amongst patients fulfilling SUTENT trial eligibility criteria, males (OR, 0.12; p=0.020) and older patients (OR, 0.92; p=0.011) were less likely to receive sunitinib after adjustment for additional patient and disease characteristics (see frequencies in Table 2.3).

Median OS of eligible patients not receiving any targeted therapy was 20.9 months (95% CI 7.4-not reached [NR]). Table 2.4 presents the median OS in subgroups of patients with a favourable or intermediate prognosis treated with first-line sunitinib. Median OS of eligible patients treated with sunitinib was 18.0 months (95% CI 10.1-NR). The mean time from diagnosis to start of first-line sunitinib was 2.1 months (SD 3.3).

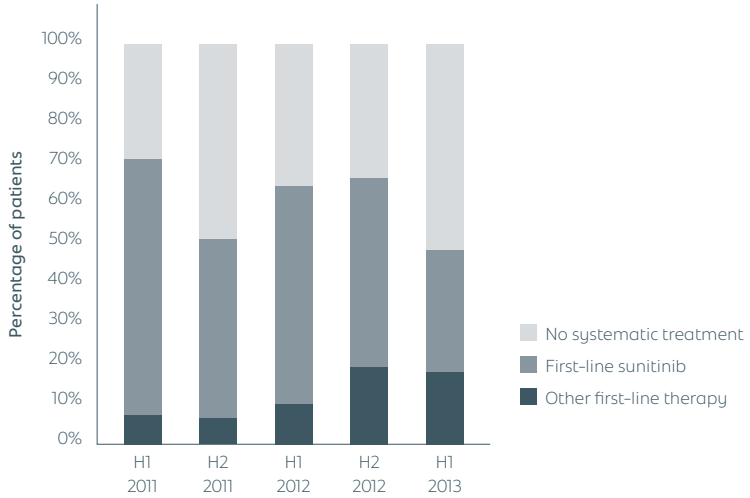
Median OS was 10.9 months (95% CI 2.7-NR) for patients treated with sunitinib but not fulfilling SUTENT trial eligibility criteria. No significant differences were observed within subgroups.

Use of targeted therapies amongst patients with a poor prognosis (2011–2013 Cohort)

97/233 patients (42%) had a poor prognosis. Forty-two patients (43%) did not receive any targeted therapy; thirteen of these 42 patients (31%) fulfilled the SUTENT trial eligibility criteria. Of the 44 patients treated with sunitinib, 26 fulfilled the SUTENT trial eligibility criteria.

Of patients fulfilling SUTENT trial eligibility criteria, older patients (OR, 0.84; p=0.012) were less likely to receive sunitinib (see frequencies in Table 2.3). The unadjusted model showed a significant association between hospital of diagnosis and the prescription of sunitinib, but this association disappeared after adjustment for demographics, clinical and laboratory factors.

Figure 2.2 Use of first-line drugs over time per half a year (2011–2013 Cohort)



Note: Patients were classified using the date of diagnosis.
Abbreviations: H1, first half year; H2, second half year

Median OS of eligible patients not receiving any targeted therapy was 3.4 months (95% CI 0.8-NR). Table 2.4 shows the median OS in subgroups of patients with a poor prognosis treated with first-line sunitinib. Median OS of eligible patients treated with sunitinib was 6.6 months (95% CI 3.8-NR). The mean time from diagnosis to start of first-line sunitinib was 1.9 months (SD 1.8).

Median OS was significantly reduced in patients not fulfilling the SUTENT trial eligibility criteria (3.5 months, 95% CI 1.3–7.8). Additionally, as in the 2008–2010 Cohort, median OS was significantly reduced in patients with brain metastases and patients with a WHO performance status of 2–4. OS was also significantly reduced in older patients.

Discussion

Since 2006, several new targeted therapies for mRCC have entered the market and randomised controlled trial (RCTs) have shown that these therapies improve survival.[25-29,34,40-48] This study examined the uptake and use of targeted therapies in the Netherlands. Not unexpected, targeted therapies, sunitinib in particular, have largely replaced IFN-α as first-line standard of care. Few patients were treated with bevacizumab (combined with IFN-α) or temsirolimus in the 2008–2013 period, even though these therapies were added to the ESMO guidelines in 2009,[30] and to Dutch guidelines in 2010.[32] Pazopanib has only been recommended since 2010,[31] which partly explains why an increase in its use was only seen from 2012. Furthermore, there was a shift in the use

of second-line therapies, where sorafenib was replaced by everolimus as the most frequent choice from 2010 onwards.

The median OS of patients with an intermediate prognosis treated with sunitinib in Dutch daily practice and fulfilling the SUTENT trial eligibility criteria was shorter than the median OS of patients in the SUTENT trial with an intermediate prognosis, i.e. 14.8 months (95% CI 10.8–16.1) in the 2008–2010 Cohort compared to 20.7 months (95% CI 18.2–25.6) in the SUTENT trial.[34] However, the difference was much smaller for the 2011–2013 Cohort (median OS, 18.0 months (95% CI 10.1-NR)) compared to the SUTENT trial patients. Median OS of patients with a poor prognosis fulfilling the SUTENT trial eligibility criteria was similar to the median OS found in the SUTENT trial, i.e. 6.8 months (95% CI 5.3–10.7) in the 2008–2010 Cohort and 6.6 months (95% CI 3.8-NR) in the 2011–2013 Cohort compared to 5.3 months (95% CI 4.2–10.0) in the SUTENT trial.[34]

The median OS of patients with an intermediate prognosis treated with sunitinib in Dutch daily practice (regardless of their SUTENT trial eligibility status) was shorter than the OS in the expanded-access trial.[35] Median OS of patients with a poor prognosis was in line with the results of the expanded-access trial. The median OS of patients with an intermediate prognosis treated with sunitinib in Dutch daily practice was also shorter than the OS in a retrospective, non-interventional study in Australia.[49] These findings may indicate that the patients in the PERCEPTION registry with an intermediate risk had a worse prognosis than the patients with an intermediate risk in other studies.

While previous studies suggest that patients fulfilling SUTENT trial eligibility criteria have a survival benefit from first-line sunitinib,[34] many eligible patients did not receive sunitinib (or any other targeted therapy) in daily practice. This was also seen in England where one in three patients with mRCC eligible for either sunitinib or pazopanib did not receive the drug.[50] Patients aged 65+ years were less likely to receive targeted therapy than younger patients after adjustment for other factors. This age factor was found in patients with an intermediate prognosis (2011–2013 Cohort) and in patients with a poor prognosis (2008–2010 Cohort and 2011–2013 Cohort). There are several explanations for this association, including medical contraindications, other grounds for physician reluctance, and patient refusal. Additionally, patients with one metastatic site were less likely to receive sunitinib (according to the 2008–2010 Cohort results), which might be explained by patients with low volume but unresectable metastases whose targeted therapy is delayed. Nevertheless, most of these patients died within the follow-up period without receiving targeted therapy at any point in time. The reasons for apparent underutilisation of targeted therapies should be examined more carefully. While hospital-level factors may also affect utilisation and lead to between-hospital

variation, we found no significant differences in the prescription of targeted therapy between hospitals, except for the patients with a poor prognosis in the 2008–2010 Cohort. However, the sample size per hospital was small and the statistical power to show a difference was therefore limited.

Although this study mainly focussed on patients fulfilling SUTENT trial eligibility criteria, we found that many patients in daily clinical practice are different from patients included in RCTs. In the total study population, only 42% and 58% fulfilled the SUTENT trial eligibility criteria in the 2008–2010 Cohort and 2011–2013 Cohort, respectively. This was partly caused by the inclusion criteria of the PERCEPTION registry, which consisted of a diagnosis of mRCC (i.e. metastases at initial presentation in the 2008–2010 Cohort) of any histological subtype. Since many patients are excluded from clinical trials, such as patients with a non clear-cell subtype, patients with a WHO performance status of 2 to 4 and patients with brain metastases, one could argue that the results of these trials only apply to a subgroup of patients.

A limitation of this study is the amount of missing data in baseline characteristics, which is inherent to an observational study. To overcome this problem, multiple imputations by chained equations were conducted, which ensure that all patients are included in the analysis but simultaneously ensure that the uncertainties from missing data are retained.[39] Additionally, eligibility criteria, such as the presence of measurable disease and adequate organ function were not taken into account when determining whether patients fulfilled the SUTENT trial eligibility criteria, since data on these criteria were lacking in the PERCEPTION registry. As a consequence, some of the patients that we labelled as eligible in this study were not in fact eligible for targeted therapy. However, since we used WHO performance status to classify patients, and since we expect a relationship between WHO performance status and organ function, we believe that this could only have had a limited effect on our conclusions about the uptake and use of targeted therapies. Furthermore, the follow-up length of the 2011–2013 Cohort was limited. As a consequence, patients might have received targeted therapy after the follow-up period, leading to an underestimate of actual targeted therapy use. However, this limitation is only relevant for patients treated later in the 2011–2013 period who did not die. Lastly, OS was calculated from the date of diagnosis (i.e. metastatic disease) for patients not receiving any targeted therapy and from the start of therapy for patients treated with targeted therapy; as a consequence a comparison between the two is impossible. This approach was based on the one used in other studies to enable comparisons between the OS of patients treated with sunitinib in our study with the OS of patients treated with sunitinib in other studies.[34,35,49]

Conclusions

In conclusion, targeted therapies, sunitinib in particular, have largely replaced IFN-α as the first-line standard of care in the Netherlands. Nevertheless, many patients in Dutch daily practice fulfilling SUTENT trial eligibility criteria did not receive sunitinib (or any other targeted therapy) even though it could improve their survival. For example, older patients were less likely to receive sunitinib, perhaps because physicians are reluctant to prescribe it. The reasons for apparent underutilisation of targeted therapies should be examined more carefully.

Figure S2.1 Patient enrolment

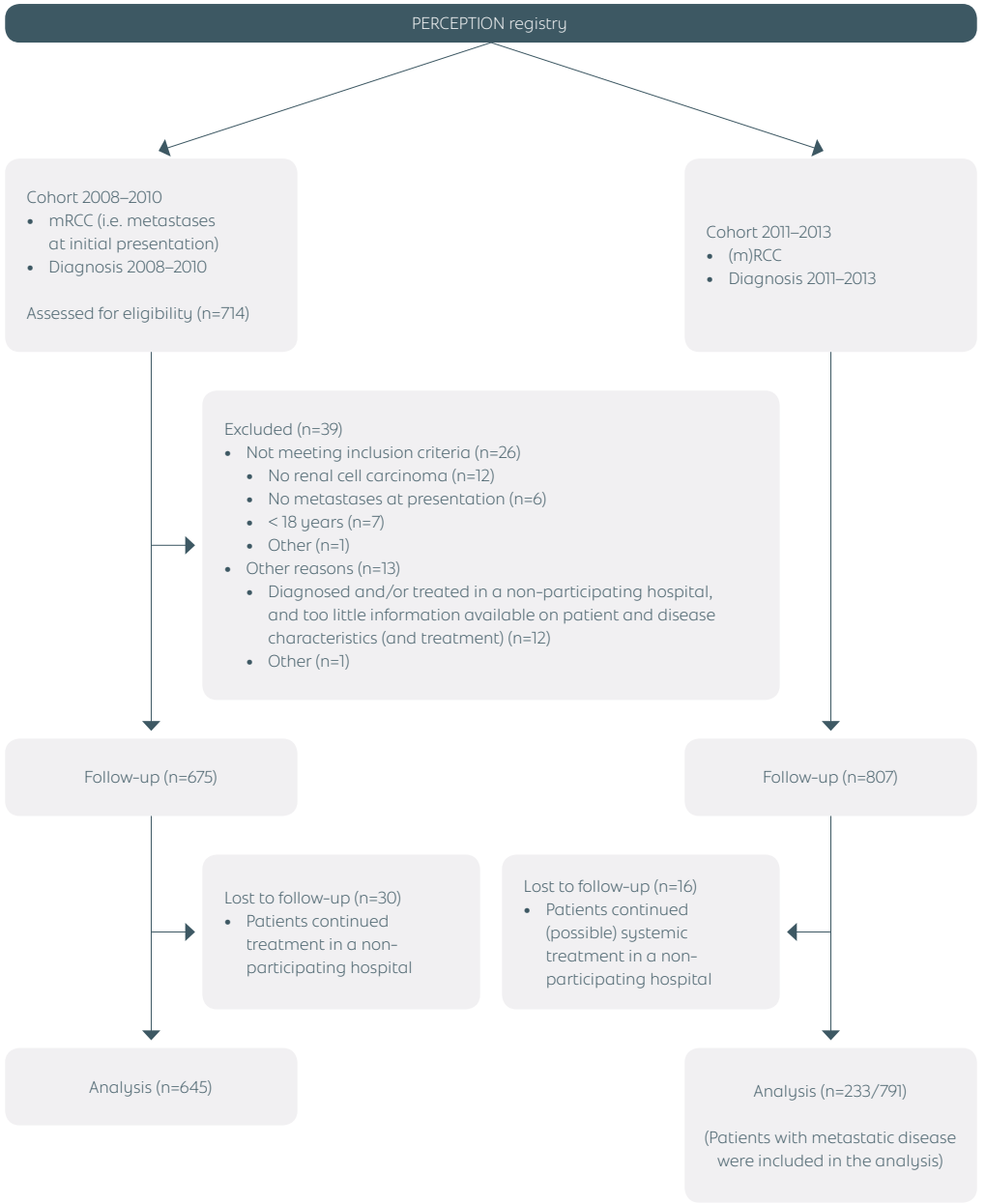


Table S2.1 Patient and disease characteristics (observed and imputed) 2008–2010 Cohort and 2011–2013 Cohort

	2008–2010 Cohort: mRCC at the initial diagnosis				2011–2013 Cohort: mRCC			
	Real-world data (n=621)		Imputed data (n=621)		Real-world data (n=221)		Imputed data (n=221)	
Sex – n (%)								
Female	213	34%	213	34%	60	27%	60	27%
Male	408	66%	408	66%	61	73%	161	73%
Median age – yr (range)	66	23-93	66	23-93	66	27-93	66	27-93
Histology – n (%)								
Clear cell	354	57%	354	57%	152	69%	152	69%
Other	267	43%	267	43%	69	31%	69	31%
WHO performance status – n (%)								
0-1	204	33%	430	69%	94	43%	178	81%
2-4	61	10%	191	31%	13	6%	42	19%
Missing	356	57%			114	52%		
Site of metastasis – n (%)								
One	195	31%	206	33%	85	38%	87	39%
More than one	398	64%	415	67%	131	59%	134	61%
Missing	28	5%			5	2%		
Liver metastasis – n (%)								
No	487	78%	509	82%	171	77%	175	79%
Yes	106	17%	112	18%	45	20%	46	21%
Missing	28	5%			5	2%		
Lung metastasis – n (%)								
No	163	26%	173	28%	72	33%	74	33%
Yes	430	69%	448	72%	144	65%	147	67%
Missing	28	5%			5	2%		
Bone metastasis – n (%)								
No	375	60%	393	63%	154	70%	158	71%
Yes	218	35%	228	37%	62	28%	63	29%
Missing	28	5%			5	2%		
Brain metastasis – n (%)								
No	546	88%	571	92%	200	90%	200	90%
Yes	47	8%	50	8%	16	7%	16	7%
Missing	28	5%			5	2%		
Haemoglobin – n (%)								
Normal	171	28%	205	33%	76	34%	85	38%
< LLN	347	56%	416	67%	122	55%	136	62%
Missing	103	17%			23	10%		
Neutrophil count – n (%)								
Normal	203	33%	383	62%	82	37%	152	69%
> ULN	108	17%	238	38%	41	19%	69	31%
Missing	310	50%			98	44%		
Platelet count – n (%)								
Normal	358	58%	452	73%	127	57%	159	72%
> ULN	140	23%	169	27%	51	23%	62	28%
Missing	123	20%			43	19%		
Albumin – n (%)								
Normal	247	40%	391	63%	86	39%	130	59%
< LLN	136	22%	230	37%	61	28%	91	41%
Missing	238	38%			74	33%		
Corrected serum calcium – n (%)								
Normal	243	39%	421	68%	88	40%	140	63%
> ULN	116	19%	200	32%	51	23%	81	37%
Missing	262	42%			82	37%		
Alkaline phosphatase – n (%)								
Normal	324	52%	432	70%	112	51%	152	69%
> ULN	139	22%	189	30%	48	22%	69	31%
Missing	158	25%			61	28%		
Lactate dehydrogenase – n (%)								
Normal	277	45%	372	60%	130	59%	179	81%
> 1.5 times ULN	174	28%	249	40%	31	14%	42	19%
Missing	170	27%			60	27%		
Comorbidities – n (%)								
Zero or one	356	57%	356	57%	151	68%	151	68%
More than one	265	43%	265	43%	67	30%	67	30%
Missing	0	0%			3	1%		
Time since RCC diagnosis – n (%)								
> One year	NA	NA	NA	NA	16	7%	16	7%
< One year	NA	NA	NA	NA	204	92%	204	92%

Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal; NA, not applicable.

Chapter 3

Survival in patients with primary metastatic renal cell carcinoma treated with sunitinib with or without previous cytoreductive nephrectomy: Results from a population-based registry

S. de Groot; W.K. Redekop; S. Sleijfer; E. Oosterwijk; A. Bex; L.A.L.M. Kiemeney; C.A. Uyl-de Groot

Urology 2016;95:121–127.

Abstract

Objective: To evaluate the effect of cytoreductive nephrectomy (CN) on overall survival (OS) in primary metastatic renal cell carcinoma (mRCC) patients treated with first-line sunitinib.

Methods: Patients with primary mRCC treated with first-line sunitinib were selected from a Dutch population-based registry. A propensity score was calculated reflecting the probability of a patient undergoing CN prior to sunitinib using a set of known covariates, such as the Memorial Sloan Kettering Cancer Center and International mRCC Database Consortium risk factors. After propensity score matching, differences in OS were analysed using the Kaplan-Meier method and a multivariable Cox proportional hazards model was used to evaluate the effect of CN on OS.

Results: A total of 227 patients met the selection criteria; 74 patients (33%) underwent CN prior to sunitinib. In the matched population, the median OS of patients who underwent CN was 17.9 months compared to 8.8 months for patients treated with sunitinib only. Multivariable analysis showed that CN was an independent predictor of OS (hazard ratio 0.61, 95% confidence interval 0.41-0.92). A subgroup analysis of patients with a time to targeted therapy of <1 year showed a median OS of 12.7 months for patients treated with CN compared to 8.0 months for patients treated with sunitinib only. The corresponding hazard ratio was 0.67 (95% confidence interval 0.46-0.98).

Conclusions: This study suggests that CN may be effective. However, the benefit was modest when correcting for time from diagnosis to sunitinib. One important limitation is the use of a registry (with retrospectively collected data), which made it impossible to correct for unmeasured characteristics that could be associated with treatment choices or survival.

Introduction

In Europe in 2012, an estimated 115,200 patients were diagnosed with kidney cancer, and 49,000 patients died from the disease.[18] The most common type of kidney cancer is renal cell carcinoma. Metastases at the initial presentation are present in 25%-30% of the patients.[51]

Following a randomised study showing that interferon-alfa (IFN-α) improved overall survival (OS) compared to medroxyprogesterone acetate and subsequent randomised studies that concluded that a cytoreductive nephrectomy (CN) further prolongs survival of patients with metastatic renal cell carcinoma (mRCC), a tumour nephrectomy followed by IFN-α has been the standard approach for many years.[52-54] However, treatment has considerably changed over the last decade. Targeted therapies for mRCC have entered the market since 2006, and sunitinib is nowadays the most frequently used treatment in the Netherlands.

Whether or not CN prolongs the survival of patients treated with targeted agents such as sunitinib is unknown given the lack of evidence from randomised controlled trials (RCTs). International guidelines recommend CN for patients with a good performance status and large primary tumours with limited volumes of metastatic disease, and for patients with a symptomatic primary lesion.[55,56] Clearly, this is based on a low level of evidence with a high risk of bias. As a consequence, clinicians remain uncertain about whether or not CN offers any benefit to mRCC patients who are about to start with sunitinib and have no complaints arising from their primary tumour.

The effect of CN compared to no surgery in primary mRCC was previously investigated in retrospective studies.[57-60] However, these studies included heterogeneous patient populations that were composed of patients from different countries treated with various systemic therapies. In addition, with a few exceptions, no attention was given as to when targeted therapy was provided. This is of importance since the time from diagnosis to targeted therapy of <1 year is a validated risk factor in the Memorial Sloan Kettering Cancer Center (MSKCC) and International mRCC Database Consortium (IMDC) risk scores.[21,23] Without correction, indiscriminate inclusion of patients with a time to targeted therapy of ≥1 year introduces substantial bias. The aim of this paper is to evaluate the effect of CN on OS in primary mRCC patients treated with first-line sunitinib. We specifically selected patients within 1 country, all treated with the same systemic treatment (i.e. sunitinib) within a limited time frame. A subgroup analysis was conducted including patients with a time to targeted therapy of <1 year.

Patients and methods

Study population

A population-based registry (PERCEPTION) was initiated to evaluate treatment of patients with mRCC in Dutch clinical practice. Eligible patients for this registry were selected from the Dutch Cancer Registry, which includes all new cancer cases in the Netherlands. Inclusion criteria for the PERCEPTION registry comprised a diagnosis of mRCC (i.e. metastases at initial presentation) of any histologic subtype. Patients diagnosed between January 2008 and December 2010 in 42 of 51 hospitals (both general and academic) in 4 regions, covering approximately half of the country, were included in the PERCEPTION registry. The research protocol was approved by the medical ethical committee of the Radboud university medical center in Nijmegen in 2010.

For the present study, we selected patients treated with first-line sunitinib to study the effect of CN on OS in a homogeneous patient population. We chose to select patients treated with first-line sunitinib because the majority of patients treated with a systemic therapy were treated with sunitinib (i.e. 84% in the PERCEPTION registry).

Variables and definitions

Data on baseline demographics and clinical and laboratory factors were retrospectively collected from individual patient records. Furthermore, data on treatment schemes and treatment endpoints (e.g. survival) were collected. Laboratory factors were standardised according to routinely used reference values.

Statistical analyses

To mimic an RCT on the effect of CN as good as possible, a propensity score was calculated reflecting the probability of a patient undergoing CN prior to sunitinib, using a set of known covariates and multivariable logistic regression analysis. The covariates were obtained from the PERCEPTION registry by selecting clinical, biochemical, and hematologic factors known to impact progression-free survival or OS (Table 3.1).[37,38] Also, baseline demographics (age at diagnosis and gender) and 3 additional clinical factors (histology, clinical tumour stage, and regional lymph node involvement) were incorporated as covariates, because these factors could have influenced the decision to conduct a nephrectomy. OS was calculated from the date of the start of treatment (i.e. CN or start of first-line sunitinib) until the date of death from any cause or last follow-up using the Kaplan-Meier method. The first approach to evaluate the effect of CN was a Cox proportional hazards model using the propensity score as a covariate.

In the second approach, we applied 2 different methods to match the treated patients (i.e. patients who underwent CN prior to sunitinib) to comparable controls (i.e. patients treated with sunitinib only): single nearest-neighbour matching (with and without caliper, with replacement) and kernel matching.[61] Single nearest-neighbour matching involves matching each treated patient to the untreated patient with the closest propensity score. Kernel matching involves matching all treated patients with a weighted average of all untreated patients.[62] In this way, a hypothetical match is constructed for each patient who underwent CN prior to sunitinib. Patient characteristics were studied to decide whether single nearest-neighbour matching or kernel-matching method achieved the best balance of covariates. Then, a multivariable Cox proportional hazards model was estimated in the matched population, as recommended by Stuart.[61] Covariates considered for inclusion were baseline demographics (age at diagnosis and gender), and clinical, biochemical, and hematologic factors (Table 3.1).[37,38] Backward selection was used to select the covariates for the model; any nonsignificant attributes (α=0.05) were excluded from the model one at a time. Forward selection was used to create an alternative model. The proportionality assumption was tested graphically using log–log plots of survival, and formally using Schoenfeld residuals.

The analyses were repeated for a subgroup of patients with a time to targeted therapy of <1 year. Again, propensity score matching was applied and a multivariable Cox proportional hazards model was estimated to evaluate the effect of CN on OS.

Missing data were handled using multiple imputations by chained equations. This method generated imputations based on a set of imputation models, one for each variable with missing values.[39] The process of multiple imputations produced 10 imputed datasets. The propensity score used in the analysis was the mean of 10 propensity scores for each patient.[63]

Data analyses were conducted using STATA statistical analysis software (StataCorp. 2013, Stata Statistical Software: Release 13. College Station, TX: StataCorp LP).

Results

Study population

Out of 714 patients with primary mRCC in the PERCEPTION registry, 282 patients treated with first-line sunitinib were selected. Finally, 227 patients were included in the analysis of whom 74 patients (33%) underwent CN prior to sunitinib (reasons for exclusion are presented in Figure S3.1).

Table 3.1 Clinical factors, biochemical and haematological factors known to impact mRCC outcomes

Clinical factors	Biochemical and haematologic factors
WHO performance status	Corrected serum calcium >ULN
Time from diagnosis to study start*	Neutrophil count >ULN
Number of metastatic sites	Platelet count >ULN
Prior nephrectomy**	Alkaline phosphatase >ULN
Lung metastases	Lactate dehydrogenase >ULN
Liver metastases	Haemoglobin <LLN
Bone metastases	Albumin <LLN

* Since all patients who did not undergo cytoreductive nephrectomy had a time from diagnosis to treatment of less than 1 year, this factor was excluded during the calculation of propensity scores.
** Prior nephrectomy was excluded during the calculation of propensity scores.
Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal.

Patient characteristics are shown in Table 3.2. Briefly, the size or extent of the primary tumour was more often limited to T1, T2, or T3a for patients who underwent CN. CN-treated patients more often had just 1 metastatic site and had bone metastases less often than patients who did not undergo CN. Lastly, patients who underwent CN more often had a normal corrected serum calcium level, and their time from diagnosis to start sunitinib was more often ≥1 year. Observed patient and disease characteristics (without imputations) are provided in Table S3.1.

Of patients who underwent CN, 3% and 54% had a favourable or intermediate prognosis, respectively, whereas 0% and 46% of patients treated with sunitinib only had a favourable or intermediate prognosis according to the MSKCC risk score.

All but 1 of the 74 patients underwent a radical nephrectomy, either performed by an open (n=38) or laparoscopic (n=27) approach; for 7 patients, the technique was unknown and 1 patient underwent an open partial nephrectomy. Eight of 74 patients also underwent an associated procedure such as an adrenalectomy during the same session. The median time from diagnosis to nephrectomy was 0 days (range: 0-135). The date of diagnosis for each patient was derived from the Dutch Cancer Registry. In this registry, dates of diagnosis are registered in a systematic way in the sense that a clinical diagnosis date is overruled with the first date at which histologic confirmation takes place. For that reason, the diagnosis date is the date of CN for more than 50% of our patient series. Median time from nephrectomy to start of first-line sunitinib was 85 days (range: 11-1048). For patients who did not undergo a nephrectomy prior to sunitinib, the median time from diagnosis to start of first-line sunitinib was 27 days (range: 0-352).

At last follow-up, 207 of 227 patients had died. Figure 3.1 shows the Kaplan-Meier curves of OS in patients treated with CN prior to sunitinib and patients treated with sunitinib only. The unadjusted median OS of patients who underwent CN was 16.7 months (95% confidence interval [CI] 11.7-20.2) compared to a median OS of 6.8 months (95% CI 5.8-8.8) for patients treated with sunitinib only ($p<0.001$). The corresponding hazard ratio (HR) was 0.53 (95% CI 0.39-0.72).

Matched population

When the propensity score was included as the sole covariate in a Cox proportional hazards model, CN was associated with an HR of 0.56 (95% CI 0.40-0.80) (Table 3.3).

Besides estimating a Cox proportional hazards model using the propensity score as a covariate, single nearest-neighbour matching (with and without caliper, with replacement) and kernel matching were applied. Differences in patient characteristics were noticeably reduced after single nearest-neighbour matching and kernel matching with kernel matching showing the largest reductions. Patient characteristics of the kernel-matched population are presented in Table 3.2. Differences in size or extent of the primary tumour, the number of metastatic sites, and the presence of bone metastases were noticeably reduced. Also, the percentages of patients with a normal corrected serum calcium level became comparable after matching. Figure 3.2 shows the Kaplan-Meier curves after matching. The median OS of patients who underwent a nephrectomy prior to sunitinib was 17.9 months compared to a median OS of 8.8 months for patients treated with sunitinib only. After matching, the multivariable Cox proportional hazard model showed that CN had an HR of 0.61 (95% CI 0.41-0.92) (Table 3.3).

A subgroup analysis of patients with a time to targeted therapy of <1 year showed a median OS of 12.7 months for patients treated with CN prior to sunitinib compared to 8.0 months for patients treated with sunitinib only. The corresponding HR associated with CN was 0.67 (95% CI 0.46-0.98) after adjustment for other factors. Figure 3.3 shows the Kaplan-Meier curves of OS after matching in this subgroup of patients.

Discussion

Two RCTs have found that CN followed by treatment with IFN-α improves survival compared to IFN-α alone in patients presenting with mRCC and a performance status of 0 or 1.[53,54] In contrast, there is no RCT evidence supporting the value of a CN prior to targeted therapies. This has led to a large variation in clinical practice, reflected by our data showing that 74 of 227 of the patients (33%) presenting with mRCC underwent CN prior to sunitinib.

The median OS of patients who underwent CN prior to sunitinib was 17.9 months compared to 8.8 months for matched patients treated with sunitinib

Table 3.2 Patient characteristics before and after propensity score matching

	Unmatched data			Matched data	
	Sunitinib only (n=153)	CN + sunitinib (n=74)	p-value	Sunitinib only (n=73)	CN + sunitinib (n=73)*
Gender – n (%)					
Female	48 (31%)	13 (18%)		13 (17%)	13 (18%)
Male	105 (69%)	61 (82%)	0.030	60 (83%)	60 (82%)
Median age – yr (range)	64 (24-89)	61 (28-77)	0.010	64 (24-89)	62 (28-77)
Histology – n (%)					
Clear cell	101 (66%)	58 (78%)		57 (78%)	57 (78%)
Other	52 (34%)	16 (22%)	0.059	16 (22%)	16 (22%)
cTNM – T – n (%)					
T1-T3a	103 (67%)	61 (82%)		59 (81%)	60 (82%)
T3b-T4	50 (33%)	13 (18%)	0.020	14 (19%)	13 (18%)
cTNM – n – n (%)					
N0	71 (46%)	37 (50%)		32 (44%)	37 (50%)
N1	82 (54%)	37 (50%)	0.657	41 (56%)	36 (50%)
WHO performance status – n (%)					
0-1	126 (82%)	58 (79%)		60 (82%)	57 (78%)
2-4	27 (18%)	16 (21%)	0.698	13 (18%)	16 (22%)
Site of metastasis – n (%)					
One	39 (25%)	34 (46%)		39 (54%)	33 (45%)
More than one	114 (75%)	40 (54%)	0.002	34 (46%)	40 (55%)
Liver metastasis – n (%)					
No	129 (84%)	68 (92%)		68 (94%)	67 (92%)
Yes	24 (16%)	6 (8%)	0.120	5 (6%)	6 (8%)
Lung metastasis – n (%)					
no	36 (24%)	16 (22%)		24 (32%)	16 (22%)
yes	117 (76%)	58 (78%)	0.749	49 (68%)	57 (78%)
Bone metastasis – n (%)					
No	88 (58%)	59 (80%)		61 (84%)	58 (79%)
Yes	65 (42%)	15 (20%)	0.001	12 (16%)	15 (21%)
Haemoglobin – n (%)					
Normal	53 (34%)	20 (27%)		17 (23%)	20 (27%)
< LLN	100 (66%)	54 (73%)	0.264	56 (77%)	53 (73%)
Neutrophil count – n (%)					
Normal	96 (63%)	51 (68%)		49 (67%)	50 (68%)
> ULN	57 (37%)	23 (32%)	0.568	24 (33%)	23 (32%)
Platelet count – n (%)					
Normal	109 (71%)	58 (79%)		58 (79%)	57 (79%)
> ULN	44 (29%)	16 (21%)	0.257	15 (21%)	16 (21%)
Albumin – n (%)					
normal	100 (66%)	56 (75%)		51 (70%)	55 (75%)
< LLN	53 (34%)	18 (25%)	0.217	22 (30%)	18 (25%)
Corrected serum calcium – n (%)					
Normal	94 (62%)	59 (79%)		57 (78%)	58 (79%)
> ULN	59 (38%)	15 (21%)	0.037	16 (22%)	15 (21%)
Alkaline phosphatase – n (%)					
Normal	108 (70%)	58 (78%)		59 (81%)	57 (78%)
> ULN	45 (30%)	16 (22%)	0.291	14 (19%)	16 (22%)
Lactate dehydrogenase – n (%)					
Normal	91 (59%)	45 (61%)		39 (54%)	45 (62%)
> 1.5 times ULN	62 (41%)	29 (39%)	0.781	34 (46%)	28 (38%)
Time from diagnosis to start sunitinib – n (%)					
>= One year	0 (0%)	12 (16%)	<0.001	0 (0%)	12 (16%)
< One year	153 (100%)	62 (84%)		153 (100%)	62 (84%)

Abbreviations: CN, cytoreductive nephrectomy; LLN, lower limit of normal; ULN, upper limit of normal.
* One patient was excluded, since his propensity score was higher than the maximum propensity scores of the controls.

Table 3.3 Multivariable analysis for factors associated with overall survival (based on covariate adjustment using the propensity score and based on kernel-matching)

Covariate	HR nephrectomy adjusted for propensity score (95% CI)	Multivariable HR in the matched population (95% CI)*
Cytoreductive nephrectomy (yes vs. no)	0.56 (0.40-0.80)	0.61 (0.41-0.92)
Propensity score	0.74 (0. 34-1.62)	
WHO performance status (2-4 vs. 0-1)		2.26 (1.04-4.95)
Liver metastasis (yes vs. no)		2.30 (1.14-4.62)
Corrected serum calcium (>ULN vs. normal)		1.87 (1.01-3.46)
Time from diagnosis to start sunitinib (< 1 year vs. >= 1 year)		4.45 (1.59-12.42)

Abbreviations: HR, hazard ratio; CI, confidence interval; ULN, upper limit of normal.
* A slightly different model was created based on forward selection with platelet count instead of WHO performance status. Variables considered for inclusion were: cytoreductive nephrectomy, gender, age, WHO performance status, site of metastasis, liver metastasis, lung metastasis, bone metastasis, haemoglobin, neutrophil count, platelet count, albumin, corrected serum calcium, alkaline phosphatase, lactate dehydrogenase and time from diagnosis to start sunitinib.

only. A multivariable Cox proportional hazard model showed that CN had an HR of 0.61 (95% CI 0.41-0.92), suggesting that CN may be effective. A comparable HR was found when the propensity score was included as the sole covariate in a Cox proportional hazards model. A subgroup analysis of patients with a time to targeted therapy of <1 year showed a median OS of 12.7 months for patients treated with CN prior to sunitinib compared to 8.0 months for matched patients treated with sunitinib only. The corresponding HR associated with CN was 0.67 (95% CI 0.46-0.98) after adjustment for other factors. All approaches indicate that CN is associated with improved survival. However, the benefit was modest when correcting for time from diagnosis to sunitinib.

Similar to our study, 3 other studies found an OS benefit in patients who underwent a nephrectomy prior to targeted therapies (HR 0.60, 95% CI 0.52-0.69; HR 0.68, 95% CI 0.46-0.99; HR 0.43, 95% CI 0.28-0.68, respectively).[57-59] In the studies by Heng et al and Choueiri et al, HRs were adjusted for the IMDC prognostic factors.[57,58] Similar to our study, the study by Bamias et al took additional factors into account besides the IMDC prognostic factors.[59] In all 3 studies, a multivariable Cox proportional hazard model was used to correct for differences in baseline characteristics between treatment groups. An advantage of our study is that propensity score matching was performed before fitting a model. As such, the design of the study was separated from the analyses; we estimated the effect of CN on OS only after patient characteristics were sufficiently balanced.[64,65] Furthermore, problems from a misspecified regression model are less severe when propensity score matching precedes conventional regression and characteristics have been balanced.[65] Lastly, all 3 studies included heterogeneous patient populations that were composed of patients from different countries treated with various systemic therapies (except the study by Bamias et al that selected patients treated with first-

Figure 3.1 Kaplan-Meier survival estimate of overall survival in primary mRCC patients treated with or without cytoreductive nephrectomy – observed patients

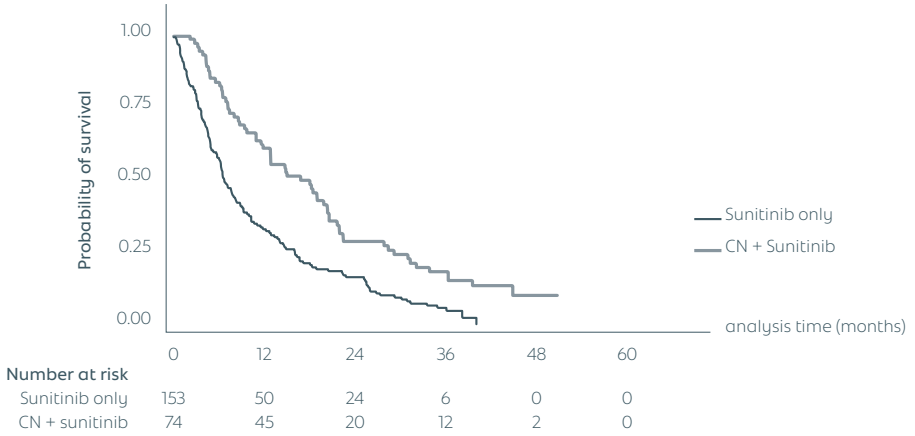


Figure 3.2 Kaplan-Meier survival estimate of overall survival in primary mRCC patients treated with or without cytoreductive nephrectomy – propensity score matched patients

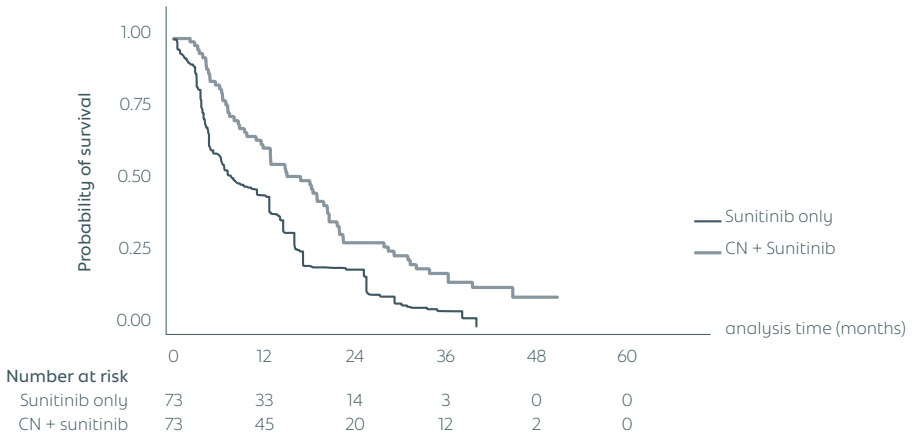
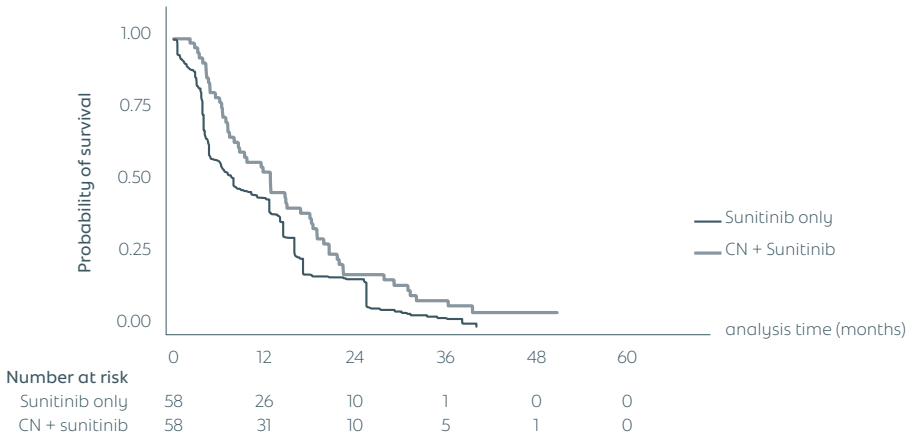


Figure 3.3 Kaplan-Meier survival estimate of overall survival in a subgroup of primary mRCC patients (time to sunitinib < 1 year) treated with or without cytoreductive nephrectomy – propensity score matched patients



line sunitinib), whereas we selected primary mRCC patients from a Dutch population-based registry, all treated with first-line sunitinib. Warren et al also showed an improved OS associated with CN prior to treatment with tyrosine kinase inhibitors (HR 0.38, 95% CI 0.19-0.74).[60] However, Warren et al's study population comprised a mix of patients (because only some patients presented with mRCC) and treatments (tyrosine kinase inhibitor could be the first- or second-line therapy). Besides these studies, other studies have examined the effect of presurgical sunitinib.[66-70]

The median OS of patients treated with sunitinib in the randomised phase III trial of sunitinib vs IFN-α was much longer than the OS observed in the current study. The pivotal phase III trial reported a median OS of 26.4 months,[34] whereas we found an OS of 16.7 months for patients who underwent CN (measured from the date of CN) and 6.8 months for patients treated with sunitinib only (in the unmatched population). This difference can be explained by the fact that the pivotal trial had many more patients with a favourable prognosis than our study (i.e. 38% vs 1%). In addition, our study exclusively focuses on primary mRCC, whereas the phase III trial mainly included patients with metachronous mRCC.

One important limitation of our study is the use of a population-based registry (with retrospectively collected data), which made it impossible to correct for unmeasured characteristics that could be associated with treatment choices or survival. Only a prospective, randomised trial could overcome this limitation.

A second limitation stems from our sample size of 74 patients who underwent CN prior to sunitinib. Although this relatively small sample resulted in fairly wide CIs for the HRs for CN, the upper limits of the CIs are lower than one, suggesting that CN may be effective.

Third, patients who underwent CN prior to sunitinib more often received a second-line therapy (27 of 74 patients [36%]) than patients treated with sunitinib only (34 of 153 patients [22%]). But since this difference is relatively small, it does not likely explain the difference in OS.

In addition, reasons for starting (or delaying) treatment with sunitinib have not been registered. Especially patients who underwent CN prior to sunitinib had a longer time from diagnosis to sunitinib (16%, ≥1 year). It is common practice in low-volume primary metastatic disease to perform a CN and to delay systemic therapy until further progression. This finding was also seen in the study by Heng et al, wherein 5% of the patients treated with sunitinib only and 29% of the patients who underwent CN prior to sunitinib had a time from diagnosis to targeted therapy of ≥1 year.[57] We conducted a subgroup analysis of patients with a time to targeted therapy of <1 year.

A further limitation results from our patient selection procedure. Specifically, since we selected patients treated with first-line sunitinib, we excluded patients who underwent a nephrectomy and were intended to receive sunitinib thereafter but ultimately did not receive it. Patients not receiving sunitinib could vary widely in prognosis from patients with deterioration of performance due to surgery to patients with low-volume metastatic disease without postoperative progression. A combined analysis of 2 prospective randomised trials comparing CN plus IFN-α with IFN-α alone revealed that only a rather low percentage of patients ended up not receiving immunotherapy (5.6% and 1.8%, respectively).[71] We therefore expect the exclusions of these patients from the analyses to have very little impact on the results, indicating that CN is associated with improved survival.

On the contrary, selecting patients treated with first-line sunitinib also provided advantages. As a result of this selection procedure, the study population consisted of relatively homogeneous patients. Especially in the subgroup analysis, patients were relatively homogeneous, since all patients had a time to sunitinib <1 year. Results of this analysis suggested that the benefit of CN is potentially less prominent than previously thought. This particular patient population is likely to represent the patients being included in the currently ongoing Clinical Trial to Assess the Importance of Nephrectomy (CARMENA). In the CARMENA trial, patients presenting with mRCC are being randomised to either nephrectomy followed by sunitinib or sunitinib alone. The final data collection date for the primary outcome measure of the CARMENA trial is expected in September 2019, meaning that the results will not be reported until perhaps 2020.[72] Furthermore, the role of presurgical targeted therapy is being evaluated in the SURTIME trial. However, this trial addresses a different research question compared to the research question we have addressed; that is, patients in the SURTIME trial are being randomised to immediate nephrectomy followed by sunitinib or to 3 cycles of presurgical sunitinib followed by nephrectomy.[73]

Conclusions

In accordance with other studies,[57-60] our study showed that CN was associated with a low HR in primary mRCC patients, suggesting that CN may be effective. However, the benefit was modest when correcting for time from diagnosis to sunitinib. While we are waiting for the results of the ongoing RCTs to be known, the potential advantages of CN combined with the current evidence support the practice of performing CN prior to treatment with sunitinib in patients presenting with mRCC.

Figure S3.1 Patient enrolment

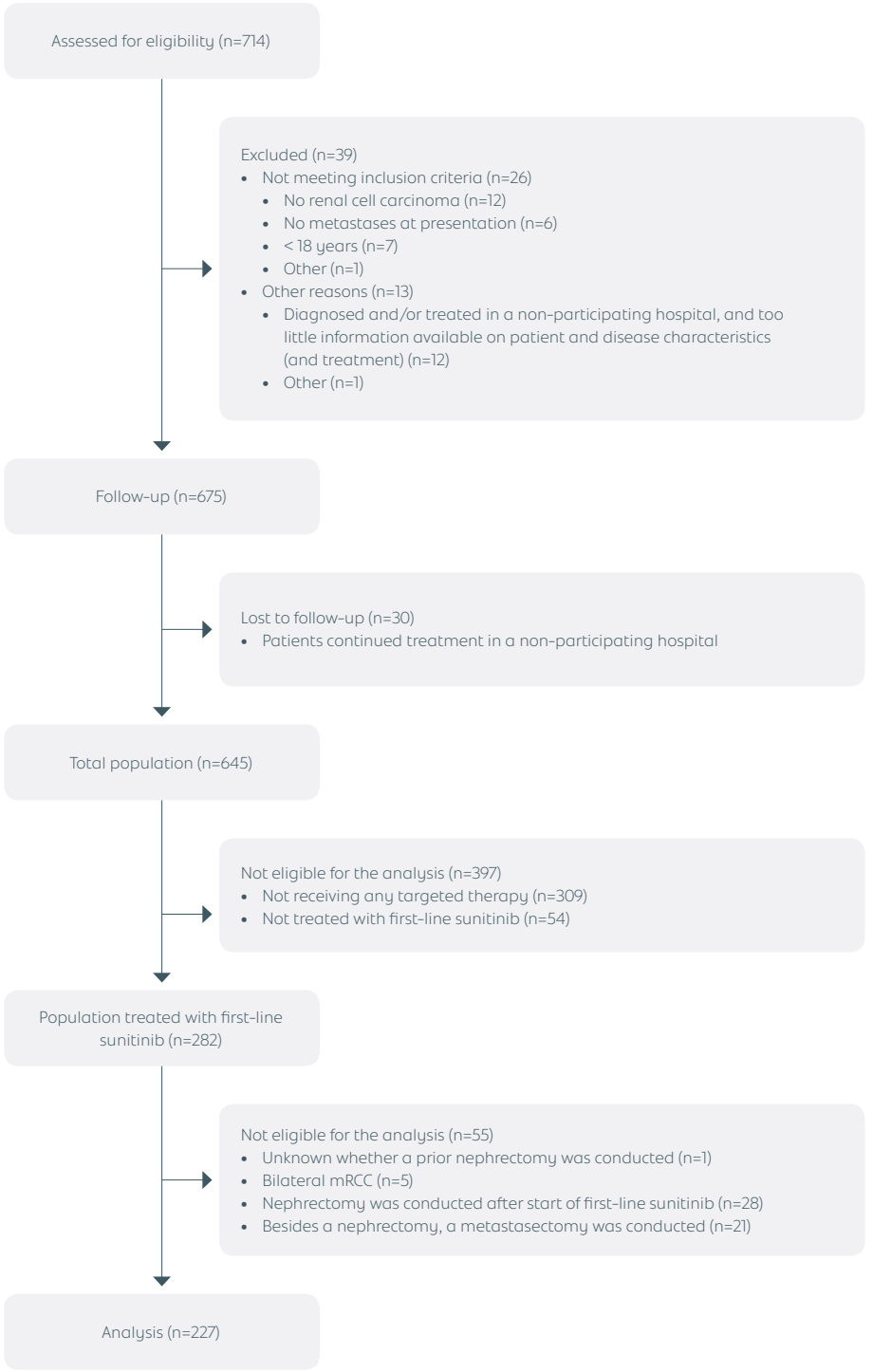


Table S3.1 Patient characteristics (observed and imputed)

	Observed data		Imputed data		
	Sunitinib only (n=153)	CN + sunitinib (n=74)	Sunitinib only (n=153)	CN + sunitinib (n=74)	p-value
Gender - n (%)					
Female	48 (31%)	13 (18%)	48 (31%)	13 (18%)	
Male	105 (69%)	61 (82%)	105 (69%)	61 (82%)	0.030
Median age – yr (range)	64 (24-89)	61 (28-77)	64 (24-89)	61 (28-77)	0.010
Histology – n (%)					
Clear cell	101 (66%)	58 (78%)	101 (66%)	58 (78%)	
Other	52 (34%)	16 (22%)	52 (34%)	16 (22%)	0.059
cTNM – T – n (%)					
T1-T3a	85 (56%)	57 (77%)	103 (67%)	61 (82%)	
T3b-T4	37 (24%)	13 (18%)	50 (33%)	13 (18%)	0.020
Missing	31 (20%)	4 (5%)			
cTNM – N – n (%)					
N0	56 (37%)	30 (41%)	71 (46%)	37 (50%)	
N1	67 (44%)	30 (41%)	82 (54%)	37 (50%)	0.657
Missing	30 (20%)	14 (19%)			
WHO performance status – n (%)					
0-1	68 (44%)	24 (32%)	126 (82%)	58 (79%)	
2-4	11 (7%)	4 (5%)	27 (18%)	16 (21%)	0.698
Missing	74 (48%)	46 (62%)			
Site of metastasis – n (%)					
One	39 (25%)	34 (46%)	39 (25%)	34 (46%)	
More than one	114 (75%)	40 (54%)	114 (75%)	40 (54%)	0.002
Liver metastasis – n (%)					
No	129 (84%)	68 (92%)	129 (84%)	68 (92%)	
Yes	24 (16%)	6 (8%)	24 (16%)	6 (8%)	0.120
Lung metastasis – n (%)					
No	36 (24%)	16 (22%)	36 (24%)	16 (22%)	
Yes	117 (76%)	58 (78%)	117 (76%)	58 (78%)	0.749
Bone metastasis – n (%)					
No	88 (58%)	59 (80%)	88 (58%)	59 (80%)	
Yes	65 (42%)	15 (20%)	65 (42%)	15 (20%)	0.001
Haemoglobin – n (%)					
Normal	52 (34%)	19 (26%)	53 (34%)	20 (27%)	
< LLN	99 (65%)	52 (70%)	100 (66%)	54 (73%)	0.264
Missing	2 (1%)	3 (4%)			
Neutrophil count – n (%)					
Normal	66 (43%)	24 (32%)	96 (63%)	51 (68%)	
> ULN	40 (26%)	11 (15%)	57 (37%)	23 (32%)	0.568
Missing	47 (31%)	39 (53%)			
Platelet count – n (%)					
Normal	106 (69%)	46 (62%)	109 (71%)	58 (79%)	
> ULN	43 (28%)	13 (18%)	44 (29%)	16 (21%)	0.257
Missing	4 (3%)	15 (20%)			
Albumin – n (%)					
Normal	78 (51%)	31 (42%)	100 (66%)	56 (75%)	
< LLN	41 (27%)	10 (14%)	53 (34%)	18 (25%)	0.217
Missing	34 (22%)	33 (45%)			
Corrected serum calcium – n (%)					
Normal	72 (47%)	31 (42%)	94 (62%)	59 (79%)	
> ULN	46 (30%)	6 (8%)	59 (38%)	15 (21%)	0.037
Missing	35 (23%)	37 (50%)			
Alkaline phosphatase – n (%)					
Normal	101 (66%)	39 (53%)	108 (70%)	58 (78%)	
> ULN	43 (28%)	11 (15%)	45 (30%)	16 (22%)	0.291
Missing	9 (6%)	24 (32%)			
Lactate dehydrogenase – n (%)					
Normal	83 (54%)	30 (41%)	91 (59%)	45 (61%)	
> 1.5 times ULN	57 (37%)	19 (26%)	62 (41%)	29 (39%)	0.781
Missing	13 (8%)	25 (34%)			
Time from diagnosis to start sunitinib – n (%)					
>= One year	0 (0%)	12 (16%)	0 (0%)	12 (16%)	<0.001
< One year	153 (100%)	62 (84%)	153 (100%)	62 (84%)	

Abbreviations: CN, cytoreductive nephrectomy; LLN, lower limit of normal; ULN, upper limit of normal.

Chapter 4

Health-related quality of life and its determinants in patients with metastatic renal cell carcinoma

S. de Groot; W.K. Redekop; M.M. Versteegh; S. Sleijfer; E. Oosterwijk; L.A.L.M. Kiemeney; C.A. Uyl-de Groot

Submitted

Abstract

Objective: Based on improvements of progression-free survival (PFS), new agents for metastatic renal cell carcinoma (mRCC) have been approved. It is assumed that one of the benefits is an improved health-related quality of life (HRQoL), or a delay in HRQoL deterioration as a result of a delay in progression of disease. However, little data are available supporting this relationship. This study aims to provide insight into the most important determinants of HRQoL (including progression of disease) of patients with mRCC.

Methods: A patient registry (PERCEPTION) was created to evaluate treatment of patients with (m)RCC in the Netherlands. HRQoL was measured, using the EORTC QLQ-C30 and EQ-5D-5L, every three months in the first year of participation in the study, and every six months in the second year. Random effects models were used to study associations between HRQoL and patient characteristics.

Results: Eighty-seven patients with mRCC completed 304 questionnaires. The average EORTC QLQ-C30 global health status was 68.88 (SD 18.92) before progression and 61.34 (SD 21.73) after progression of disease. Similarly, the average EQ-5D utility was 0.75 (SD 0.19) before progression and 0.66 (SD 0.30) after progression of disease. Presence of fatigue, nausea and vomiting, pain, dyspnoea, and the application of radiotherapy were associated with significantly lower EQ-5D utilities.

Conclusions: Key drivers for reduced HRQoL in mRCC are disease symptoms. Since symptoms increase with progression of disease, targeted therapies that increase PFS can postpone reductions in HRQoL in mRCC.

Introduction

Renal cell carcinoma (RCC) accounts for 90% of all kidney cancers.[17] While the prognosis of patients with localised disease treated with surgery is relatively good, the prognosis of patients with advanced or metastatic disease (mRCC) is poor. Median overall survival (OS) ranges from 7.8 months for patients with a poor risk to 43.2 months for patients with a favourable risk according to the Heng criteria.[22] Besides the impact of mRCC on survival, mRCC can be associated with severe symptoms, such as cachexia and/or anorexia, asthenia and/or fatigue, pain, anaemia, and venous thromboembolism.[74] These symptoms can impair the health-related quality of life (HRQoL).

Since 2006, several new targeted therapies have been approved for the treatment of mRCC such as sunitinib, sorafenib, pazopanib and everolimus. In phase III studies, these therapies improved progression-free survival (PFS) of patients with mRCC over the diverse comparators,[25-27,29,40,44,46,47] but the effect on OS was less pronounced, likely (partly) due to treatment crossover. It is assumed that one of the benefits of the new therapies is an improved HRQoL, or a delay in HRQoL deterioration as a result of a delay in progression of disease. However, little data are available supporting this relationship, although clinicians feel that a better PFS translates into a better HRQoL.[75] In the context of the high prices of targeted therapies which form a strain on healthcare budgets, it is important to establish whether indeed a delay in progression delays HRQoL deterioration. Besides this clinical perspective, it is important to know if the potential benefit of improved HRQoL has significance from an economic perspective. This economic perspective is generally captured through cost-utility analyses.

In cost-utility analyses benefits are expressed in terms of quality-adjusted life years (QALYs), which incorporate both changes in length of life and changes in HRQoL. Utilities are used to adjust the life-years lived for the quality of life that is lived, and these can take on several values generally ranging from 0 (death) to 1 (full health). They are interpreted as proxies for HRQoL. The EQ-5D is a generic questionnaire, which measures the health states patients are in. It considers various dimensions of quality of life (i.e. mobility, self-care, usual activities, pain/discomfort and anxiety/depression).[76] To derive utilities, a country specific algorithm ('tariff') is applied to the health states generated by the EQ-5D. This method to measure the health states patients are in, and value health (gains) is recommended by several reimbursement authorities across Europe.[77,78]

Relatively little is known about HRQoL of patients with mRCC. This study is the first to provide insight into the most important determinants of HRQoL (including progression of disease) of patients with mRCC using data from a

patient registry in the Netherlands.[79] Additionally, this study aims to assess if those measures used in economic evaluations to assess benefit (i.e. EQ-5D) can detect relevant changes in HRQoL of patients with mRCC.

Patients and methods

Study population

A patient registry (i.e. PERCEPTION) was created to evaluate treatment of patients with (m)RCC in the Netherlands. Patients with RCC (all stages) of any histological subtype diagnosed from 2011 until June 30th 2013 in 25 of 32 hospitals (both general and academic) in three regions in the Netherlands were invited to participate, and fill out HRQoL questionnaires. Eligible patients were identified through the hospitals' registration systems. Additionally, the Netherlands Cancer Registry (NCR), which maintains a cancer registration database of all cancer patients in the Netherlands, was used to ensure that no patients were missed.

The research protocol was approved by the medical ethics committee of Radboud university medical center in Nijmegen (CMO Region Arnhem-Nijmegen) in May 2010.

Data collection

Cancer-specific HRQoL was measured using the EORTC (European Organisation for Research and Treatment of Cancer) QLQ-C30 questionnaire (v3.0).[80] This measure includes five functional scales (physical, role, emotional, social and cognitive), three symptom scales (fatigue, nausea & vomiting and pain), a global health status/QOL scale and six single items (dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties). In addition to the EORTC QLQ-C30, the EQ-5D-5L was used to measure HRQoL. The EQ-5D-5L is a preference-based generic measure, and measures HRQoL on five dimensions, i.e. mobility, self-care, usual activities, pain/discomfort and anxiety/depression.[76] Each dimension includes five severity levels. Patients were sent a HRQoL questionnaire every three months in the first year of participation in the study, and every six months in the second year.

In addition to data on HRQoL, data on demographics, clinical and laboratory factors were collected retrospectively from individual patient records by using uniform case report forms. Clinical and laboratory factors were collected before the start of each new treatment (i.e. surgery, radiotherapy or targeted therapy). Furthermore, data on treatment schemes and treatment endpoints (e.g. survival) were derived from patient records. Data collection stopped at the end of 2013.

Statistical analyses

For each scale of the QLQ-C30, the average of the items that contributed to that scale was calculated. They were then linearly transformed in line with the EORTC QLQ-C30 scoring manual.[81] EQ-5D utilities were derived by combining the answers to the EQ-5D-5L with the Dutch EQ-5D-5L tariff.[82] Mean EQ-5D utilities and HRQoL according to the EORTC QLQ-C30 were calculated by taking the average of the observations for each patient. The proportion of reported problems for each EQ-5D dimension were presented by taking the modus (i.e. the level reported most frequently) across observations for each patient. If two or more modes exist, the highest level was taken.

HRQoL was evaluated separately for the periods before and after progression of disease. Moreover, within the period before progression of disease, a further distinction was made between the period in which patients did not receive any therapy (i.e. wait-and-see) and the period in which they were treated with (first-line) targeted therapy. The treatment period was assumed to last until progression of disease.

Since data on HRQoL were clustered, random effects models were used to study associations between HRQoL (i.e. EORTC QLQ-C30 global health status and EQ-5D utility) and patient and disease characteristics, symptoms and treatment. Use of random effects models ensured that multiple measurements from the same patient were analysed appropriately and made it possible to distinguish between intraindividual and interindividual variation. Backward selection was used to select the covariates for the models; any non-significant covariates were excluded from the models one at a time (significance level of 0.20 for entering and 0.10 for removing the explanatory variables). To control for heteroskedasticity, random effects models with robust standard errors were estimated.

Additionally, random effects logit models were used to study associations between the individual EQ-5D dimensions and patient and disease characteristics, symptoms and treatment. EQ-5D levels were dichotomised into 'no problems/ (i.e. level 1) and 'problems' (i.e. levels 2 to 5).

Missing data regarding patient and disease characteristics were handled using multiple imputations by chained equations. This method generated imputations based on a set of imputation models, one for each variable with missing values.[39]

The significance level was set at $\alpha=0.10$. Data analyses were conducted using STATA statistical analysis software (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.).

Results

Four hundred-eleven patients participating in the study completed 1630 questionnaires. The average EORTC QLQ-C30 global health status of patients with localised disease (365 patients, 1,326 questionnaires) was 76.34 (SD 14.73), and the average EQ-5D utility was 0.82 (SD 0.17).

Eighty-seven patients had mRCC (i.e. metastatic disease at initial presentation or after an initial diagnosis with localised disease). Of these patients, eighty-two percent was male, and the median age at diagnosis was 63 years (Table 4.1). Twenty-six percent of the population did not receive any systemic therapy during follow-up. Of the patients receiving systemic therapy, the majority (80%) was treated with first-line sunitinib. Twenty-three patients also received a second-line therapy within the follow-up period; the majority of these patients was treated with everolimus (13/23). Thirty-one percent of the population received radiotherapy during follow-up.

In total, 304 questionnaires were completed by patients with mRCC and the median number of questionnaires per patient was three (range; 1-7). Table 4.2 shows HRQoL during the different stages of the disease. The EORTC QLQ-C30 global health status was 67.24 (SD 19.19). Patients experienced most problems with role functioning (i.e. doing daily activities and pursuing leisure time activities). Symptoms most commonly reported were fatigue, pain, insomnia and dyspnoea. A significant difference was found in the EORTC QLQ-C30 global health status before and after progression of disease, i.e. 68.88 (SD 18.92) and 61.34 (SD 21.73) (p=0.022). All functioning scales significantly decreased, except for emotional and cognitive functioning. Two symptom scales significantly increased; patients reported more problems regarding dyspnoea (p=0.031) and diarrhoea (p=0.057) after progression than before progression of disease.

Within the period before progression of disease, a similar HRQoL was found for a period without therapy (i.e. wait-and-see) and a period with therapy; mean EORTC QLQ-C30 global health statuses were 69.28 (SD 22.29) and 69.52 (SD 16.63), respectively. However, within the period before progression of disease, patients experienced fewer problems with emotional functioning during a period with therapy compared to a period without therapy (p=0.067). Additionally, patients reported fewer problems regarding constipation (p=0.072), but more problems regarding diarrhoea during a period with therapy compared to a period without therapy (p=0.005).

The average EQ-5D utility was 0.74 (SD 0.19). As with the EORTC QLQ-C30 global health status, a significant difference was found in EQ-5D utility before progression of disease and after progression of disease; the average EQ-5D utility before progression of disease was 0.75 (SD 0.19), whereas the average EQ-5D utility after progression of disease was 0.66 (SD 0.30) (p=0.032). Within

Table 4.1 Baseline characteristics at diagnosis

Variable	Patients (n=87)
Male sex - n (%)	71 (82%)
Age, median [range]	63 [40-79]
Non-clear cell pathology - n (%)	17 (20%)
WHO performance status - n (%)	
0-1	82 (94%)
2-4	5 (6%)
More than one metastatic site - n (%)	48 (55%)
Liver metastasis - n (%)	15 (17%)
Lung metastasis - n (%)	48 (56%)
Bone metastasis - n (%)	21 (24%)
Brain metastasis - n (%)	3 (3%)
Haemoglobin < LLN - n (%)	46 (52%)
Neutrophil count > ULN - n (%)	18 (21%)
Platelet count > ULN - n (%)	19 (22%)
Corrected serum calcium > ULN - n (%)	26 (30%)
Lactate dehydrogenase > 1.5 times ULN - n (%)	11 (12%)
Time since RCC diagnosis < one year	78 (90%)
MSKCC risk score – n (%)	
favourable	6 (7%)
intermediate	54 (62%)
poor	27 (31%)

Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal.

the period before progression of disease, no significant difference was found between a period without therapy (i.e. wait-and-see) and a period with therapy; mean utilities were 0.76 (SD 0.21) and 0.76 (SD 0.18), respectively.

Figures 4.1 and 4.2 show the proportions of patients reporting levels 1 to 5 by EQ-5D dimension, before progression of disease and after progression of disease. Both before and after progression of disease, most problems were reported on the mobility, usual activities and pain/discomfort dimensions.

Univariable analyses show several relationships between disease characteristics, symptoms and treatment, and HRQoL (Table 4.3). Patients with brain metastases and patients with progression of disease reported a significantly lower HRQoL (according to the EORTC QLQ-C30 global health status and EQ-5D). Patients with more than one metastatic site or bone metastases reported a significantly lower EQ-5D utility than other patients, a relationship that was not seen in the QLQ-C30 global health status. Additionally, symptoms (i.e. fatigue, nausea and vomiting, pain, dyspnoea, insomnia, appetite loss, constipation and diarrhoea) were associated with a lower HRQoL (according to the EORTC QLQ-C30 global health status and EQ-5D). Lastly, patients treated with radiotherapy reported a significantly worse HRQoL (according to the EORTC QLQ-C30 global health status and EQ-5D).

Multivariable analysis showed that fatigue, pain and appetite loss were significantly associated with the EORTC QLQ-C30 global health status. Besides these symptoms, the presence of more than one metastatic site, brain metastases and progression of disease were significantly associated with the EORTC

Table 4.2 Health-related quality of life based on the EQ-5D and QLQ-C30

	Total n=87 patients (304 obs.)	Before progression n=81 patients (246 obs.)			After progression n=27 patients (58 obs.)
		Total	No systemic therapy n=47 (125 obs.*)	First-line therapy n=50 (119 obs.)	Total
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*
EQ-5D					
Utility	0.74 (0.19)	0.75 (0.19)	0.76 (0.21)	0.76 (0.18)	0.66 (0.30)
EORTC QLQ-C30					
Global health status	67.24 (19.19)	68.88 (18.92)	69.28 (22.29)	69.52 (16.63)	61.34 (21.73)
Functioning scales					
Physical functioning	69.31 (23.20)	70.89 (22.62)	72.74 (22.48)	69.36 (23.43)	62.08 (29.18)
Role functioning	58.83 (28.21)	61.04 (29.21)	60.84 (29.71)	61.85 (28.66)	52.37 (32.75)
Emotional functioning	79.03 (16.29)	80.31 (17.50)	77.40 (19.05)	81.67 (18.88)	73.35 (18.95)
Cognitive functioning	79.70 (20.12)	80.12 (21.81)	81.27 (21.44)	79.32 (24.67)	76.34 (21.83)
Social functioning	75.57 (22.23)	77.50 (21.96)	77.42 (20.36)	78.12 (22.03)	67.44 (27.95)
Symptom scales					
Fatigue	40.78 (25.26)	39.14 (26.51)	36.44 (26.70)	40.76 (26.74)	47.94 (29.93)
Nausea & vomiting	12.11 (17.35)	12.70 (20.46)	7.58 (13.48)	16.69 (23.84)	10.34 (11.51)
Pain	28.64 (23.71)	27.30 (23.86)	23.92 (24.60)	28.77 (26.09)	34.47 (30.16)
Single items					
Dyspnoea	24.31 (24.27)	23.15 (23.88)	23.26 (24.91)	25.74 (28.35)	29.32 (33.56)
Sleeping	27.59 (26.04)	26.37 (27.46)	24.32 (27.43)	26.51 (29.61)	35.08 (31.17)
Appetite loss	19.18 (25.50)	17.70 (25.07)	15.04 (25.90)	21.00 (25.99)	22.12 (32.44)
Constipation	9.56 (16.71)	8.96 (17.43)	12.35 (23.82)	4.50 (10.14)	12.14 (21.45)
Diarrhoea	19.70 (25.95)	18.88 (27.45)	12.75 (26.67)	23.15 (27.91)	22.43 (26.25)
Financial difficulties	9.79 (18.00)	9.41 (18.14)	9.30 (20.64)	10.80 (19.00)	7.82 (21.09)

Abbreviations: Obs, observations; SD, standard deviation.
* Observations of patients who died within 90 days after being diagnosed with mRCC were excluded from this subgroup (n=2), since these measurements would not contribute to the estimation of the HRQoL of a patient awaiting therapy.

QLQ-C30 global health status. An association was also found between fatigue and pain, and the EQ-5D utility. Furthermore, nausea and vomiting, dyspnoea and treatment with radiotherapy appeared to be significantly associated with HRQoL according to EQ-5D.

Although the univariable analyses showed several relationships between disease characteristics (e.g. the presence of bone or brain metastases and progression of disease) and HRQoL, these characteristics were no longer associated with a deterioration of HRQoL in multivariable analyses after correction for symptoms (at a significance level of 0.05 and 0.01, except for the presence of brain metastases in the model with the EORTC QLQ-C30 global health status as the dependent variable). This seems to imply that symptoms might increase due to progression of disease (and/or due to the spread of the cancer to the bone or brain), which explains the reduced HRQoL.

Figure 4.1 Proportion of patients reporting levels 1 to 5 by dimension, before progression of disease

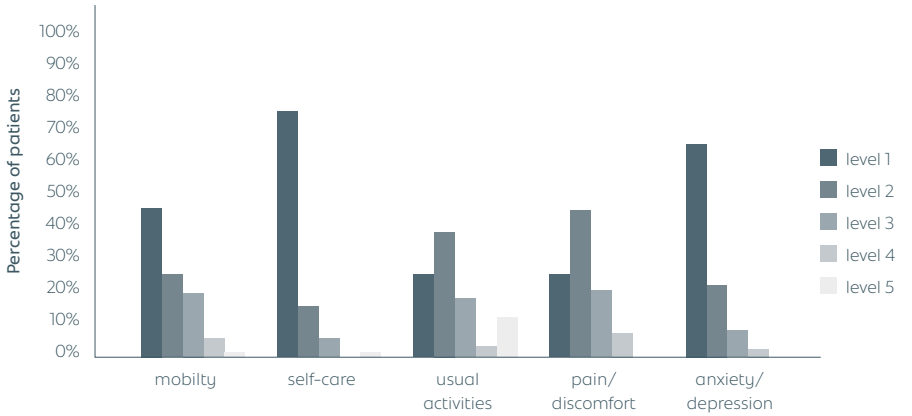


Figure 4.2 Proportion of patients reporting levels 1 to 5 by dimension, after progression of disease

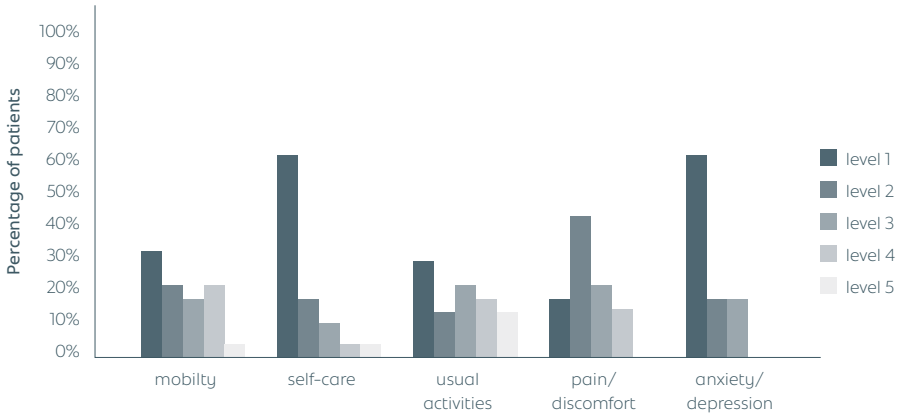


Table 4.4 shows that fatigue was significantly associated with all EQ-5D dimensions, except with the mobility dimension. Pain was also significantly associated with all EQ-5D dimensions, except with the anxiety/depression dimension.

Discussion

This study on health-related quality of life (HRQoL) of patients with metastatic renal cell carcinoma (mRCC) showed that the average EORTC QLQ-C30 global health status was 67.24 (SD 19.19). A significant difference was found between the global health status before and after progression of disease, i.e. 68.88 (SD 18.92) and 61.34 (SD 21.73), respectively. Based on these findings, targeted therapies that increase progression-free survival (PFS) of patients with mRCC, can postpone reductions in HRQoL. The difference between HRQoL before and after progression of disease was also detected by the EQ-5D. The

Table 4.3 Associations between HRQoL and patient and disease characteristics, symptoms and treatment

	EQ-5D utility				EORTC QLQ-C30 global health status			
	Univariable analysis		Multivariable analysis		Univariable analysis		Multivariable analysis	
	Coefficient	*SE*	*Coefficient*	*SE*	*Coefficient*	*SE*	*Coefficient*	*SE*
Patient characteristics								
Male sex	0.077	0.069	NS		2.748	5.198	NS	
Age (per year)	-0.001	0.002	NS		-0.257	0.223	NS	
WHO performance score								
0-1								
2-4	-0.080	0.072	NS		-5.919	7.304	NS	
Disease characteristics								
More than one metastatic site	-0.068*	0.035	NS		-5.048	3.342	4.048*	2.276
Presence of liver metastases	-0.027	0.050	NS		-3.992	4.779	NS	
Presence of lung metastases	-0.021	0.041	NS		0.465	4.074	NS	
Presence of bone metastases	-0.085**	0.040	NS		-3.390	3.915	NS	
Presence of brain metastases	-0.285*	0.170	NS		-21.143*	10.239	-13.586***	2.438
MSKCC risk score								
favourable								
intermediate	0.015	0.062	NS		-0.431	8.924	NS	
poor	0.054	0.063	NS		2.485	9.220	NS	
Progression of disease	-0.082**	0.036	NS		-6.897*	3.000	-3.859*	2.249
Disease duration (in months)	-0.002	0.001	NS		-0.081	0.117	NS	
Symptoms								
Fatigue	-0.004***	0.001	-0.003***	0.001	-0.451***	0.035	-0.316***	0.042
Nausea & vomiting	-0.001*	0.001	0.001**	0.001	-0.360***	0.050	NS	
Pain	-0.004***	0.000	-0.002***	0.000	-0.324***	0.036	-0.143***	0.035
Dyspnoea	-0.003***	0.000	-0.001***	0.000	-0.222***	0.040	NS	
Sleeping	-0.002***	0.000	NS		-0.219***	0.035	NS	
Appetite loss	-0.002***	0.000	NS		-0.274***	0.034	-0.111***	0.035
Constipation	-0.002***	0.001	NS		-0.186***	0.054	NS	
Diarrhoea	-0.001*	0.000	NS		-0.089**	0.040	NS	
Treatment								
Systemic therapy vs. no systemic therapy	0.026	0.027	NS		-0.487	2.408	NS	
Radiotherapy	-0.150***	0.042	-0.115***	0.036	-10.017***	3.306	NS	
Model intercept			0.943***	0.016			85.380***	1.903
R2 (overall)				0.559				0.534
Wald test (p-value)				<0.001				<0.001

Note: Several comorbidities at diagnosis were considered for inclusion in the multivariable analyses, but all appeared to be not significantly associated with HRQoL.
Abbreviations: SE, standard error; NS, not significant.
* Significant at $\alpha = 0.1$ ** Significant at $\alpha = 0.05$ *** Significant at $\alpha = 0.01$.

average EQ-5D utility of patients with mRCC is 0.74 compared to an average of 0.84 (SD 0.18) in the Dutch population aged 60 to 69.[82] Whereas the average EQ-5D utility was 0.75 before progression of disease, the average EQ-5D utility was 0.66 after progression of disease.

The differences between HRQoL before and after progression of disease were found in univariable analyses. Progression of disease was no longer associated with a deterioration of HRQoL in multivariable analyses after correction for symptoms (at a significance level of 0.05 and 0.01). In line with Wilson and Cleary,[83] a relationship between disease characteristics and symp-

Table 4.4 Associations between the EQ-5D dimensions and patient and disease characteristics, symptoms and treatment

	Mobility		Self-care		Usual activities		Pain/ Discomfort		Anxiety/ Depression	
	OR	SE	OR	SE	OR	SE	OR	SE	OR	SE
Patient characteristics										
Male sex	0.149**	0.112	NS		0.095**	0.110	NS		NS	
Age (per year)	1.078**	0.032	NS		NS		NS		NS	
Disease characteristics										
Presence of liver metastases	4.427*	3.395	NS		NS		NS		NS	
Presence of lung metastases	NS		NS		NS		NS		0.300*	0.191
Presence of bone metastases	4.733**	2.961	NS		15.054***	14.768	NS		NS	
MSKCC risk score										
favourable										
intermediate	NS		NS		NS		0.041***	0.049	NS	
poor	NS		NS		NS		0.143	0.176	NS	
Disease duration	NS		1.073**	0.033	NS		NS		NS	
Symptoms										
Fatigue	NS		1.044***	0.012	1.128***	0.028	1.034***	0.012	1.021**	0.010
Nausea & vomiting	0.967**	0.015	NS		NS		NS		NS	
Pain	1.029***	0.009	1.030***	0.010	1.029*	0.015	1.143***	0.023	NS	
Dyspnoea	1.025***	0.009	NS		1.024*	0.014	NS		NS	
Sleeping	NS		NS		NS		NS		1.016*	0.009
Appetite loss	1.031***	0.011	NS		NS		NS		NS	
Treatment										
Radiotherapy	NS		6.062***	3.971	NS		NS		NS	

Note: Odds ratios based on models created using multivariable logistic regression.
Abbreviations: OR, odds ratio; SE, standard error.
* Significant at $\alpha = 0.1$ ** Significant at $\alpha = 0.05$ *** Significant at $\alpha = 0.01$.

toms was expected, which could explain why disease characteristics (such as progression) were no longer statistically significant in the multivariable analyses. Indeed, this study showed that symptoms increase as the disease progresses, and in addition, symptoms appeared to be significantly associated to reduced HRQoL. Similarly, bone metastases were no longer associated with a deterioration of HRQoL in multivariable analyses. However, pain appeared to be significantly associated to reduced HRQoL. Since bone pain is a symptom of cancer that has spread to the bone, this might explain the reduced HRQoL. This seems to imply that symptoms increase due to progression of disease (and/or due to the spread of the cancer to the bone), which explains the reduced HRQoL.

Besides the relationship between symptoms and HRQoL, a significant association was found between radiotherapy and HRQoL (in the model with the EQ-5D utility as the dependent variable). It is possible that this observed association is not due to radiotherapy itself, but instead due to the selection of which mRCC patients are to receive radiotherapy. That is, radiotherapy is mostly reserved for palliation of local and symptomatic disease or to prevent the progression of metastatic disease in critical sites (i.e. bones and brain).[55] Either way, radiotherapy appears to be a significant determinant of HRQoL,

even after correction for patient and disease characteristics (including bone and brain metastases) and symptoms.

74% of the study population was treated with a targeted therapy (the majority received sunitinib). The average EQ-5D utility of these patients was 0.76 before progression of disease. In a study by Cella et al, a similar EQ-5D utility was reported for patients treated with sunitinib (i.e. 0.75).[84] In the economic evaluation of bevacizumab and sunitinib by Thompson-Coon et al,[85] a health state utility of 0.78 (95% CI 0.76-0.80) was used for progression-free survival and 0.70 (95% CI 0.66-0.74) for progressive disease. These utilities were derived from the data presented in the sunitinib submission to NICE and are somewhat higher than the utilities that we found in our study. The economic evaluation of sunitinib by Remák et al[86] was based on the results of a phase II trial of sunitinib as second-line treatment in mRCC;[87] utilities of 0.72 and 0.76 were used for progression-free survival (i.e. during treatment or rest, respectively), whereas utilities of 0.63 and 0.55 were used for progressive disease (i.e. during second-line treatment or after termination of second-line treatment, respectively). The latter utilities are below the utilities found in our study, but this might be explained by differences in the study population, e.g. patients with progression on first-line cytokine therapy were enrolled in the phase II trial.

This study has several limitations. First, a significance level of 0.10 instead of 0.05 was used to study associations between patient characteristics and HRQoL in order to avoid missing any relevant association. Although we might have reported associations that achieved statistical significance due to chance alone, we did not find associations between patient characteristics and HRQoL that we could not interpret. Second, a significant association between WHO performance status and HRQoL, and the MSKCC risk score and HRQoL was not found, although such a relationship would have been expected. The MSKCC risk score divides patients into three risk groups, and gives an indication of the life expectancy of patients with mRCC.[23] Whereas HRQoL was measured several times during the follow-up period, data on patient characteristics (e.g. WHO performance status) and disease characteristics (e.g. laboratory factors, which are part of the MSKCC risk score) were collected only once before the start of a new treatment. As a consequence, too few observations on patient and disease characteristics might have been available to detect a significant association between WHO performance status and the MSKCC risk score, and HRQoL.

A third limitation is that our study sample was too small to find any difference in EQ-5D utilities between different types of targeted therapies, while these therapies differ in toxicity profiles.[88] Nevertheless, although adverse

events have a high impact on HRQoL, an association between adverse events and HRQoL would not be found if the proportion of patients with grade 3 or 4 adverse events is relatively low. Hypertension and fatigue are the most commonly reported grade 3 or 4 adverse events in the randomised phase 3 trial of sunitinib,[25] but these adverse events occurred in only 8% and 7% of the population. Therefore, a very large sample size is needed to find any difference in EQ-5D utilities between different types of targeted therapies. Additionally, it is unknown whether the improved HRQoL due to prolonged PFS outweighs reductions in HRQoL due to treatment-related adverse events. Importantly, this study did not find differences in HRQoL of patients treated with systemic therapy and patients not treated with systemic therapy, or between periods with or without systemic therapy. However, this study may have been underpowered to find such differences.

To conclude, key drivers for reduced HRQoL in mRCC are symptoms of the disease. Since this study showed that symptoms increase with progression of disease, targeted therapies that increase PFS can help to delay loss in HRQoL. This study also showed that the EQ-5D is able to detect changes in HRQoL of patients with mRCC, as it found associations between well-known symptoms of mRCC and EQ-5D utilities. Similar associations were found between these symptoms and the disease-specific EORTC QLQ-C30.

Chapter 5

Potential health gains for patients with metastatic renal cell carcinoma in daily clinical practice: A real-world cost-effectiveness analysis of sequential first- and second-line treatments

S. de Groot; H.M. Blommestein; W.K. Redekop; S. Sleijfer; L.A.L.M. Kiemeney; E. Oosterwijk; C.A. Uyl-de Groot

Submitted

Abstract

Objective: Randomised controlled trials have shown that targeted therapies like sunitinib are effective in metastatic renal cell carcinoma (mRCC). Little is known about the current use of these therapies, and their associated costs and effects in daily clinical practice. We estimated the real-world cost-effectiveness of different treatment strategies comprising one or more sequentially administered drugs.

Methods: Analyses were performed using patient-level data from a Dutch population-based registry including patients diagnosed with primary mRCC from January 2008 to December 2010 (i.e. treated between 2008 and 2013). The full disease course of these patients was estimated using a patient-level simulation model based on regression analyses of the registry data. A healthcare sector perspective was adopted; total costs included healthcare costs related to mRCC. Cost-effectiveness was expressed in cost per life-year and cost per quality-adjusted life year (QALY) gained. Probabilistic sensitivity analysis was conducted to estimate the overall uncertainty surrounding cost-effectiveness.

Results: In current daily practice, 54% (336/621) of all patients was treated with targeted therapies. Most patients (84%; 282/336) received sunitinib as first-line therapy. Of the patients receiving first-line therapy, 30% (101/336) also received second-line therapy; the majority was treated with everolimus (40%, 40/101) or sorafenib (28%, 28/101). Current treatment practice (including patients not receiving targeted therapy) led to 0.807 QALYs; mean costs were €58,912. This resulted in an additional €105,011 per QALY gained compared to not using targeted therapy at all. Forty-six percent of all patients received no targeted therapy; of these patients, 24% (69/285) was eligible for sunitinib. If these patients were treated with first-line sunitinib, mean QALYs would improve by 0.062-0.076 (where the range reflects the choice of second-line therapy). This improvement is completely driven by the health gain seen amongst patients eligible to receive sunitinib but did not receive it, who gain 0.558-0.684 QALYs from sunitinib. Since additional costs would be €7,072-9,913, incremental costs per QALY gained are €93,107-111,972 compared to current practice.

Conclusions: Health can be gained if more treatment-eligible patients receive targeted therapies. Moreover, it will be just as cost-effective to treat these patients with sunitinib as current treatment practice.

Introduction

Attention for the cost-effectiveness of cancer treatments is swiftly increasing, particularly prompted by the advent of so-called molecularly targeted agents. This class of agents has clearly improved outcomes in several tumour types, but also substantially increased costs.[89] One of the tumour types for which targeted treatments are available is metastatic renal cell carcinoma (mRCC).

In 2008, there were an estimated 88,400 new cases of kidney cancer in Europe.[90] The European mean age-standardised 5-year survival was 60.6%, but substantial differences were seen within European regions.[91] Besides registration artefacts, differences in cancer biology, the use of diagnostic tests and screening, and access to high-quality care might explain the differences in cancer survival.[91]

While previous studies demonstrated a survival benefit from targeted therapies in metastatic renal cell carcinoma,[28,29,34,42,43] a Dutch population-based registry showed that many treatment-eligible patients do not receive sunitinib (or any other targeted therapy) in daily practice.[92] This was also seen in England where one in three patients with mRCC eligible for either sunitinib or pazopanib did not receive the drug.[50] Patient and disease characteristics might play a role in the decision to not prescribe targeted therapy. Another possible reason is that it is not cost-effective to treat these patients.

There is little known about the effect that the potential underuse of targeted therapy in daily clinical practice has on health outcomes and costs. The aim of this study was to estimate the real-world cost-effectiveness of several treatment strategies applied in patients with mRCC comprising one or more sequentially administered drugs.

Patients and methods

Study population and data

From the Dutch Cancer Registry, all patients newly diagnosed with mRCC, i.e. metastatic disease at first presentation, from January 2008 until December 2010 in 42 hospitals (both general and academic) in four regions, covering approximately half of the Netherlands, were included in the PERCEPTION registry. In this registry, data on patient characteristics, treatment schemes, treatment endpoints and resource use were retrospectively collected from patient records. Data had been anonymised and de-identified prior to analyses, thus no written informed consent was required. The research protocol was approved by the medical ethics committee of Radboud university medical center in Nijmegen (CMO Region Arnhem-Nijmegen) in May 2010.

Model structure and design

A patient-level simulation (PLS) model was developed to model the full disease course of patients newly diagnosed with mRCC. The model comprised entities (i.e. patients), attributes assigned to the entities, and events. Attributes were obtained from patient-level data from the PERCEPTION registry by selecting clinical factors, biochemical and hematologic factors known to impact mRCC outcomes.[37,38] Events were either second-line treatment or death. The time horizon of the model spanned the patients' lifetime. The total structure of the model is presented in Figure S5.1.

Parameter estimation and time-to-event

For each patient in the PERCEPTION registry, time from diagnosis of mRCC until the first event (TTE1) (i.e. second-line treatment or death) was calculated. Similarly, the time from start of second-line treatment until the second event (TTE2) (i.e. death) was calculated.

We then compared a range of parametric models to extrapolate the survival data. The fit of different models was assessed systematically by performing Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) tests. Additionally, visual inspection was performed by comparing the parametric survival models with the Kaplan-Meier curves.

Clinical factors, biochemical and hematological factors, and the type of targeted treatment were considered for inclusion in the models; TTE1 was also considered as a covariate to estimate TTE2. Backward selection was used to select the attributes for the model; any non-significant attributes (α=0.10) were excluded from the model one at a time. Forward selection was used to create an alternative model. When two different models were created, AIC and BIC tests were performed, and visual inspection was used to decide on the final model.

Missing data were handled using multiple imputations by chained equations.[39] All statistical analyses were conducted using STATA statistical analysis software (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP).

Model calculation

Populations of 621 patients (i.e. the same sample size as the original study population) were repeatedly simulated, one population at a time. Each simulation started with assigning patient and disease characteristics to each patient, based on patient profiles observed in the PERCEPTION registry. That is, random numbers were drawn from predefined distributions for all patient and disease characteristics; similar distributions were used for patients with a similar WHO performance status. For example, the probability of having more than one metastatic site was 64% for patients with a WHO performance status of 0-1, but 73% for patients with a WHO performance status of 2-4. In addition, the previous measurement (e.g. number of metastatic sites before first-line treatment) was taken into account when simulating patient and disease characteristics before second-line treatment.

The TTE1 for each simulated patient was determined by drawing random numbers from two parametric survival models; that is, one model was used to calculate TTE1 until 12 months while a second model was used to calculate TTE1 after 12 months. Two models were used since the probability of an event (i.e. second-line treatment or death) from 12 months onwards was underestimated when only one single model was used. The type of event (i.e. second-line treatment or death) was determined using a separate model. The TTE2 was calculated in a similar manner.

If a patient's modelled time to an event was longer than the remaining life expectancy based on national vital statistics data,[93] we used national vital statistics data to estimate TTE1 and TTE2 because it is not plausible for someone with mRCC to have a longer than average life expectancy.

Treatment scenarios

In the base-case scenario, patients were treated as they were in the real world (Figure 5.1). A multinomial logistic regression model based on patient-level data from the PERCEPTION registry was used to predict the type of treatment in both first- and second-line, using values for WHO performance status, haemoglobin, corrected serum calcium and lactate dehydrogenase (i.e. 4/5 Memorial Sloan Kettering Cancer Center [MSKCC] criteria).[23]

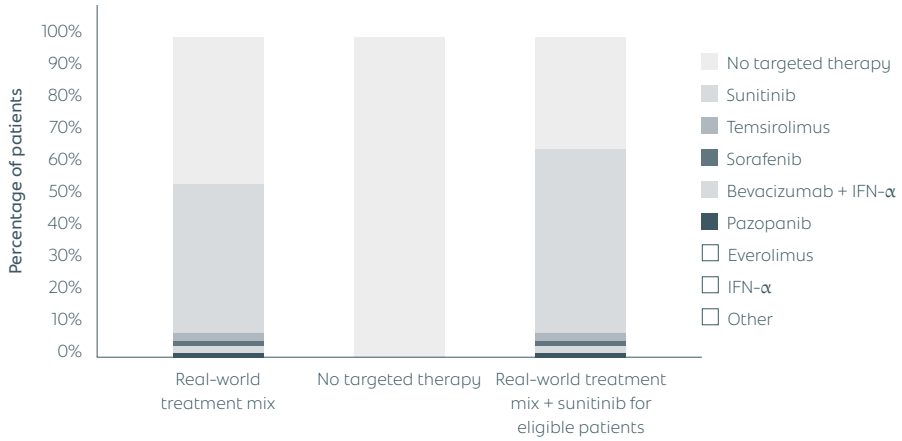
Alternative scenarios included no targeted therapy for all patients, or treating all patients just as they were in reality, except for one difference, namely that first-line sunitinib followed by sorafenib, everolimus or another second-line treatment was given to patients who did not receive any targeted treatment even though they fulfilled the SUTENT trial eligibility criteria (Figure 5.1). A patient was classified as fulfilling SUTENT trial eligibility criteria if he had a clear-cell subtype, a WHO performance status of 0 or 1 and no brain metastases.[25]

The potential health outcomes and costs of all treatment scenarios were calculated by running the model for 621 simulated patients.

Health outcomes

Health outcomes were estimated in terms of life-years (LYs) and quality-adjusted life years (QALYs). QALYs were calculated by weighting LYs for the qual-

Figure 5.1 First-line therapies in the various treatment scenarios



ity of life during these years using utility weights derived from the published literature (Table S5.1).[94-96]

Resource use and costs

This cost-effectiveness analysis (CEA) was conducted from a healthcare sector perspective, but only included healthcare costs related to mRCC, i.e. drug costs as well as resource utilisation costs, such as hospitalisations, outpatient visits and medical imaging services. Hospitalisations due to adverse events were included while other types of costs due to adverse events, such as concomitant medications, were not. The calculation of drug costs and resource utilisation costs is described in the supplementary information.

Costs were reported in Euro 2014. Wherever necessary, costs were adjusted to 2014 using the general price index derived from Statistics Netherlands.[97] Costs and effects were discounted at 4% and 1.5%, respectively.[98]

Model validation

The model was internally validated by comparing patient characteristics and OS observed in the PERCEPTION registry to patient characteristics and health outcomes according to the model.[99] Health outcomes from the model were presented as the mean of 1,000 iterations with a 95% confidence interval (CI) using the standard deviation of 1,000 iterations as standard error of the mean. The model's internal validity was assessed by evaluating whether OS observed in the PERCEPTION registry fell within the 95% CI of the OS according to the base-case scenario of the model.

Sensitivity analysis

Univariate sensitivity analyses were performed to examine the impact of alternative input parameters on the incremental cost-effectiveness ratios (ICERs).

Probabilistic sensitivity analysis (PSA) was conducted to examine the impact of the joint uncertainty regarding all input parameters on the results.

Results

Study population and treatment

714 patients in the Dutch Cancer Registry fulfilled the inclusion criteria. 39 patients were excluded (Figure S5.2), and an additional 30 patients were lost to follow-up. Complete follow-up up to three years after diagnosis was available for 645 patients. Twenty-four of these patients received a metastasectomy (combined with a nephrectomy) with a possible curative intention, making targeted treatment redundant. These patients were therefore excluded from the analyses.

Patient characteristics are shown in Table 5.1, along with the characteristics after multiple imputation. The distribution of patients according to the MSKCC risk score showed a high proportion of patients (55%) with a poor prognosis. 42% of the patients had an intermediate prognosis. Since all patients presented with metastatic disease, very few patients (3%) had a favourable prognosis (e.g. 86% of the patients had a time from diagnosis to treatment, which is one of the MSKCC criteria, of less than one year).

Fifty-four percent (336/621) of all patients was treated with targeted therapies. Of these patients, 84% (282/336) received sunitinib as first-line therapy. Other first-line treatments given were temsirolimus (7%, 24/336) and sorafenib (3%, 11/336). 101 patients also received a second-line therapy; the majority was treated with everolimus (40%, 40/101) or sorafenib (28%, 28/101). Median overall survival (OS) of patients treated with targeted therapies was 12.6 months (95% CI 10.5-14.8).

Almost half (46%, 285/621) of all patients did not receive any targeted therapy. Of these 285 patients, 69 patients (24%) fulfilled the SUTENT trial eligibility criteria. Most patients (76%) did not fulfill the SUTENT trial eligibility criteria; 168 patients (78%) did not have a clear-cell subtype, 46 patients (21%) did not have a WHO performance status of 0 or 1 and 2 patients (1%) had brain metastases. Median OS of patients not treated with targeted therapies was 2.6 months (95% CI 2.1-3.5); 10.6 months (95% CI 3.8-18.6) for patients fulfilling the SUTENT trial eligibility criteria and 1.9 months (95% CI 1.6-2.6) for patients not fulfilling the SUTENT trial eligibility criteria.

Table 5.2 shows the final models with their covariates (e.g. patient and disease characteristics) and corresponding coefficients to estimate TTE1, the type of event after TTE1 and TTE2. For example, a WHO performance status of 2-4 means a shorter TTE1.

Table 5.1 Patient and disease characteristics before start of first- and second-line treatment

	First-line		Second-line	
	Real-world data (n=621)	Average of imputed datasets (n=621)	Real-world data (n=101)	Average of imputed datasets (n=101)
Sex – n (%)				
Female	213 (34%)	213 (34%)	27 (27%)	27 (27%)
Male	408 (66%)	408 (66%)	74 (73%)	74 (73%)
Median age – yr (range)	66 (23-93)	66 (23-93)	62 (23-79)	62 (23-79)
Histology – n (%)				
Clear cell	354 (57%)	354 (57%)	69 (68%)	69 (68%)
Other *	267 (43%)	267 (43%)	32 (32%)	32 (32%)
WHO performance status – n (%)				
0-1	204 (33%)	430 (69%)	34 (34%)	73 (72%)
2-4	61 (10%)	191 (31%)	9 (9%)	28 (28%)
Missing	356 (57%)		58 (57%)	
Site of metastasis – n (%)				
One	195 (31%)	206 (33%)	19 (19%)	19 (19%)
More than one	398 (64%)	415 (67%)	82 (81%)	82 (81%)
Missing	28 (5%)		0 (0%)	
Liver metastasis – n (%)				
No	487 (78%)	509 (82%)	74 (73%)	74 (73%)
Yes	106 (17%)	112 (18%)	27 (27%)	27 (27%)
Missing	28 (5%)		0 (0%)	
Lung metastasis – n (%)				
No	163 (26%)	173 (28%)	21 (21%)	21 (21%)
Yes	430 (69%)	448 (72%)	80 (79%)	80 (79%)
Missing	28 (5%)		0 (0%)	
Bone metastasis – n (%)				
No	375 (60%)	393 (63%)	58 (57%)	58 (57%)
Yes	218 (35%)	228 (37%)	43 (43%)	43 (43%)
Missing	28 (5%)		0 (0%)	
Brain metastasis – n (%)				
No	546 (88%)	571 (92%)	92 (91%)	92 (91%)
Yes	47 (8%)	50 (8%)	9 (9%)	9 (9%)
Missing	28 (5%)		0 (0%)	
Prior nephrectomy – n (%)				
No	452 (73%)	453 (73%)	43 (43%)	43 (43%)
Yes	168 (27%)	168 (27%)	58 (57%)	58 (57%)
Missing	1 (0%)		0 (0.0%)	
Haemoglobin – n (%)				
Normal	171 (28%)	205 (33%)	20 (20%)	20 (20%)
< LLN	347 (56%)	416 (67%)	78 (77%)	81 (80%)
Missing	103 (17%)		3 (3%)	
Neutrophil count – n (%)				
Normal	203 (33%)	383 (62%)	67 (66%)	88 (87%)
> ULN	108 (17%)	238 (38%)	10 (10%)	13 (13%)
Missing	310 (50%)		24 (24%)	
Platelet count – n (%)				
Normal	358 (58%)	452 (73%)	66 (65%)	70 (69%)
> ULN	140 (23%)	169 (27%)	29 (29%)	31 (31%)
Missing	123 (20%)		6 (6%)	
Albumin – n (%)				
Normal	247 (40%)	391 (63%)	51 (51%)	75 (74%)
< LLN	136 (22%)	230 (37%)	18 (18%)	26 (26%)
Missing	238 (38%)		32 (32%)	
Corrected serum calcium – n (%)				
Normal	243 (39%)	421 (68%)	45 (45%)	72 (71%)
> ULN	116 (19%)	200 (32%)	18 (18%)	29 (29%)
Missing	262 (42%)		38 (38%)	
Alkaline phosphatase – n (%)				
Normal	324 (52%)	432 (70%)	65 (64%)	74 (73%)
> ULN	139 (22%)	189 (30%)	24 (24%)	27 (27%)
Missing	158 (25%)		13 (13%)	
Lactate dehydrogenase – n (%)				
Normal	277 (45%)	372 (60%)	63 (62%)	71 (70%)
> 1.5 times ULN	174 (28%)	249 (40%)	28 (28%)	30 (30%)
Missing	170 (27%)		10 (10%)	

Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal.
* mRCC was clinically established without histopathological confirmation in 17% of patients and mRCC was classified as not otherwise specified without further subtyping in 13% of patients. It is likely that a substantial proportion of these patients had a clear cell subtype.

Internal validation

Observed data from the PERCEPTION registry showed a median OS of 7.3 months (0.6 LYs) (95% CI 6.3-8.4) for the total population (in which 54% of the patients received a targeted therapy). Median OS in the model was 7.0 months (0.6 LYs) (95% CI 5.7-8.3) if patients were treated as they were in the real world (i.e. base-case scenario). The OS derived from the PERCEPTION registry fell within the 95% CI of the outcome of the model. Additionally, the observed Kaplan Meier curves (TTE1, TTE2 and OS) were closely followed by the survival curves derived from the model (Figures S5.3-S5.5).

Effectiveness

The model yielded an estimated mean survival of 1.2 LYs (14.4 months) if patients were treated as they were in the real world (in which 54% of the patients received a targeted therapy). If all treatment-eligible patients would be treated with first-line sunitinib followed by sorafenib, everolimus or another second-line treatment (if they did not die after first-line treatment), mean survival would increase to 1.3 LYs (15.6 months). If none of the patients were to be treated with any targeted therapy, mean survival would decrease to 0.9 LYs (10.8 months) (Table 5.3).

If patients were treated as they were in the real world, mean QALYs are 0.807. If all treatment-eligible patients would be treated with first-line sunitinib followed by sorafenib, everolimus or another second-line treatment, mean QALYs would increase to 0.883, 0.868 or 0.841 respectively. If none of the patients were to be treated with any targeted therapy, mean QALYs would decrease to 0.576 (Table 5.3).

Costs

Mean total costs per treatment strategy are presented in Table 5.3. Mean total costs per patient amount to €58,912 if patients were treated as they were in the real world. If all treatment-eligible patients were to be treated with first-line sunitinib followed by sorafenib, everolimus or another second-line treatment, mean total costs would increase to €65,984, €65,825 or €65,062 respectively. If none of the patients would be treated with any targeted therapy, mean total costs would decrease to €34,733.

Cost-effectiveness

Compared to a scenario in which none of the patients receives a targeted therapy, the real-world treatment mix results in a QALY gain of 0.230 and a cost increase of €24,179. Thus, an additional €105,011 per QALY gained is spent compared to the scenario of not using targeted therapy.

Table 5.2 Covariates and corresponding coefficients of the survival models and logistic regression model

	Time to event 1 (first 12 months)	Time to event 1 (> 12 months)	Type of event 1*	Time to event 2
Model type	Loglogistic	Exponential	Logistic	Weibull
Covariate	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Constant	1.060 (0.393)	5.227 (0.663)	-1.778 (0.790)	2.804 (0.249)
Age (yr)	0.012 (0.005)	-0.027 (0.009)	0.037 (0.013)	
Sex (male vs. female)				-0.382 (0.185)
Histology (non-clear cell vs. clear cell)	-0.229 (0.112)			
Prior nephrectomy (yes vs. no)	0.783 (0.130)			
Number of metastatic sites (more than 1 vs. 1)	-0.306 (0.120)			-0.387 (0.199)
WHO performance status (2-4 vs. 0-1)	-0.585 (0.136)			
Liver metastases (yes vs. no)	-0.459 (0.137)			0.632 (0.164)
Bone metastases (yes vs. no)	0.277 (0.113)	-0.330 (0.170)		-0.421 (0.150)
Brain metastases (yes vs. no)				-0.657 (0.238)
Thrombocytes (>ULN vs. normal)			0.587 (0.320)	-0.360 (0.183)
Neutrophil count (>ULN vs. normal)	-0.258 (0.136)			
Albumin (<LLN vs. normal)	-0.290 (0.137)		1.074 (0.381)	-0.388 (0.196)
Alkaline phosphatase (>ULN vs. normal)	-0.269 (0.129)			
First-line sunitinib (vs. no targeted therapy)	0.915 (0.124)	-0.602 (0.201)		
First-line temsirolimus (vs. no targeted therapy)	0.580 (0.244)	-1.529 (0.611)		
First-line other (vs. no targeted therapy)	0.992 (0.273)	0.339 (0.375)		
Second-line everolimus (vs. sorafenib)				-0.143 (0.191)
Second-line other (vs. sorafenib)				-0.388 (0.181)
TTE 1 (TTE 1 > 12 months vs. TTE 1 <= 12 months)			-0.641 (0.270)	0.537 (0.147)
Shape parameter	0.669 (0.028)			1.625 (0.137)

Note: Lung metastases, haemaglobin, corrected serum calcium and lactate dehydrogenase were considered for inclusion in the survival models and logistic regression model, but excluded through backward and/or forward selection.
Abbreviations: LLN, lower limit of normal; ULN, upper limit of normal; SE, standard error; TTE 1, time to event 1.
*0 = second-line therapy/ 1 = death.

Compared to the real-world treatment mix, first-line sunitinib followed by sorafenib, everolimus or another second-line treatment leads to a QALY gain of 0.076, 0.062 and 0.035 at the population level, respectively. When combined with the corresponding incremental costs (i.e. €7,072, €6,913 and €6,150), the incremental costs per QALY gained are €93,107, €111,972 and 177,226, respectively (Table 5.3). Note that the health gains are achieved by changing the treatment of a relatively small group of patients representing 11% of the population. These are patients who were eligible to receive sunitinib but did not receive it in real life; in this group, sunitinib leads to a gain of 0.684, 0.558 or 0.315 QALYs per patient.

Sensitivity analyses

The tornado diagram (Figure 5.2) shows the variability in the ICER of sunitinib followed by everolimus compared to the real-world treatment mix as a consequence of changes in the values of various input parameters. Varying the unit costs of first-line sunitinib has the highest impact on the ICER.

Figure 5.3 shows the uncertainty around the total costs and QALYs as obtained from the PSA. For sunitinib followed by sorafenib, everolimus or another second-line treatment, 88.3%, 82.3% and 72.1% of all simulations fell in the north-east quadrant indicating more QALYs and higher costs compared to the real-world treatment mix. For the scenario in which none of the patients received a targeted therapy, 99.4% of all simulations fell in the south-west quadrant indicating less QALYs and lower costs.

Cost-effectiveness acceptability curves are presented in Figure 5.4, showing the likelihood that treatment strategies would be cost-effective at a given willingness-to-pay threshold. Treating according to the real-world treatment mix never attains more than 16% of simulations. Treating all patients with sunitinib followed by sorafenib or everolimus would be favoured; these scenarios have a probability of 34% and 15%, respectively, of being cost-effective at a willingness-to-pay threshold of €106,000. Not treating any patient with a targeted therapy would be preferred at willingness-to-pay thresholds below €106,000, but this scenario results in health loss.

Discussion

To the best of our knowledge, this is the first study that models the full disease course of patients with mRCC using real-world data. We found that real-world treatment of mRCC patients yields a QALY gain of 0.230 with incremental costs of €24,179 compared to a scenario in which none of the patients would receive a targeted therapy. Thus, we currently pay €105,011 per QALY gained. However, only 54% of the patients in our study population received a targeted therapy and this raises the question about what the potential impact would be if all treatment-eligible patients were to receive targeted therapy. Compared to real-world treatment, health can be gained if all eligible patients were to be treated with first-line sunitinib followed by sorafenib or everolimus. The costs to gain health by treating all eligible patients with these treatment strategies, i.e. €93,107 and €111,972 per QALY gained, respectively, are similar to the current costs per QALY gained. These costs include the costs of both first- and second-line therapy.

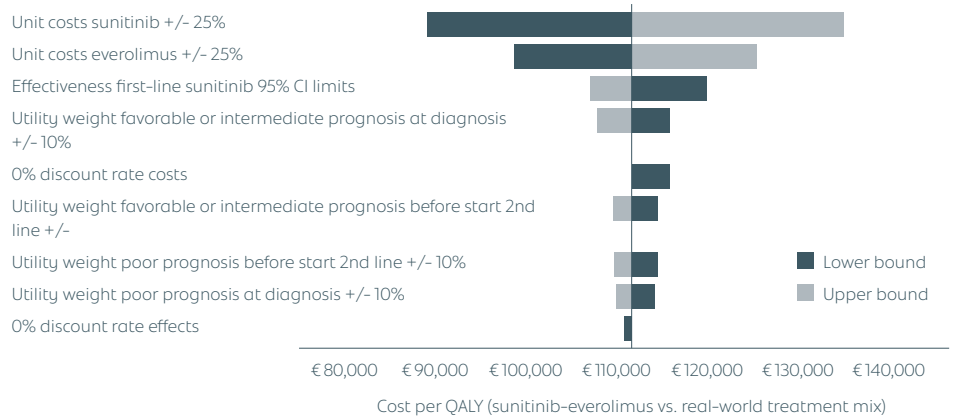
The proportion of patients not being treated in this series is high at 46% (285/621). However, not all these patients were eligible for targeted therapy. We found that one in four patients (69/268; 26%) fulfilling SUTENT trial eligibility criteria did not receive any targeted therapy. Also in England, one in three patients with mRCC eligible for either sunitinib or pazopanib did not receive the drug.[50] Previous analyses indicated that patients aged 65+ years were less likely to receive targeted therapy than younger patients after adjustment for other factors.[92] However, the exact causes underlying the remarkably high proportion of non-treated patients deserve further study. Importantly, all drugs studied in this project were available in the Netherlands during the study period without any limitations for patients or prescribers, so this could

Table 5.3 Summary of the cost-effectiveness results

	Real-world treatment mix	No targeted therapy	Real-world treatment mix + sunitinib for eligible patients (followed by sorafenib)	Real-world treatment mix + sunitinib for eligible patients (followed by everolimus)	Real-world treatment mix + sunitinib for eligible patients (followed by other)
Per strategy					
Time to event 1 (years) - mean (95% CI)	1.1 (1.0-1.3)	0.9 (0.6-1.1)	1.1 (1.0-1.3)	1.1 (1.0-1.3)	1.1 (1.0-1.3)
Time to event 2 (years) - mean (95% CI)*	0.7 (0.5-0.9)	NA	0.9 (0.6-1.2)	0.8 (0.5-1.0)	0.6 (0.4-0.8)
LYs - mean (95% CI)	1.2 (1.0-1.4)	0.9 (0.6-1.1)	1.3 (1.1-1.5)	1.3 (1.1-1.5)	1.3 (1.1-1.5)
QALYs - mean (95% CI)	0.807 (0.647-0.966)	0.576 (0.403-0.749)	0.883 (0.719-1.046)	0.868 (0.709-1.027)	0.841 (0.687-0.996)
Total costs - mean (95% CI)	€ 58,912 (€ 48,393-€ 69,431)	€ 34,733 (€ 23,164-€ 46,301)	€ 65,984 (€ 55,009-€ 76,959)	€ 65,825 (€ 54,661-€ 76,989)	€ 65,062 (€ 54,106-€ 76,018)
Compared to the real-world treatment mix					
LYs gained - mean (95% CI)	NA	-0.4 (-0.6- -0.1)	0.1 (-0.1-0.3)	0.1 (-0.1-0.3)	0.1 (-0.1-0.2)
QALYs gained - mean (95% CI)	NA	-0.230 (-0.390 - -0.070)	0.076 (-0.056-0.208)	0.062 (-0.068-0.191)	0.035 (-0.094-0.163)
Incremental costs - mean (95% CI)	NA	-€ 24,179 (-€ 34,856 - -€ 13,502)	€ 7,072 (-€ 2,070-€ 16,214)	€ 6,913 (-€ 2,252-€ 16,079)	€ 6,150 (-€ 2,867-€ 15,167)
Cost/ LYG	NA	NR	€ 60,716	€ 73,485	€ 117,814
Cost/ QALY	NA	NR	€ 93,107	€ 111,972	€ 177,226
Compared to no targeted therapy					
LYs gained - mean (95% CI)	0.4 (0.1-0.6)	NA	0.5 (0.2-0.7)	0.4 (0.2-0.7)	0.4 (0.1-0.7)
QALYs gained - mean (95% CI)	0.230 (0.070-0.390)	NA	0.306 (0.126-0.486)	0.292 (0.114-0.470)	0.265 (0.088-0.442)
Incremental costs - mean (95% CI)	€ 24,179 (€ 13,502-€ 34,856)	NA	€ 31,251 (€ 18,848-€ 43,654)	€ 31,093 (€ 18,391-€ 43,794)	€ 30,329 (€ 17,734-€ 42,924)
Cost/ LYG	€ 69,068	NA	€ 66,983	€ 70,003	€ 75,393
Cost/ QALY	€ 105,011	NA	€ 102,058	€ 106,483	€ 114,469

Note: Results are discounted (benefits 1.5% and costs 4%).
Abbreviations: LYs, life years; QALYs, quality-adjusted life years; LYG, life years gained; CI, confidence interval; NA, not applicable; NR, not reported.
* Time to event 2 is only relevant for patients who received a second-line therapy.

Figure 5.2 Results of the univariate sensitivity analyses



not explain why many eligible patients in Dutch daily clinical practice did not receive targeted therapy.

Multiple economic evaluations of targeted therapies in mRCC have been published using data from RCTs, [86,100-103] two of which examined the cost-effectiveness of sunitinib. Several explanations exist for differences regarding the cost-effectiveness of sunitinib between our study and these two studies (by Remák et al[86] and Benedict et al[102]). To start with, we used real-world data (using the PERCEPTION registry) while the other studies used data from key clinical trials. In addition, we looked at all patients with mRCC (at the initial presentation), while the other two studies just studied one subgroup (i.e. sunitinib-eligible patients).

Some limitations to the data and methods deserve mentioning. First, since data from all patients newly diagnosed with mRCC at the initial presentation in 42 hospitals (both general and academic) were collected, it is likely that these patients are representative of average patients with mRCC at initial diagnosis and the average treatment in the Netherlands. However, these patients account for only 40%-70% of the total mRCC population.[104] This total population also includes patients who initially presented with non-metastatic disease and later developed distant metastases. It is likely that a lot more of these patients are being treated; a CEA based on these patients will likely yield different results.

Second, the model was populated using data from primary mRCC patients diagnosed between January 2008 and December 2010 (i.e. treated between 2008 and 2013), whereas new treatments have become available since then. The PERCEPTION registry also included a cohort of (m)RCC patients diagnosed between January 2011 and June 2013, but since inclusion criteria differed, patients were identified differently and the duration of follow-up varied, it was not feasible to include these patients in the current study. Nevertheless, we

Figure 5.3 Cost-effectiveness plane for various treatment scenarios versus real-world treatment mix

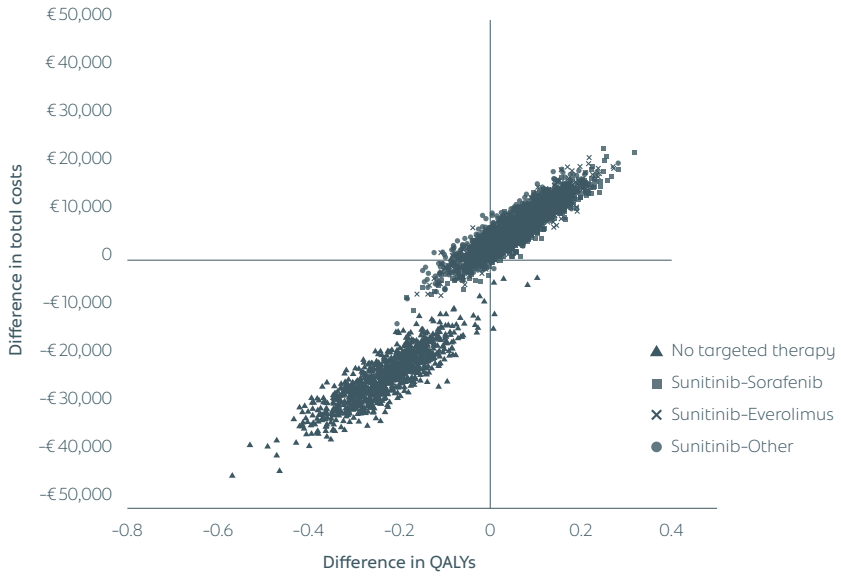
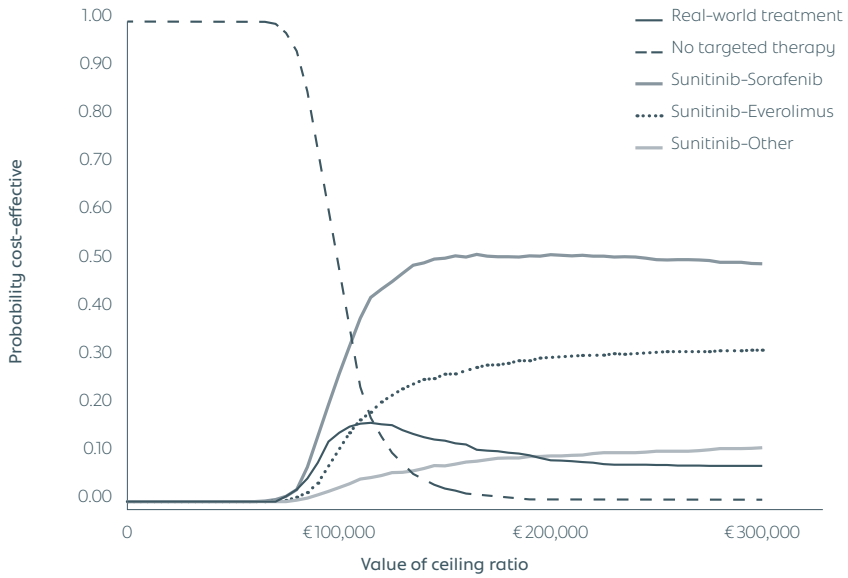


Figure 5.4 Cost-effectiveness acceptability curves representing the probability that each treatment strategy is cost-effective for a given maximum willingness-to-pay threshold per QALY gained



Note: The y-axis shows the likelihood that strategies would be considered cost-effective for a given cost-effectiveness willingness to pay threshold.

did not observe an increase in the proportion of patients receiving targeted therapies in this population.[92] Therefore, we believe that the conclusion of this study still applies, and health can be gained if more treatment-eligible patients receive targeted therapies.

Third, the alternative scenarios included in this model assume that all treatment-eligible patients can be treated with a certain targeted therapy. This assumption may overestimate the number of eligible patients since some of these patients eligible on the basis of the data captured in the PERCEPTION registry, may not actually be eligible because of poor organ function or comorbidities.

Fourth, treatment costs were overestimated somewhat since we did not adjust for dose reductions. However, the effect on the incremental costs and ICERs will be minimal since the treatment costs of 54% of the patients in the base-case real-world scenario were also overestimated somewhat. Another limitation of this study is the amount of missing data in baseline characteristics. Multiple imputations by chained equations were conducted to overcome this problem.[39] This method ensures that all patients are included in the analysis but simultaneously guarantees that the uncertainties from missing data are retained.

In conclusion, RCTs have shown that targeted therapies like sunitinib are effective in mRCC treatment. RCT-based cost-effectiveness analyses with a lifetime time horizon provide important information about the cost-effectiveness of these therapies. However, these analyses are limited in scope, since they are conducted in a selected population. A full disease model and real-world data as presented here are essential in estimating cost-effectiveness ratios that are externally valid. We found that one in four patients eligible for sunitinib did not receive it. It is difficult to state with certainty why these patients did not receive sunitinib. One possible reason is a limited health gain from treatment with sunitinib, but this reasoning seems unlikely since we estimated that its use may add 0.684 QALYs (or eight months in perfect health) to individual patients. Another possible reason is that it is not cost-effective to treat these patients. However, we found that it is just as cost-effective to treat these patients with sunitinib as current treatment practice.

Supplementary information: Calculation of drug costs and resource utilisation costs

Drug costs were calculated by multiplying monthly costs of a therapy[105] by the time to an event (Table S5.2). The PERCEPTION registry showed that, if patients did not undergo a cytoreductive nephrectomy prior to drug treatment, they started drug treatment 1.4 months (SD 1.7, n=219) after diagnosis. If patients underwent cytoreductive nephrectomy prior to drug treatment, they started drug treatment 8.6 months (SD 9.8, n=116) after diagnosis. Therefore, in the model we assumed patients started drug treatment one month after diagnosis, except for patients who underwent a nephrectomy prior to drug treatment; we assumed that these patients started drug treatment nine months after diagnosis.

We also assumed patients discontinued drug treatment either one month before start of second-line treatment or three months before death. These assumptions are also based on the PERCEPTION registry, but since the date of discontinuation was often lacking in the PERCEPTION registry, it was not possible to calculate the time between discontinuation and either start of second-line treatment or death for all patients in the registry. This assumption was therefore verified by a clinical expert who participated in the registry.

A maximum first-line treatment duration of 41 months, and a maximum second-line treatment duration of 34.1 months was assumed based on the results from other studies.[34,48]

Monthly resource use per treatment strategy was derived from patient-level data (Table S5.3). Total costs were calculated by multiplying monthly resource use by unit costs,[106,107] and the time to an event.

Table S5.1 Utility weights

Prognosis	Utility (mean)	Source
Favourable or intermediate prognosis at diagnosis	0.725*	94
Poor prognosis at diagnosis	0.590	95
Favourable or intermediate prognosis before start second-line therapy	0.700**	96
Poor prognosis before start second-line therapy	0.590	95

* Average of utility at baseline for patients receiving pazopanib and patients receiving placebo
** Average of utility for patients treated with axitinib and patients treated with sorafenib.

Table S5.2 Drug costs

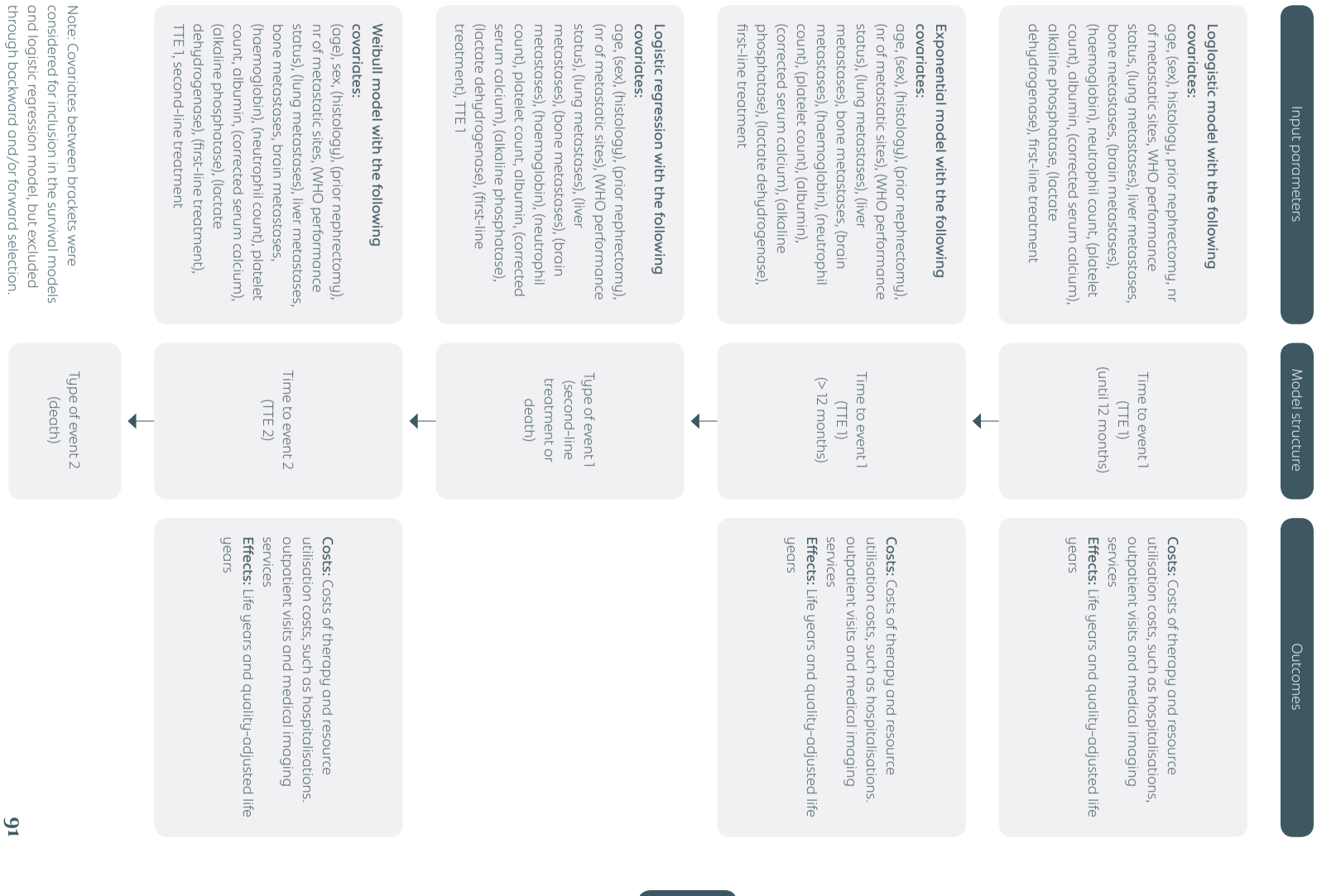
Prognosis	Dose and frequency	Costs (€)	Costs per month (€)	Source
Sunitinib	50 mg daily for 4 weeks, followed by 2-week rest period	184 per 50 mg	3,727	105
Temsirolimus	25 mg once per week	928 per dose	4,030	105
Sorafenib	400 mg twice daily	35 per 200 mg	4,243	105
Everolimus	10 mg daily	129 per 10 mg	3,931	105

Table S5.3 Unit costs and resource use per treatment strategy per month

Resources	Unit costs (€)	No targeted therapy n=215 Mean (SE)	End of life (<1 month) n=72 Mean (SE)	First-line sunitinib n=281 Mean (SE)	First-line temsirolimus n=23 Mean (SE)	Second-line sorafenib n=28 Mean (SE)	Second-line everolimus n=40 Mean (SE)
Inpatient days – non – ic	530	5.1 (0.5)	18.4 (1.3)	2.7 (0.2)	4.4 (1.0)	1.5 (0.6)	2.6 (0.5)
Inpatient days – ic	2,401	0.1 (0.0)	1.2 (0.5)	0.0 (0.0)	0.1 (0.1)	0.0 (0.0)	0.0 (0.0)
Outpatient visits	94	1.4 (0.1)	1.9 (0.3)	1.9 (0.1)	2.4 (0.3)	1.4 (0.2)	2.2 (0.3)
Day care treatments	276	0.1 (0.0)	0.1 (0.0)	0.2 (0.0)	2.0 (0.2)	0.3 (0.1)	0.3 (0.1)
Emergency room visits	166	0.2 (0.0)	0.8 (0.1)	0.1 (0.0)	0.2 (0.1)	0.1 (0.0)	0.2 (0.0)
Laboratory	5	6.2 (1.0)	29.0 (5.7)	4.2 (0.4)	5.9 (1.0)	2.4 (0.6)	4.2 (0.7)
X-ray	53	0.8 (0.1)	4.1 (0.7)	0.6 (0.1)	1.4 (0.3)	0.5 (0.2)	0.8 (0.2)
CT-scan	189	0.5 (0.0)	1.6 (0.3)	0.5 (0.0)	0.7 (0.1)	0.4 (0.1)	0.5 (0.1)
MRI	281	0.1 (0.0)	0.2 (0.1)	0.1 (0.0)	0.1 (0.0)	0.0 (0.0)	0.0 (0.0)
Ultrasound	88	0.2 (0.0)	0.9 (0.2)	0.2 (0.0)	0.2 (0.0)	0.1 (0.1)	0.1 (0.0)
PET-CT	1,163	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Skeletal scintigraphy	244	0.1 (0.0)	0.1 (0.1)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Heart scintigraphy (MUGA)	314	0.0 (0.0)	0.1 (0.1)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Electrocardiogram	16	0.0 (0.0)	0.2 (0.1)	0.1 (0.0)	0.1 (0.0)	0.0 (0.0)	0.1 (0.0)

Note: Unit costs of inpatient days, day treatments and outpatient visits were based on detailed microcosting studies.[106] Resource use related to imaging services was valued using the tariffs as issued by the Dutch Healthcare Authority.[107]
Abbreviations: SE, standard error; ic, intensive care.

Figure S5.1 Model structure



Note: Covariates between brackets were considered for inclusion in the survival models and logistic regression model, but excluded through backward and/or forward selection.

Figure S5.2 Patient enrollment

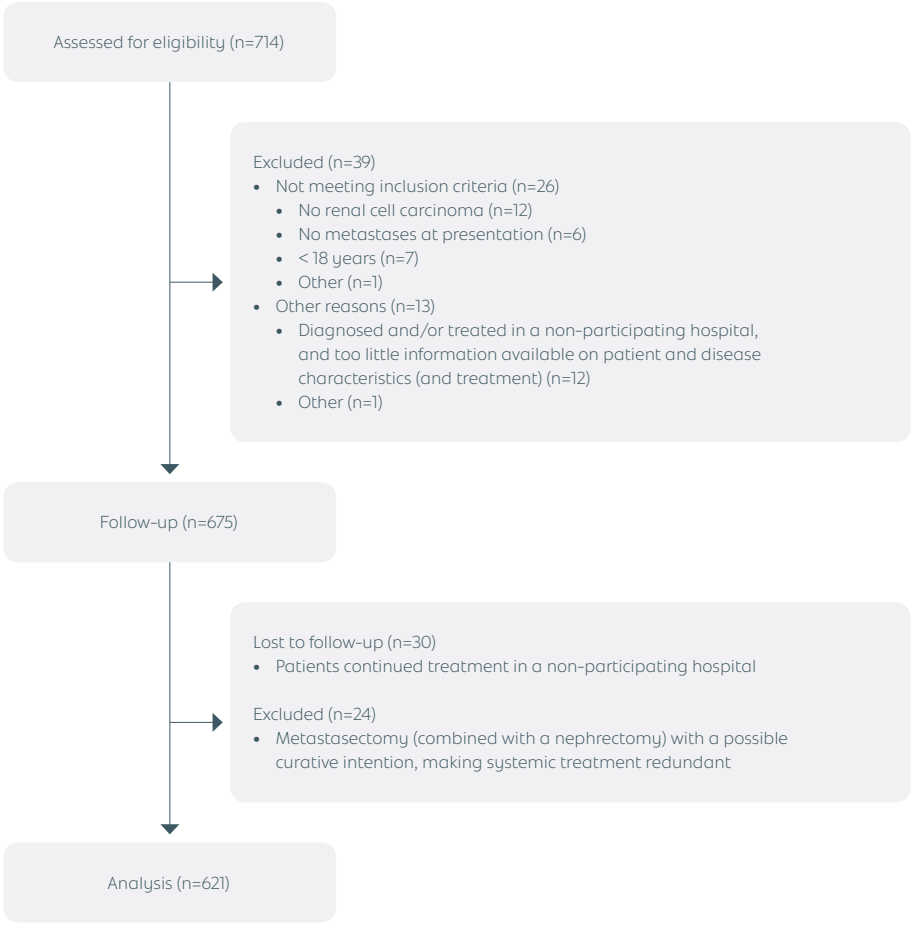


Figure S5.3 Comparison of time to event 1 (i.e. time from diagnosis to second-line treatment or death) between original and simulated data

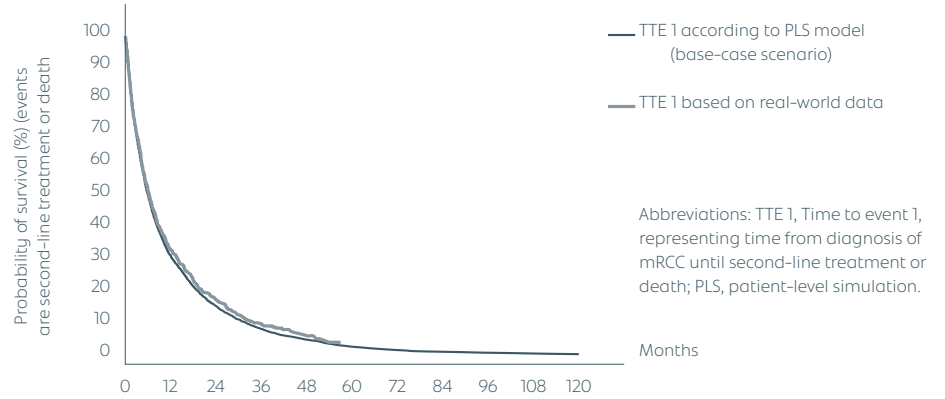


Figure S5.4 Comparison of time to event 2 (i.e. time from start of second-line treatment to death) between original and simulated data

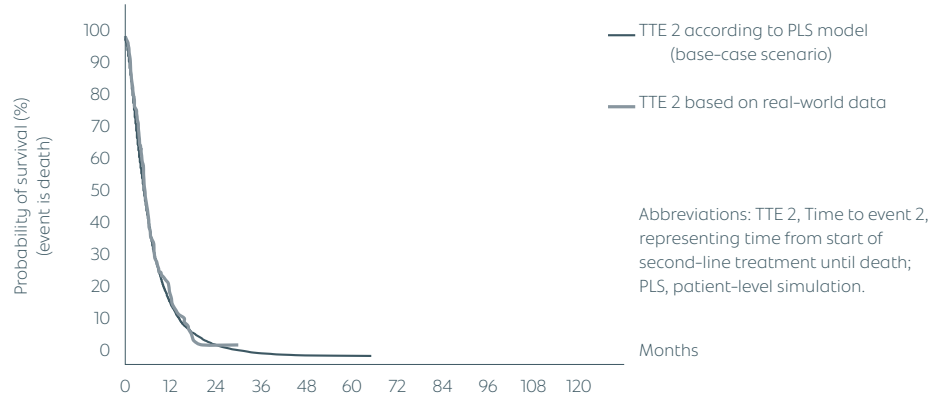
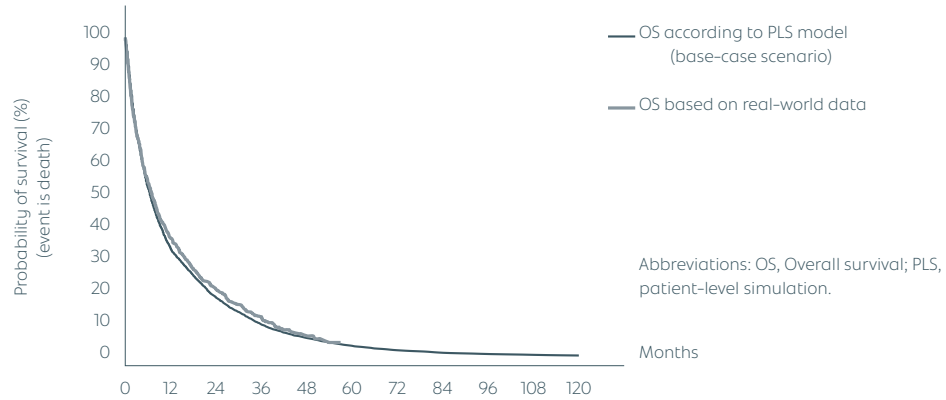


Figure S5.5 Comparison of overall survival (i.e. time from diagnosis to death) between original and simulated data



Chapter 6

Balancing the optimal and the feasible: A practical guide for setting up patient registries for the collection of real-world data for healthcare decision making based on Dutch experiences

S. de Groot*; N. van der Linden*; M.G. Franken; H.M. Blommestein; B. Leeneman; E. van Rooijen; J.J.M. van der Hoeven; M.W. Wouters; H.M. Westgeest; C.A. Uyl- de Groot

Value in health 2016 [In Press]

* Both authors contributed equally to this manuscript

Abstract

Objective: The aim of this article was to provide practical guidance in setting up patient registries to facilitate real-world data collection for healthcare decision making.

Methods: This guidance was based on our experiences and involvement in setting up patient registries in oncology in the Netherlands. All aspects were structured according to 1) mission and goals ('the Why'), 2) stakeholders and funding ('the Who'), 3) type and content ('the What'), and 4) identification and recruitment of patients, data handling, and pharmacovigilance ('the How').

Results: The mission of most patient registries is improving patient health by improving the quality of patient care; monitoring and evaluating patient care is often the primary goal ('the Why'). It is important to align the objectives of the registry and agree on a clear and functional governance structure with all stakeholders ('the Who'). There is often a tradeoff between reliability, validity, and specificity of data elements and feasibility of data collection ('the What'). Patient privacy should be carefully protected, and address (inter-) national and local regulations. Patient registries can reveal unique safety information, but it can be challenging to comply with pharmacovigilance guidelines ('the How').

Conclusions: It is crucial to set up an efficient patient registry that serves its aims by collecting the right data of the right patient in the right way. It can be expected that patient registries will become the new standard alongside randomised controlled trials due to their unique value.

Introduction

Globally, there is an increasing trend to use real-world data to inform decision making in healthcare. Real-world data are often collected using a patient registry. A patient registry can be defined as "an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes".[108]

Regulatory authorities (United States Food and Drug Administration and European Medicines Agency [EMA]) can require real-world data collection for safety surveillance and risk assessment (e.g. Risk Evaluation and Mitigation Strategy by the Food and Drug Administration and risk management plan by the EMA).[109] Furthermore, reimbursement agencies increasingly use real-world data in decision making. This was, for example, seen in the Netherlands where a coverage with evidence development policy was implemented in 2006.[9] This policy aims to guarantee early access to expensive drugs that have an added therapeutic value and an expected budget impact of at least 2.5 million Euros.[13] In exchange, it is required to collect data regarding appropriate drug use, effectiveness, and cost-effectiveness in real-world clinical practice. These data are intended to complement the findings from clinical trial(s), and to evaluate a drug's real-world value after 4 years of initial reimbursement. As a consequence of the introduction of this policy, the number of patient registries has been rapidly increasing in the Netherlands.

In this article, we provide practical guidance in setting up patient registries for the collection of real-world data. Although guidance for designing patient registries exists,[108] we specifically address practical issues. This article is based on our involvement in setting up patient registries in the Netherlands for various types of cancer (i.e. melanoma, lung, prostate, renal cell, haematological, colorectal, and head and neck cancer). We first discuss the mission and goals ('the Why') of patient registries and highlight issues related to stakeholders and funding ('the Who'). After that, challenges and solutions will be discussed regarding the type and content of a patient registry ('the What') and the identification and recruitment of patients, data handling, and pharmacovigilance ('the How'). Last, we discuss the main challenges in balancing the optimal and the feasible in setting up patient registries.

Mission and goals ('the Why')

Why use a patient registry and how to guarantee valorisation of outcomes?

The mission of most registries is improving patient health by improving the quality of patient care; monitoring and evaluating patient care is therefore

often the primary goal. This goal may be operationalised in several ways. For example, patient registries are one of EMA's tools to gain insight into risks of a product in real-world clinical practice.[109] Patient registries can also provide information on appropriate use (i.e. is a product used in the right way in the right patients), effectiveness, costs, and cost-effectiveness in real-world clinical practice.[110] Furthermore, registries can include essential information on patient-reported outcome measures (PROMs) in case data are prospectively collected. Moreover, patient registries can inform public health planning (e.g. registering causes of disease to illustrate the need for a prevention program).[111] It is important to be very specific about how the primary goal of monitoring and evaluating patient care will be operationalised and/or interpreted. Ultimately, this will ease the other steps in setting up patient registries.

Monitoring and evaluating patient care may not immediately improve patient health but may improve the health of future patients. It is essential to frequently discuss findings with clinicians and ensure a quality-of-care feedback loop. Furthermore, outcomes can be used in the development of clinical guidelines. Table 6.1 provides an overview of the mission and goals of the registries in which we are involved. All registries ensure transparency to the public through presentations and publications.[112-119] However, only the melanoma registry (Dutch Melanoma Treatment Registry [DMTR]) fortnightly provides clinicians with online benchmarked feedback regarding a predefined set of quality indicators developed by the professional organisation. These quality indicators will be shared at a hospital level with healthcare insurers, patient organisations, and the general public in the near future. Quality-of-care improvement by using a structured feedback loop to clinicians was not part of the initial aims of most of the registries. This may be explained by the fact that most of the registries in which we are involved were funded by manufacturers and mainly set up for reimbursement purposes. Besides reimbursement purposes, the melanoma registry (DMTR) was set up for monitoring quality of care, which was obligated by the professional organisation.

Important lessons to feedback loops are that agreement needs to be reached on the type of indicators that will be collected, how they will be measured, and the way they will be presented. In addition, the data need to be representative for all patients within a certain hospital (e.g. starting data collection on patients with a worse prognosis will initially lead to biased feedback) and the data need to be case-mix corrected to allow valid comparisons between hospitals (or clinicians), especially when it concerns outcomes indicators. To correct for differences between patients at baseline, the registry should contain a sufficient number of observations and sufficient data on the

relevant prognostic factors. Last, a user-friendly (Web-based) application is needed to facilitate a quality-of-care feedback loop.

Stakeholders and funding ('the Who')

Who are involved in the registry?

Broad support for the registry is needed to maximize its benefits. Identifying and engaging relevant stakeholders is key to the success of a patient registry. Stakeholders include clinicians, patients, researchers, governmental parties, healthcare insurers, and manufacturers. Involvement from professional organisations and clinical experts (including key opinion leaders) improves the valorisation of results. Involvement of patient representatives secures patient participation and may help ensure that the aims of the registry are pursued with minimal burden to patients. Participation of manufacturers may support funding of the registry. Table 6.2 illustrates the involvement of stakeholders in the registries in which we are involved.

Stakeholders can, however, have conflicting interests. An essential and potentially time-consuming step is aligning the aims of the registry with these interests. It is important to determine the main objectives with key stakeholders at an early stage. It is also crucial to establish a clear and functional governance structure including a description of tasks, responsibilities, and decision-making processes. In the prostate cancer registry (CAstration-resistant Prostate cancer RegIstry), clinical data and health-related quality of life data are collected in two separate projects with separate funding and study protocols; however, both projects are carried out by the same project team. The project team is the core executive body, responsible for the day-to-day management of the registry, coordination, and adherence to the planning and protocol. The project team is advised by a clinical steering committee as well as a general assembly. The clinical steering committee has decision-making power regarding the clinical and scientific aspects of the registry (e.g. data collection and publication of results) and includes balanced representatives of urologists, medical oncologists, and radiotherapists of the participating hospitals and the Dutch uro-oncology study group. The general assembly represents all relevant stakeholders (including all involved manufacturers and representatives of the Dutch prostate cancer patient organisation). Scientific proposals are judged by the steering committee, and the writing team includes the involved project team members and a selection of the steering committee and the subinvestigators from the participating hospitals.

Another issue may be related to data ownership (including publishing rights), (level of) data access, and data sharing. For example, when multiple manufacturers fund the registry, they may not be willing to share prod-

Table 6.1 Mission and goals ('the Why')

Name of registry The Why	PHAROS 1	CAPRI and PRO-CAPRI	DMTR	Melanoma	Metastatic colorectal carcinoma	PERCEPTION	Non-small cell lung carcinoma	POSEIDON (not running)	Locally advanced Head & Neck	Recurrent and/or metastatic Head & Neck
Disease	CLL, MM, NHL	CRPC	Melanoma (unresectable stage IIIc/IV)	Melanoma (stage I–IV)	mCRC	mRCC	NSCLC	NSCLC	(LA) SCCHN	(RM) SCCHN
Aim: • Providing insights into patient and disease characteristics and treatment patterns	X	X	X	X	X	X	X	X	X	X
• Providing insights into clinical outcomes and economic outcomes	X	X	X	X	X	X	X	X	X	X
• Providing insights into patient reported outcomes										
• Related to health-related quality of life	X*	X	X			X		X		
• Related to costs (direct and/or indirect)		X	X					X		
• Providing online benchmarked feedback to clinicians, hospitals and manufacturers			X							
• Identifying prognostic groups based on patient material			Future aim			X		To be decided		

Abbreviations: PHAROS, Population-based HAematological Registry for Observational Studies; CAPRI, CAstration-resistant Prostate cancer RegIstry; PRO-CAPRI, Patient Reported Outcomes in the CAstration-resistant Prostate cancer RegIstry; DMTR, Dutch Melanoma Treatment Registry; PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry; POSEIDON, Prospective Observational Study Examining Investments and Derived Outcomes in NSCLC treatment; CLL, chronic lymphocytic leukemia; MM, multiple myeloma; NHL, Non Hodgkin lymphoma; CRPC, castration-resistant prostate cancer; mCRC, metastatic colorectal carcinoma; mRCC, metastatic renal cell carcinoma; NSCLC, Non-small-cell lung carcinoma; LA SCCHN, locally advanced Squamous Cell Carcinoma of the Head and Neck; RM SCCHN, recurrent and/or metastatic Squamous Cell Carcinoma of the Head and Neck.
* Data on health-related quality of life was collected in The Profiles registry[120]

uct-specific data. In this case, detailed product-specific data can be shared with the product owner, whereas aggregated data can be shared with other companies. By allowing variation in the level of data sharing,[121] competing parties can participate and benefit from collaboration within the same registry.

Who funds the registry?

It is crucial to secure sufficient funding for all activities related to the registry to ensure viability and sustainability. Activities include designing the registry (e.g. stakeholder meetings, writing and revising the study protocol, defining data sets, and ethical approval) and running the registry (e.g. data collection, data analyses, writing, and reporting). Ensuring funding can be challenging, especially in case of extensive data collection and/or long-term follow-up. Long-term funding arrangements are essential for the sustainability of a registry.

Registries can be funded from one or multiple sources including public and private sources. Potential funding sources are manufacturers, healthcare insurers, governmental parties, patient organisations, professional associations, private foundations, and advocacy groups. Funding for the registries in which we are involved was often provided by multiple manufacturers. These registries were largely motivated by the need to collect real-world data on the performance of drugs in line with the Dutch coverage with evidence development policy. Some of these registries also received governmental funding (including [unrestricted] research grants).

Multisponsor registries have the advantage of decreasing the financial burden for each party and securing wider support. However, sponsors may have conflicting interests and different ideas about the design and planning of the registry. For example, multiple manufacturers were involved in the haematological registry (Population-based HAematological Registry for Observational Studies 1). They had products for various indications in different treatment lines. Because the optimal approach to collect data may differ per party (e.g. dependent on treatment line), priorities needed to be set and needed to be acceptable for all parties.

Another example is the lung cancer registry (Prospective Observational Study Examining Investments and Derived Outcomes in NSCLC), aimed to start in four hospitals. Although the set-up started 3 years ago, it is currently unknown whether data collection will actually commence. Over time, more stakeholders became involved and the objectives became concurrently broader. For example, one of the objectives was to collect detailed biomarker information for scientific purposes and in order to conduct economic evaluations

Table 6.2 Stakeholders and funding ('the Who')

Name of registry		PHAROS 1	CAPRI and PRO-CAPRI	DMTR	Melanoma	Metastatic colorectal carcinoma	PERCEPTION	Non-small cell lung carcinoma	POSEIDON (not running)	Locally advanced Head & Neck	Recurrent and/or metastatic Head & Neck
The Who											
Disease		CLL, MM, NHL	CRPC	Melanoma (unresectable stage IIIc/IV)	Melanoma (stage I–IV)	mCRC	mRCC	NSCLC	NSCLC	(LA) SCCHN	(RM) SCCHN
Consultation*	• Clinicians and/or hospitals	X	X	X	X	X	X	X	X	X	X
	• Governmental party	X	X	X			X				X
	• Manufacturer(s)	X	X	X	X	X	X	X	X	X	X
	• Patients		X	X					X		
	• Researchers/academia	X	X	X	X	X	X	X	X	X	X
Decision making/ governance**	• Clinicians and/or hospitals	X	X	X			X		X		
	• Governmental party	X	X	X			X				X
	• Manufacturer(s)	X		X	X		X				
	• Patients		X	X							
	• Researchers/academia	X	X	X	X	X	X	X	X	X	X
Funding	• Clinicians and/or hospitals	X									
	• Governmental party	X	X	X			X				X
	• Manufacturer(s)	X	X	X	X	X	X	X	X	X	X

Abbreviations: PHAROS, Population-based HAematological Registry for Observational Studies; CAPRI, CAstration-resistant Prostate cancer RegIstry; PRO-CAPRI, Patient Reported Outcomes in the CAstration-resistant Prostate cancer RegIstry; DMTR, Dutch Melanoma Treatment Registry; PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry; POSEIDON, Prospective Observational Study Examining Investments and Derived Outcomes in NSCLC treatment; CLL, chronic lymphocytic leukemia; MM, multiple myeloma; NHL, Non Hodgkin lymphoma; CRPC, castration-resistant prostate cancer; mCRC, metastatic colorectal carcinoma; mRCC, metastatic renal cell carcinoma; NSCLC, Non-small-cell lung carcinoma; LA SCCHN, locally advanced Squamous Cell Carcinoma of the Head and Neck; RM SCCHN, recurrent and/or metastatic Squamous Cell Carcinoma of the Head and Neck.
* Stakeholders involved with the registry initiative and/or design. ** Stakeholders who have a formal say in decisions regarding the project when it is running.

of targeted therapies. However, collecting data on biomarkers increases the requirements for infrastructure and funding. Furthermore, different stakeholders had different ideas about the type of biomarker data to be included. Agreement between all stakeholders has not yet been reached.

A practical solution for future registries is to carefully consider the number and type of stakeholders and their specific role in decision making. The inclusion of more stakeholders increases potential benefits, but it can also complicate decision making.

Type and content ('the What')

What is a suitable type and content?

A patient registry can be intervention-based or disease-based.[108] An intervention-based registry addresses research questions regarding appropriate use, effectiveness, cost-effectiveness, and safety. Disease-based registries provide additional information and facilitate studying the full disease course including (sequential) treatment pathways.[116] Furthermore, such a registry provides information on the number of untreated patients and whether these patients would have been eligible for treatment. It should be noted, however, that this also adds to complexity, time, and costs of a registry. Table 6.3 provides an overview of the type and content of the registries in which we are involved.

Both intervention-based and disease-based registries can include all patients who meet the inclusion criteria or include a sample of this population. Including all patients adds to time and costs, whereas selecting a sample can be more efficient but can have pitfalls as well. In particular, the representativeness of the patient population may be hampered (external validity). Although causal studies about how nature works do not necessarily need a representative sample, representativeness is crucial in studies describing a specific population at a specific point in time.[122] As a consequence, a representative sample is needed when monitoring and evaluating patient care. A random sample or a cluster sample can enhance representativeness. A cluster sample includes patients in a certain cluster (e.g. a region or a hospital) based on the assumption that the cluster is representative of other clusters.

To increase efficiency, it may be an option to use multiple-phase sampling. For example, in a two-phase design, limited data are first collected in a large sample, after which detailed data are collected in a subsample. The melanoma registry (DMTR) uses such an approach. Minimal data are collected on patients who are not treated in a melanoma centre (due to a worse prognosis), whereas full data (clinical, economic, PROMs) are collected for all patients who received treatment in 1 of the 14 melanoma centres. In addition, more

Table 6.3 Type and content of the registry ('the What')

Name of registry		PHAROS 1	CAPRI and PRO-CAPRI	DMTR	Melanoma	Metastatic colorectal carcinoma	PERCEPTION	Non-small cell lung carcinoma	POSEIDON (not running)	Locally advanced Head & Neck	Recurrent and/or metastatic Head & Neck
The What											
Disease		CLL, MM, NHL	CRPC	Melanoma (unresectable stage IIIc/IV)	Melanoma (stage I–IV)	mCRC	mRCC	NSCLC	NSCLC	(LA) SCCHN	(RM) SCCHN
Type:	• Disease-based	X	X	X	X		X	X	X		
	• Intervention-based					X				X	X
Scope:	• Population-based			X							
	• Sample-based	X	X		X	X	X	X	X	X	X
Content:	• Patient and disease characteristics/ treatment	X	X	X	X	X	X	X	X	X	X
	• Clinical outcomes	X	X	X	X	X	X	X	X	X	X
	• Economic outcomes	X	X	X	X	X	X	X	X	X	X
	• Patient reported outcomes	X*	X	X			X		X		
	• Quality of care indicators**			X							
	• Patient material			Future aim			X		To be decided		
Data-collection:	• Prospective		X	X			X		X		
	• Retrospective	X	X	X	X	X	X	X		X	X
	• Start and end date	From 2010	2012–2017	From 2013	2012–2015	2010–2013	2011–2014	2012–2014	To be decided	2011	2011–2013
	• Years of diagnosis	From 2004	2010–2015	From 2012	2003–2011	2003–2013	2008–2013	2009–2011	To be decided	2007–2010	2006–2013

Abbreviations: PHAROS, Population-based HAematological Registry for Observational Studies; CAPRI, CAstration-resistant Prostate cancer RegIstry; PRO-CAPRI, Patient Reported Outcomes in the CAstration-resistant Prostate cancer RegIstry; DMTR, Dutch Melanoma Treatment Registry; PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry; POSEIDON, Prospective Observational Study Examining Investments and Derived Outcomes in NSCLC treatment; CLL, chronic lymphocytic leukemia; MM, multiple myeloma; NHL, Non Hodgkin lymphoma; CRPC, castration-resistant prostate cancer; mCRC, metastatic colorectal carcinoma; mRCC, metastatic renal cell carcinoma; NSCLC, Non-small-cell lung carcinoma; LA SCCHN, locally advanced Squamous Cell Carcinoma of the Head and Neck; RM SCCHN, recurrent and/or metastatic Squamous Cell Carcinoma of the Head and Neck.
* Data on health-related quality of life was collected in The Profiles registry[120] ** Quality of care indicators can be derived from all registries (e.g. length of a stay in a hospital). However, the DMTR is the only registry providing online benchmarked feedback to clinicians, hospitals and manufacturers.

detailed data (additional healthcare resource use, productivity losses, and informal care) are only collected in a selection of 4 of the 14 centres.

Despite the sampling procedures, which initially enhance representativeness, representativeness is hampered in case patients who do not want to participate differ from those who participate, or in case patients are not randomly lost to follow-up. In addition, sampling from a complete sampling frame is not always feasible, especially for registries using a prospective design.

What data elements?

What data elements to include largely depends on the goal of the registry. If the goal is to improve the quality of patient care by providing information on appropriate use, effectiveness, and cost-effectiveness in real-world clinical practice, comprehensive data are needed on patient and disease characteristics, treatment, and outcomes (health and economic outcomes). However, if the goal is explicitly focused on effectiveness and safety in order to improve the quality of patient care, the choice of data elements can be more selective. To select the most important data elements, an analysis plan can be created. Describing the future data analyses helps identifying those data elements that are essential and those elements that are academically 'interesting'.[123]

Data elements should, preferably, be based on data standards (e.g. Clinical Data Interchange Standards Consortium), current data sets (e.g. national disease registry), and/or standard terminology (e.g. Systematized Nomenclature of Medicine). This facilitates comparison to other studies and creates the opportunity to link different data sets.

Consultation of experts ensures the selection of appropriate data elements.[124] It is important to involve clinical experts as well as experts in using real-world data. Clinical experts who are not experienced with real-world data may advise on data elements that are difficult to collect in a real-world setting. It is always recommended to test the availability of data elements. In case there is a lack of reliable data about a certain variable, it may be possible to use a proxy (e.g. time to next treatment as a proxy for time to progression).

Using real-world data always implies balancing between reliability, validity, and specificity of data elements on the one hand and the feasibility of data collection (affordability and completeness) on the other hand. The available sources will set boundaries to what can be collected and influence the manner of data collection. For example, data on adverse events in clinical trials are commonly reported using the Common Terminology Criteria for Adverse Events as graded by the clinician. This is, however, often not feasible in a registry, unless the Common Terminology Criteria for Adverse Events are consistently used and concisely reported in medical charts in clinical practice.

Table 6.4 Identification and recruitment of patients, handling data and pharmacovigilance ('the How')

Name of registry		PHAROS 1	CAPRI and PRO-CAPRI	DMTR	Melanoma	Metastatic colorectal carcinoma	PERCEPTION	Non-small cell lung carcinoma	POSEIDON (not running)	Locally advanced Head & Neck	Recurrent and/or metastatic Head & Neck
The How											
Disease		CLL, MM, NHL	CRPC	Melanoma (unresectable stage IIIc/IV)	Melanoma (stage I–IV)	mCRC	mRCC	NSCLC	NSCLC	(LA) SCCHN	(RM) SCCHN
Identification of patients:	• National database(s)	X			X		X		To be decided		
	• Hospital database(s)		X			X	X	X	To be decided		X
	• Clinicians (database(s))			X					To be decided	X	X
Handling data:	• Paper-based case report form	X			X					X	
	• Electronic case report form	X	X	X	X		X	X	X		X
Handling PROMS:	• Paper-based questionnaire	X	X	X*			X		X		
	• Electronic questionnaire	X		X					X		
Patient privacy protection:	• Anonymisation and/or pseudonomisation	X	X	X	X		X	X		X	X
	• Trusted third party			X					X		
(S)AE:	• Collection	X	X	X	X	X	X	X	X	X	X
	• Reporting to pharmacovigilance authority	X	Yes, (S)AE level						To be decided		

Abbreviations: PHAROS, Population-based HAematological Registry for Observational Studies; CAPRI, CAstration-resistant Prostate cancer RegIstry; PRO-CAPRI, Patient Reported Outcomes in the CAstration-resistant Prostate cancer RegIstry; DMTR, Dutch Melanoma Treatment Registry; PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry; POSEIDON, Prospective Observational Study Examining Investments and Derived Outcomes in NSCLC treatment; CLL, chronic lymphocytic leukemia; MM, multiple myeloma; NHL, Non Hodgkin lymphoma; CRPC, castration-resistant prostate cancer; mCRC, metastatic colorectal carcinoma; mRCC, metastatic renal cell carcinoma; NSCLC, Non-small-cell lung carcinoma; LA SCCHN, locally advanced Squamous Cell Carcinoma of the Head and Neck; RM SCCHN, recurrent and/or metastatic Squamous Cell Carcinoma of the Head and Neck; (S)AE, (serious) adverse event.
* At the convenience of the patient.

In the lung cancer study, data were retrospectively collected from medical charts. Only 8.5% of adverse events (81 of 956) were graded by a clinician using a standardised grading system and reported in the medical chart. Only 51% were sufficiently reported to retrospectively derive a grade, as judged by data managers. Therefore, a tension may exist between optimizing reliability (register and grade an adverse event only if recorded by the treating clinician) and optimizing other properties of the registry such as data completeness. When selecting the data elements, one has to be aware of such trade-offs so as to optimize the attributes most important to the registry.

Identification and recruitment of patients, data handling, and pharmacovigilance ('the How')

How to identify patients?

Any type of registry may have issues regarding the identification of eligible patients. In population-based patient registries, it is essential to identify and include all eligible patients (e.g. with the diagnosis of interest or treated with the intervention of interest). In contrast, a sample of the population can be drawn, and existing databases can be used to identify eligible patients. It is crucial to ensure representativeness when using an existing database (e.g. national databases, hospital databases, and clinicians [databases]). Drawing a sample from patients joining a patient association may, for example, lead to selection bias (e.g. a higher educated group of patients). The potential for bias can be evaluated by examining different studies addressing similar research questions and comparing patient and disease characteristics to the characteristics of the patients in the registry. Table 6.4 illustrates how patients were identified in the registries in which we are involved.

In the retrospective part of the renal cancer registry (Pharmaco-Economics in Renal CEll carcinoma: a PopulaTION-based registry [PERCEPTION]), eligible patients were identified through the Netherlands Cancer Registry, which includes basic information on 95% of all cancer patients. A cluster sample was selected for inclusion in this registry (i.e. all patients with metastatic renal cell carcinoma in 42 from 51 hospitals in four regions, covering approximately half the country). A practical hurdle arises when (sufficient) information is not available on the population. For the prospective part of this registry, the Netherlands Cancer Registry could not provide a timely and complete list of eligible patients. Therefore, lists of patients diagnosed with metastatic renal cell carcinoma were fortnightly derived from hospitals' financing systems, in addition to the Netherlands Cancer Registry.

How to recruit patients?

The recruitment of patients can be a serious challenge. Participation can be voluntary or compulsory for patients and/or clinicians. To increase participation rates, it could be made compulsory to gain access to and/or reimbursement of a product (e.g. an expensive drug). This was partly the case in the melanoma registry (DMTR). The Dutch minister made the financing of an expensive melanoma drug conditional on the set-up of a population-based registry and centralisation of melanoma care in 14 specialist centres (endorsed by health insurers).

However, participation in most registries is voluntary. Patients can have multiple incentives to participate. Because a registry most likely does not change current treatment, improving future patients' health may be the most important incentive. Clinicians or hospitals may be incentivised by a particular research interest or the ability to achieve other goals (e.g. reimbursement, transparency, and improvement in quality of care).[108] Furthermore, a (financial) compensation for time invested by either clinicians or patients may help to increase participation.

How to handle the data?

Paper or electronic case report forms (CRFs) can be used to record information. Electronic CRFs offer the advantage of automatic validation checks and do not require transferring data from paper to an electronic database. The database needs to be suitable for the registry, including the level of detail of the data.

Furthermore, electronic and paper-based patient questionnaires can be used to collect PROMs. In the PERCEPTION registry, patients were sent a health-related quality-of-life questionnaire every 3 months in the first year of participation in the study and every 6 months in the second year. Experiences from the PERCEPTION registry showed that most patients who gave informed consent returned the questionnaire on a short notice; response rates varied between 80% and 90%. However, response rates can vary substantially between studies, and may depend on the study population and the burden of the questionnaire(s). To increase participation and response, it may be an option to use both electronic and paper-based patient questionnaires especially in case most patients are elderly. In addition, in case this matches the required measuring moments, questionnaires can be completed at clinic visits, for example, in the waiting room (e.g. by using a tablet). Furthermore, especially in case of immobile or terminally ill patients, telephone calls or house visits by study staff may be needed to collect the required patient-reported data.

The process of data collection should be designed to maximize participation and response, data quality, and efficiency while minimizing patient burden.

To improve the quality of clinical data, clinicians can be requested to register or verify data. This is, however, often not feasible because clinicians often lack time to review large volumes of patient data. In case registry data are used for the evaluation of the quality of care in multiple hospitals, external data managers may increase objectivity and may ensure uniformity of data collection. In the melanoma registry (DMTR), all data recorded by data managers need to be validated by clinicians. This validation process is, however, time-consuming. Validation efforts should therefore preferably focus at the most important variables (such as toxicities) that may not reliably be captured by data managers. Uniformity of data collection in the DMTR was improved by initially recording data on 10% of all patients by two data managers (one external).

It is essential to adequately and continuously train data managers supported by a detailed and up-to-date manual. This also includes guidance on when to record a value as missing, unknown, or as negative. For example, there is a difference between a patient who had no test for locating metastases and a patient who had a test but no metastases were found. Inconsistencies in data recording hamper a valid interpretation of the results. Training data managers and preliminary analyses of the collected data allow for identification of and sharing information on common mistakes.

Furthermore, it is crucial to ensure patients' privacy, in particular for patient identifiers. Training in Good Clinical Practice (to the extent the principles are relevant for patient registries) and awareness of (inter-)national and local regulations will help in designing a registry that guarantees patient privacy. This includes anonymisation or pseudonymisation of data to ensure that information cannot be traced to an individual patient. Anonymisation may hamper specific registry functionalities (e.g. combining different data sources). Pseudonymisation involves replacing identifying items by artificial identifiers, or pseudonyms. Pseudonymisation can be performed by a trusted third party, guarding the encryption to the procedure while enabling re-identification when required. However, even in case a trusted third party is used, the inclusion of patient identifiers in the CRF should be carefully scrutinised and allowed only when absolutely necessary; approval should be obtained from a medical-ethical committee.

How should pharmacovigilance be incorporated?

Patient registries have the potential to reveal unique pharmacovigilance information because their follow-up allows identification of long-term toxicity.

Moreover, real-world toxicities may differ from toxicity profiles in clinical trials because of differential populations, treatment patterns, adverse event handling, and clinician experience.[125] However, it can be challenging to comprehensively collect safety data within a registry, especially in case data are collected retrospectively.

With respect to pharmacovigilance requirements, the EMA guideline on good pharmacovigilance practices differentiates between noninterventional postauthorisation studies with primary data collection and noninterventional postauthorisation studies based on secondary use of data.[126] First, in case of postauthorisation studies with primary data collection, 'for all collected adverse events comprehensive and high quality information should be sought in a manner which allow for valid individual case safety reports to be reported within the appropriate timeframes'.[126] These time frames are intended to allow manufacturers and authorities to take immediate action when needed to prevent serious adverse events occurring in other patients. However, this requires a clear workflow and an appropriate infrastructure. Second, in case of secondary use of data (e. g., medical chart reviews), the reporting of suspected adverse reactions in the form of individual case safety reports is not required; 'reports of adverse events should be summarized as part of any interim safety analysis and in the final study report unless the protocol provides for different reporting'.[126] The distinction between noninterventional postauthorisation studies with primary data collection and noninterventional postauthorisation studies based on secondary use of data, and its consequences regarding pharmacovigilance, was not always interpreted similarly between stakeholders in some of the registries in which we are involved. This has resulted in substantial registration burden (e.g. reporting within 24 hours of recording) under pressure from manufacturers.

Designing a solid plan for pharmacovigilance is part of setting up any patient registry. This plan needs to be consistent with national and international guidelines, and agreed upon by all stakeholders and the relevant medical-ethical bodies. Ideally, all safety information should be registered and reported by the clinician at the moment of occurrence.

It may be difficult to comprehensively collect safety information within a registry, while being dependent on the available data sources. It may be impossible to determine causality without involving the treating clinician. It is therefore crucial to have short communication lines with treating clinicians, and ensuring medical expertise in the study team is recommended. Alternatively, adverse event reporting can be outsourced to knowledgeable hospital personnel.

Interim analyses in the prostate cancer registry (CAstration-resistant Prostate cancer RegIstry) revealed that about half of the patients had a recorded hospitalisation or death during treatment. Although this percentage included both related and unrelated adverse events, all needed to be reported (see Table 6.4). This illustrates that serious adverse events are common and may significantly add to data management time and thus costs of running a registry. However, it also emphasizes that pharmacovigilance may be an important aspect in improving patient health.

Lessons learned

Patient registries provide valuable information on real-world patients, real-world practice, real-world costs, real-world effects, and real-world cost-effectiveness. If well designed and well executed, registries can support decision making at different levels. Regulatory authorities and local reimbursement agencies can use real-world data in market access and reimbursement decisions. Furthermore, sharing real-world outcomes can improve decision making at the patient level, and, ultimately, can improve patient health.

Because patient registries can serve multiple goals and inform decision making at different levels, practical guidance in setting up a registry is important to ensure a proper design and execution. This article provided practical guidance on 'the Why,' 'the Who,' 'the What,' and 'the How' in setting up a patient registry, which is based on our experiences and involvement in multiple registries in the Netherlands for various types of cancer. It is essential to cooperate with all relevant stakeholders and collect the right data from the right patients in the right way. The 'right' is, however, not always the most extensive approach. It is crucial that the registry is designed in such a way that it serves its aims and is as efficient as possible. It is, therefore, particularly important to balance the optimal and the feasible to maximize the gains within the constraints of the available resources.

This article has a number of limitations. First, our experiences in setting up patient registries are based on registries in cancer only; nevertheless, we believe that this practical guidance is applicable to patient registries in other disease areas. In addition, in most of the registries in which we are involved, patients were selected using existing databases, such as the Netherlands Cancer Registry, and most of the registries were largely informed by chart reviews conducted by trained data managers. Nevertheless, we believe that our experiences in the Netherlands will benefit researchers in other contexts and other countries.

Future prospects of registries

The number of patient registries will continue to rise in the near future.[127] Their importance was shown in many areas including general practice,[128] neurology,[129,130] orthopedics,[131,132] and oncology.[133,134]

Various initiatives exist that facilitate designing high-quality registries, such as the High-Value Health Care Project[135] and the cross-border PAtient REgistries iNiTiative (PARENT) project. The PARENT project supports member states of the European Union with the implementation of interoperable patient registries and created a registry of registries available online.[136]

Several trends may influence the design of future patient registries. First, there will be a further evolution of data standards and an improvement in interoperability of registries with electronic health records.[137] Moreover, there is an increasing trend in setting up multi-institution and multicountry registries.[138] Especially in rare diseases, multicountry registries are needed to include sufficient numbers of (comparable) patients. Finally, the content of registries will reflect important clinical developments (e.g. biobanking).[139]

Considering the unique value of and increasing demand for real-world evidence, we expect that patient registries will become the new standard alongside randomised controlled trials.

Chapter 7

Guidance beyond the guidelines: Practical recommendations in constructing a full disease model spanning multiple treatment lines to support cost-effectiveness analyses

H.M Blommestein*; S. de Groot*; W.K. Redekop; C.A. Uyl- de Groot

Submitted

* Both authors contributed equally to this manuscript

Abstract

Objective: Although full disease models comprising multiple treatment lines are increasing in importance, experience is scarce and while guidelines for good modelling practices are available, they are sometimes too brief. This study provides practical recommendations in constructing a discrete event simulation (DES) model to support real-world cost-effectiveness analyses (CEAs) of treatment strategies spanning multiple treatment lines.

Methods: Based on experiences with two DES models used in CEAs of treatment strategies in cancer, we discuss how best practices, mainly derived from the ISPOR-SMDM Task Force, can best be implemented. Additional recommendations were provided wherever best practices were unavailable or not applicable.

Results: Modelling multiple treatment lines using a DES model and real-world data imposes several challenges. First, it is necessary to correct effectiveness and costs for patient characteristics and the effect of previous treatment. This could be addressed by including patient characteristics and the effectiveness of previous treatments as covariates in survival models. Second, when modelling a heterogeneous population, valid extrapolation of survival outcomes beyond observation is required. This could be achieved by using multiple survival models. Third, the timing of competing events needs to be addressed appropriately. As recommended by the Task Force, one single survival model should be used together with a regression technique to determine event type.

Conclusions: Developing good-quality models comprising multiple treatment lines requires guidance beyond the existing guidelines and practical recommendations are currently lacking. The guidance based on hands-on experience with two DES models can improve validity and credibility of future disease models and CEAs.

Introduction

As more treatments become available (within and beyond treatment lines), traditional economic evaluations may not provide sufficient information, since these do not assess costs and effects of treatment strategies spanning multiple treatment lines and are not able to determine the optimal order (i.e. sequence) in which treatments should be provided. As a consequence, full disease models comprising multiple treatment lines are expected to increase in importance, but experience is scarce. Tosh et al called for a methodological framework for economic evaluations of sequential therapy for chronic conditions, since they found that methods have not been consistently applied, which has led to varied estimates of cost-effectiveness and uncertainty in respect of the most appropriate analytic methods.[140] Although guidelines for good modelling practices are available including the series commissioned by the ISPOR-SMDM Task Force,[141] they are sometimes too brief to help researchers develop models that are valid and credible.

This study provides practical recommendations in constructing a discrete event simulation (DES) model to support cost-effectiveness analyses of treatment strategies spanning multiple treatment lines. Best practices derived from the ISPOR-SMDM Task Force and additional sources are cited, followed by a description of how these were implemented in our DES models to estimate the real-world cost-effectiveness of new treatments in metastatic renal cell carcinoma (mRCC) and multiple myeloma (MM).

Case studies

Metastatic Renal Cell Carcinoma

115,200 patients were diagnosed with kidney cancer in Europe in 2012.[18] Renal cell carcinoma represents 80% of all kidney cancers. Median overall survival (OS) of patients with advanced disease is 43, 27 and 8.8 months for patients with a favourable, intermediate or poor prognosis, respectively.[55] Health outcomes are influenced by prognostic factors.[38]

A number of first- and second-line targeted therapies (e.g. sunitinib, sorafenib and everolimus) for mRCC have been introduced since 2006.[55] These therapies improve health outcomes, such as progression-free survival (PFS) and OS.[25,34,40,41,44,45] However, a Dutch population-based registry showed that almost half of the patients presenting with mRCC did not receive any targeted therapy.[142] A DES model was developed to study the real-world cost-effectiveness of several treatment strategies applied in patients with mRCC comprising one or more sequentially administered drugs. Potential health outcomes and costs of hypothetical treatment scenarios were calculated by

assuming that all treatment-eligible patients were treated according to a particular treatment strategy.

Multiple myeloma

In 2012, 38,900 patients were diagnosed with MM in Europe.[18] MM is a heterogeneous disease with a wide variation in OS.[143,144] Depending on the stage of the disease, median OS ranges from 29–62 months.[145]

Like for many cancers, treatment of MM is characterised by sequential treatment lines aiming to prolong PFS and OS. In the past decade, several treatment options have become available including the thalidomide-, bortezomib- and lenalidomide-based regimens. While most of these treatments were first recommended as treatment for third or subsequent lines, they are now recommended as induction therapy. Health outcomes are also influenced by prognostic factors, mainly patient and disease characteristics.[146,147] A DES model was developed to study the real-world cost-effectiveness of sequential use of novel agents for elderly MM patients.[116] Furthermore, by studying treatment sequences, we aimed to identify the optimal treatment strategy.

Comprehensive data on patient and disease characteristics of patients with mRCC and MM, as well as data on treatments and outcomes were collected in two population-based registries, the mRCC registry (PERCEPTION) and the MM registry (PHAROS).[92,148]

Model structure and design

Best practice:

> *"If, […], a valid representation of any aspect of the decision problem would lead to an unmanageable number of states, then an individual-level state-transition model is recommended".*[149]

> *"DES is an attractive option in nonconstrained models […] when individual pathways through the model are influenced by multiple characteristics of the entity; and when recording individual entity experience is desirable".*[150]

The first stage in developing a decision model involves choosing an appropriate model structure. According to the best practice commissioned by the ISPOR-SMDM Task Force, DES is the preferred modelling method if it is difficult to model the disease course of the average patient, and when the course of the disease, including its treatment, would require too many health states. As stated in the previous paragraph, patients with mRCC and MM in daily practice represent a heterogeneous population, and characteristics of these patients have a large impact on the costs and effects of treatment. In order to incorporate individual patients and allow for variability between patients, DES models were developed to calculate the cost-effectiveness of various treatment scenarios as recommended by the ISPOR-SMDM Task Force.[149,150] DES models allow individual patients to have their own characteristics, such as age and health state, which may also change over time.[151]

Furthermore, treatment of both mRCC and MM is characterised by sequentially administered drugs. Instead of modelling single treatment options, a comparison of complete treatment strategies was needed. DES models can easily include the effect of previous therapies, in contrast to Markov models, which cannot incorporate history of patients without constructing a large amount of health states. Therefore, a DES model seemed a better choice for modelling treatment strategies spanning multiple treatment lines for mRCC and MM.[149] Caro et al also argued that a DES provides an alternative, more natural, way to simulate clinical reality, whereas a Markov model requires all aspects of a disease including patient and disease characteristics and treatment history to be captured in a health state.[141] Although various methods exist which can include memory in Markov models (e.g. tracker variables), the required number of tracking variables would have been quite large in a model of sequential therapies.

In addition, data from the mRCC registry and MM registry revealed that some patients died very soon after treatment was initiated while some patients survived much longer. In a microsimulation Markov model, patients can only experience one transition per cycle and this would require many cycles with a small cycle length, which favoured a DES model allowing to include time continuously.

The DES models for mRCC and MM comprised entities (i.e. patients), attributes assigned to the entities, and events. Attributes were obtained from patient-level data from either the mRCC or MM registry by selecting clinical factors, biochemical and haematologic factors known to impact mRCC or MM outcomes, respectively. Events were either second-line treatment, third-line treatment (in the MM model only) or death. The time horizon of the models spanned the patients' lifetime. The structures of the mRCC and MM model are presented in Figures 7.1 and 7.2. Characteristics and sources for input parameters of both models are presented in Table 7.1.

Time-to-event

Best practice:

> *"It is [...] very important to justify the particular extrapolation approach chosen, to demonstrate that extrapolation has been undertaken appropriately and so that decision makers can be confident in the results of the associated economic analysis".*[152]

As survival data is often not fully observed, extrapolation beyond the observation period is needed. The method to extrapolate this data should be chosen in a systematic way in order to ensure valid and clinical plausible extrapolation. Time-to-event data derived from either the mRCC or MM registry were extrapolated using a range of parametric models (Exponential, Weibull, Loglogistic, Log-normal, Gamma and Gompertz). These models were assessed for their goodness of fit to the data using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Furthermore, each parametric function was assessed graphically as recommended by Latimer.[152]

For the mRCC model, a loglogistic distribution best fitted the time to the first event (TTE1) and the time to the second event (TTE2). Nevertheless, visual inspection showed that TTE1 was underestimated after 12 months. Additionally, as a consequence of the functional form of this distribution, mean TTE1 was highly influenced by a small proportion of the population with very long TTE1 estimates. Therefore, an alternative model (i.e. exponential) was chosen for TTE1 after 12 months, based on the AIC and BIC. This approach was also conducted by Leunis et al, who specified different survival models for different time periods.[153] For the MM model, a Weibull distribution best fitted the time to the first, second and third event. A Weibull distribution had the best goodness of fit based on the AIC/BIC and was also considered appropriate based on visual inspection.

Competing events

Best practice:

> *"Where feasible, when estimating times to competing events, methods of analysis that estimate the timing of competing events jointly are preferred to approaches that estimate separate time to event curves for each event".*[150]

In survival analysis, competing events are present when an individual is at risk of several different types of events but can have only one event at a time.[150,154] For the mRCC and MM models, time to next treatment was calculated from patient level data. Since some patients died before a new treatment was initiated, next treatment and death were competing events. Generally, there

Figure 7.1 Model structure – metastatic renal cell carcinoma model

Note: Covariates between brackets were considered for inclusion in the survival models and logistic regression model, but excluded through backward and/or forward selection.

are two approaches to analyse competing events.[150] The first approach is to perform separate survival analysis for each event where the other event is treated as censored. Then, for each event a time is sampled, with the patient moving to the event with the shortest time. The second approach is to perform one single survival analysis but make no distinction between the competing events; a separate sampling process in the model determines which event a patient will experience.

The second approach was adopted in the mRCC and MM models, as recommended by the ISPOR-SMDM Task Force.[150] Whereas survival analysis assumes that censoring is non-informative, we hypothesised that death was mostly treatment-related and, as a consequence, that censoring the patients who died might have altered the probability of experiencing a next treatment. Furthermore, the graphic presentation and interpretation of the single survival analysis are straightforward whereas the interpretation of the two separate survival curves is less intuitive.

While the second approach was adopted in the mRCC and MM models, we applied the first approach to validate our results and hypothesis. Interestingly, both methods yielded very similar results in the models. Since there is no difference between the two methods in terms of ease or speed (i.e. both methods require the estimation of two statistical models), we decided to align with current guidelines.

Assigning patient and disease characteristics

Best practice:

> *"The expected costs and benefits across the sampled group [...] provide an unbiased estimate provided that a sufficiently large sample is simulated and any covariance between the different patient characteristics is correctly taken into account".*[155]

While the mRCC and MM registries provided patient level data, simulation of the population including patient and disease characteristics was needed to study what would have happened to a patient if that patient had been treated differently. No recommendations were made by the ISPOR-SMDM Task Force about simulating a population. However, Davis et al emphasises the need to account for the covariance between patient and disease characteristics.[155]

In the mRCC and MM models, patient and disease characteristics were simulated similarly; random numbers were drawn from predefined distributions. These distributions were derived from patient-level data from the mRCC registry or the MM registry, respectively.

Figure 7.2 Model structure – multiple myeloma model

To account for the covariance between characteristics, different distributions were used for patients with a different prognosis. For example, in the mRCC model, first WHO performance status before treatment was simulated by drawing random numbers from a predefined distribution as obtained from the PERCEPTION registry. In simulating additional patient and disease characteristics; for each characteristic, a different distribution was used for patients with a WHO performance status of 0-1, and for patients with a WHO performance status of 2-4. This method was adopted to increase the likelihood that the combination of patient and disease characteristics per individual matched the original data.

Besides patient and disease characteristics, treatment needed to be assigned to each patient. Two multinomial logistic regression models (i.e. one to assign first-line treatment and one to assign second-line treatment) were used, including patient and disease characteristics as well as treatment history as covariates, to assign real-world treatment patterns to the patients in the mRCC model. This process guaranteed that the patients who received the treatments in the model were similar to the patients who received these treatments in daily clinical practice. This method was not feasible in the MM model, since some novel agents (i.e. bortezomib and lenalidomide) were prescribed to very few patients during the follow-up period. These numbers were too small to run a multinomial logistic regression model. Treatment in the MM model was therefore simulated in the same way as patient and disease characteristics; the probability of receiving a certain treatment was based on the distribution of treatments as observed in daily clinical practice using different distributions for patients with a different WHO performance status.

Having simulated patient and disease characteristics, and treatment for all patients, the patient's time to an event (either TTE1, TTE2 and TTE3) was estimated taking these characteristics into account.

Accounting for previous therapies

The mRCC and MM model aimed to calculate the cost-effectiveness of several treatment strategies comprising one or more sequentially administered treatments. Therefore, it was important to correct for the effectiveness of previous treatments when estimating the effectiveness of second- and third-line treatments. In addition, the effectiveness of subsequent therapies should be taken into account when estimating overall survival of first- and second-line treatments. Ideally, the effectiveness of, for example, bortezomib after thalidomide is solely based on patients treated with thalidomide followed by treatment with bortezomib. Although the registries included a substantial number of patients, these were not adequate to provide a stable estimate of

Table 7.1 Model characteristics and sources for input parameters of the DES models

	Metastatic renal cell carcinoma	Multiple myeloma
Model characteristics		
Aim	To model real-world cost-effectiveness for patients with metastatic renal cell carcinoma	To model real-world cost-effectiveness for elderly patients with multiple myeloma
Perspective	Healthcare	Healthcare
Patients	Patients with metastatic renal cell carcinoma	Elderly patients with multiple myeloma
Outcomes	Effects (OS and QALYs) and costs (€)	Effects (OS and QALYs) and costs (€)
Model type	Discrete event simulation	Discrete event simulation
Time horizon	Lifetime	Lifetime
Parametric distribution	Loglogistic and exponential distribution (line one) and loglogistic distribution (line two)	Weibull distribution for all lines
Disease pathways (base-case)	Real-world treatment including two subsequent lines	Real-world treatment including three subsequent lines
Disease pathways (scenarios)	Hypothetical pathways including two lines of treatment No targeted therapy Sunitinib - Sorafenib Sunitinib - Everolimus Sunitinib - Other	Hypothetical pathways including three lines of treatment MP-thalidomide-bortezomib MP-thalidomide-lenalidomide Thalidomide-bortezomib-lenalidomide Thalidomide-lenalidomide-bortezomib
Sensitivity analysis	Univariate and probabilistic sensitivity analyses (1,000 simulations)	Univariate and probabilistic sensitivity analyses (1,000 simulations)
Sources for input parameters		
Data (Patient and disease characteristics, treatment effects and patterns, healthcare utilisation)	Real-world data from the mRCC registry (PERCEPTION)	Real-world data from the MM registry (PHAROS)
Unit prices	Dutch reference price lists and literature	Dutch reference price lists and literature
Discount rates	Dutch guidelines (effects 1.5%, costs 4%)	Dutch guidelines (effects 1.5%, costs 4%)
Utilities	Literature	Cross-sectional study

Abbreviations: OS, overall survival; QALYs, quality-adjusted life years; MP, melphalan prednison.

the effectiveness of all treatment sequences. Recommendations to account for previous therapies in full disease models do not exist.

Therefore, we chose to correct for the effectiveness of previous therapies by including the TTE of the previous line in estimating the TTE of the subsequent therapy. For example, TTE1 was included in the parametric survival model estimating TTE2. This allowed us to obtain the effectiveness of second-line treatment accounting for the effectiveness of first-line treatment given the patient's characteristics.

In the MM model TTE1 had a significant association with TTE2 as well as with the type of event. The coefficient corresponding to TTE1 was not treatment specific, i.e. a TTE1 of 2 months obtained by treatment with thalidomide is similar to a TTE1 of 2 months obtained by treatment with bortezomib. Since adding type of treatment to the model did not improve its explanatory value,

we believe this method can be used to correct for the effectiveness of previous therapies.

Costs and outcomes

Best practice:

> *"Costs and quality of life weights are attached to events and time spent with different health conditions to estimate long term costs and health outcomes".*[150]

Total life years (i.e. OS) were calculated by summing TTE1, TTE2 (and TTE3). Besides total life years, total quality-adjusted life years (QALYs) were calculated by weighting LYs for the quality of life during these years using utility weights. In the mRCC model, various utility weights were used for patients with a favourable or intermediate prognosis, and patients with a poor prognosis before either first-line therapy or second-line therapy since their quality of life was expected to differ. Treatment-specific (including the effect of adverse events) or utility weights for different risk groups (or disease stages) were unavailable for elderly patients with MM. Therefore, an average utility weight was used in the MM model, obtained from a Dutch population-based cross-sectional study in MM.

Based on real-world data, average treatment-specific resource use per month was obtained in order to calculate total costs per month, e.g. the number of outpatient visits per month for mRCC patients treated with sunitinib or the number of hospital days per month for MM patients treated with thalidomide. Average total costs per patient were calculated by multiplying treatment specific total costs per month with TTE.

Discounting

Best practice:

> *"A (common) real discount rate should be applied to future costs and, when used in a cost-effectiveness analysis, to future outcomes".*[156]

> *"Discounting methods should accord with general guidelines for economic evaluation".*[157]

Future effects and costs should be converted to their present value in order to account for factors such as time preferences and uncertainty. As recommended, future costs and effects in the mRCC and MM model were discounted to their present value using discount rates based on the Dutch guideline for pharmacoeconomic research.[98]

While a Markov model with a fixed cycle length provides a convenient structure to discount future costs and effects, discounting future costs and effects in a DES model including treatment strategies comprising one or more sequentially administered treatments is more challenging. First, a DES model produces individual TTE estimates, and as a consequence LYs and QALYs need to be discounted for each patient separately. Furthermore, different utility values were assigned to the treatment lines in the mRCC model and therefore, total QALYs needed to be discounted for each treatment line separately. Second, unit costs per month differed between treatment lines. For example, a patient with MM treated with melphalan-prednisone, followed by a bortezomib-based regimen, and then followed by a lenalidomide-based regimen, incurs different hospital and drug costs per month during first-, second- and third-line treatment. As a consequence, total costs need to be discounted for each treatment line separately.

Since total costs per treatment line were obtained by multiplying unit costs per month by the corresponding TTE, it was decided to discount TTE and multiply unit costs per month by the discounted TTE. The same approach was adopted to discount future QALYs in the mRCC model. While discounting time was a convenient approach in our models, this is not possible for DES models where costs are obtained from multivariable regression models. Since these models include undiscounted time as an explanatory variable, it is not possible to calculate and discount the total costs per line. Instead, costs should be calculated and discounted for different time frames, e.g. per year, which adds both complexity and computational burden to the model.

Probabilistic sensitivity analysis

Best practice:

> *"The inner loop evaluates the outcomes across the simulated population for the given parameter values, and the outer loop samples those parameter values to reflect uncertainty in the model inputs. In a cohort-level model, only the outer loop is required, thus PSA [probabilistic sensitivity analysis] computation time for a cohort-level model is likely to be lower than for an equivalent patient-level model".*[155]

In other words, the inner loop aims to calculate costs and effects for one simulated population (with constant patient and disease characteristics), whereas the outer loop changes all input parameters according to their probability distributions to examine the impact of the joint uncertainty across all input parameters. In the model for mRCC and MM, the values of input parameters varied across simulations. Due to the probabilistic structure of the models,

the values of input parameters could also vary within one simulation. For example, an inpatient day could cost €402 while calculating costs of treatment scenario A and €646 while calculating costs of treatment scenario B. Furthermore, patient and disease characteristics could vary within one simulation. For instance, 26% of the population could be assigned a WHO performance status of 2-4 while calculating costs and effects of treatment scenario A and 35% could be assigned a WHO performance status of 2-4 while calculating costs and effects of treatment scenario B. However, in each single simulation, parameters that are not related to a certain treatment scenario should have the same values in all treatment scenarios. This approach reduces the 'noise' or random variation that is introduced by setting unit costs and patient and disease characteristics twice in each simulation, once for Scenario A and once again for scenario B. It also increases the model's efficiency, since fewer simulations are needed to get a stable estimate of the incremental cost-effectiveness ratio (ICER).

Apart from probabilistic sensitivity analysis, univariate sensitivity analyses can be performed to examine the impact of alternative input parameters on the ICERs as illustrated in the mRCC model.

Discussion

Economic evaluations mostly require a lifetime time horizon in order to capture all health and economic consequences.[3] Such a time horizon makes a full disease model including treatment strategies spanning multiple treatment lines inevitable. However, the development of the full disease models for mRCC and MM revealed several challenges, including the optimal ways to correct effectiveness and costs for patient characteristics (including the effectiveness of previous treatments), extrapolate survival outcomes beyond observation for a heterogeneous patient population, and estimate the timing of competing events. Best practices, including solutions to these challenges, were not always found in the literature. Also, existing disease models did not often provide suitable solutions to these challenges since these models differed in aim, characteristics of the disease or treatment varied and comprehensive data was unavailable. Therefore, guidance beyond the existing guidelines and practical recommendations are necessary to improve validity and credibility of future disease models.

Based on hands on experiences with two DES models the following recommendations can be made in constructing a DES model to support real-world cost-effectiveness analyses of treatment strategies spanning multiple treatment lines. First, the inclusion of patient characteristics (including the effectiveness of previous treatments) as covariates in survival models, makes

it possible to derive more valid estimates of costs and effectiveness. Second, using multiple survival models per treatment line ensures valid extrapolation of survival outcomes beyond observation for a heterogeneous population. Third, as recommended by the ISPOR-SMDM Task Force, when competing events exist, one single survival model should be used together with a regression technique to determine which event type will occur.

Although the mRCC and MM models enabled the estimation of the cost-effectiveness of several treatment strategies comprising one or more sequentially administered drugs,[116,142] these models could have been improved further. Based on our experiences, the following recommendations can be made to improve future models of treatment strategies spanning multiple treatment lines. First, in our models, patient and disease characteristics were simulated by drawing random numbers from predefined distributions. Covariance between characteristics was accounted for by using different distributions for patients with a different prognosis. This method does, however, not guarantee valid relationships between all patient and disease characteristics. Multivariable regression models could overcome this problem, as illustrated by Goossens et al in a study on propensity score matching.[65] This method generated patient and disease characteristics based on a set of regression models; one for each characteristic. While the first characteristic was defined using a predefined distribution, all other characteristics were simulated using regression models that included as covariates all of the characteristics already assigned to the patient; this preserved the covariance between the different characteristics that was observed in the original data.

Second, in the mRCC and MM models, total costs per patient were derived by multiplying mean monthly costs (per treatment) by the individual patient's time to an event. However, the distribution of cost data is skewed, which means that a limited number of patients is responsible for a high proportion of the costs.[3] As a consequence, by multiplying mean monthly by the individual patient's time to an event, total costs per patient might be overestimated. Again, multivariable regression models could solve this problem. Besides type of treatment and time to event, these models could include patient and disease characteristics as covariates to estimate total costs per patient. In this way, overestimation of costs will be prevented.

This practical guide is a first attempt to document how best practices in modelling, derived from the ISPOR-SMDM Task Force and additional sources, can be interpreted and implemented. Experiences in implementing best practices were based on two studies only; these studies had rather similar aims, they both focussed on treatment strategies in cancer, and the available data was comparable. Although treatment strategies spanning multiple treatment

lines are common in other disease areas (e.g. rheumatoid arthritis), we recommend further research to be done to ascertain whether this practical guide helps others with different goals working in other disease areas with different data sources. We therefore recommend them to share their findings from constructing and using a DES model including treatment strategies spanning multiple treatment lines.

It should be clear that a DES model is not necessarily the best choice when constructing a full disease model. While a DES model was a feasible option to study the cost-effectiveness of several treatment strategies comprising one or more sequentially administered drugs for patients with mRCC and MM, disadvantages of this model structure may include the type and amount of required data as well as the time needed for model building and simulation.[158] Data for the mRCC and MM model were derived from population based registries. In these registries comprehensive data were collected on patient and disease characteristics. In addition, compared to randomised trials, both registries had a long follow-up duration. This enabled us to study the impact of multiple treatment lines on overall survival. If comprehensive data is not available, a different model structure might be more appropriate. Additionally, the time needed for model building and simulation should be balanced against the benefits of modelling patients individually in a DES model.

Conclusions

In order to secure the validity and credibility of models, guidelines were developed summarising best practices in modelling.[149,150,152,155-157] Unfortunately, these guidelines are sometimes too brief to be used in constructing full disease models comprising multiple treatment line. Extra instruction is therefore needed. This study aimed to help in filling this gap by providing practical guidance on constructing a DES model. Specifically, it explores how to apply the guidelines by describing how they were actually implemented in two DES models, and it provides additional recommendations, which may help to further improve the validity of full disease models.

Chapter 8

General discussion

The growing pressure on healthcare budgets increases the need to make choices between the reimbursement of healthcare procedures, services and programs.[3] These choices are made using many criteria. Two of these criteria are effectiveness and cost-effectiveness.[4] However, information about effectiveness and cost-effectiveness is often associated with uncertainty at the time a new procedure, service or program is introduced.

The effectiveness of drugs is often demonstrated in large clinical trials (strictly speaking, the *efficacy* of drugs is demonstrated in clinical trials, i.e. the effect of drugs under ideal circumstance). Data from these trials are often used to estimate a drug's cost-effectiveness. However, as with other procedures, services and programs, the effectiveness and cost-effectiveness of new drugs are uncertain at the time they are introduced. As a consequence, conditional funding was implemented in the Netherlands in 2006 to guarantee equal and early access to expensive drugs, while data about uptake, use, and outcomes in clinical practice were to be collected. This funding policy also applied to two expensive inpatient drugs for metastatic renal cell carcinoma (mRCC), that is, bevacizumab and temsirolimus. The collection and analyses of data about uptake, use and outcomes of these new drugs for mRCC in clinical practice form the focus of the present thesis and this final chapter shows how this evidence contributed to the decision about future funding of these drugs. Additionally, this chapter illustrates the possibilities and impossibilities (or limitations) of evaluating uptake, use and outcomes in clinical practice, and reflects on the difficulties of making decisions about future funding (based on evidence from clinical practice).
This chapter is written along the lines of statements.

Positive decisions about reimbursement do not directly translate into uptake and use in clinical practice

As outlined in chapter 1, randomised controlled trials (RCTs) demonstrated that targeted therapies for mRCC improve progression-free survival.[25-27,29,40,44,46,47] As a consequence of these findings, several targeted therapies for mRCC have entered the market since 2006, such as the first-line therapies sunitinib, temsirolimus, bevacizumab and pazopanib, and the second-line therapies sorafenib, everolimus and axitinib.

Indications of the cost-effectiveness of bevacizumab and temsirolimus were provided to the Dutch National Health Care Institute by the manufacturers of the drugs in line with the then prevailing policy for expensive drugs. However, the Dutch Committee of Pharmaceutical Aid [in Dutch: Commissie Farmaceutische Hulp] concluded that the cost-effectiveness of these drugs was insufficiently substantiated.[159,160]

The manufacturer of sunitinib was exempted from providing evidence supporting the cost-effectiveness of the drug.[161] The manufacturer of pazopanib was also exempted from providing evidence on cost-effectiveness, because the therapeutic value of pazopanib was found to be similar to the therapeutic value of sunitinib, which was already added to the Medicine Reimbursement System [in Dutch: Geneesmiddelenvergoedingssysteem, GVS].[162] Despite the absence of (valid) estimates of the cost-effectiveness of the drugs for mRCC, all are reimbursed in the Netherlands. Bevacizumab and temsirolimus were initially funded on the condition that data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice were to be collected.

In contrast to the Netherlands, the UK National Institute for Health and Care Excellence (NICE) did not recommend bevacizumab and temsirolimus as first-line treatment options for patients with advanced and/ or metastatic renal cell carcinoma. Results from the Assessment Group model showed an incremental cost-effectiveness ratio (ICER) of £171,301 per quality-adjusted life year (QALY) gained for bevacizumab plus interferon alfa (IFN-α) compared with IFN-α alone, and an ICER of £94,385 per QALY gained for temsirolimus compared with IFN-α. NICE concluded that these therapies would not be a cost-effective use of National Health Service (NHS) resources.[163]

Similar to the Netherlands, sunitinib is recommended by NICE as a possible first-line treatment, based on an assessment of the drug's effectiveness and cost-effectiveness. Results from the Assessment Group model showed an ICER of £104,715 per QALY gained for sunitinib compared with IFN-α (including the agreed patient access scheme of the first cycle of sunitinib being free to the NHS). Since sunitinib met the criteria for being a life-extending end-of-life treatment, the appraisal committee concluded that sunitinib could be recommended as a cost-effectiveness use of NHS resources.[164]

Decisions about reimbursement guide the use of drugs in clinical practice, and this ultimately determines a drug's real-world value.[4] To assess a drug's real-world value, manufacturers of expensive inpatient drugs in the Netherlands were obliged to collect data regarding appropriate drug use, effectiveness and cost-effectiveness in clinical practice, during a four-year period of conditional funding. A population-based registry (i.e. PERCEPTION) was, therefore, created to include patients with mRCC. Despite the evidence of effectiveness and the indications of cost-effectiveness of bevacizumab and temsirolimus in the Netherlands, the PERCEPTION renal cell cancer showed that few patients with mRCC were treated with one of these drugs in Dutch

clinical practice. Most patients were treated with sunitinib; this targeted therapy has largely replaced IFN-α as first-line standard of care in the Netherlands. However, as discussed in chapter 2, many treatment-eligible patients did not receive sunitinib (or any other targeted therapy) in clinical practice. Since bevacizumab, temsirolimus and sunitinib are all recommended in Dutch clinical guidelines and reimbursed in the Netherlands, other factors must explain the low uptake and use. Positive decisions about reimbursement, thus, do not directly translate into uptake and use of new cancer drugs in clinical practice, as was stated before by Jönsson et al.[4]

Also in the United Kingdom (UK), the actual use of cancer drugs with endorsements from NICE is rather low.[165] For example, one in three patients with mRCC eligible for either sunitinib or pazopanib did not receive the drug.[50]

Sufficient patient numbers are needed to provide evidence on outcomes in clinical practice

Although a drug's effectiveness and cost-effectiveness (only for certain types of drugs) are assessed at the time the drug is introduced, evidence on outcomes in clinical practice are relevant in order to assess a drug's real-world value. Additionally, these outcomes might be able to reduce the uncertainty about a drug's effectiveness and cost-effectiveness. Information on outcomes in clinical practice was required for bevacizumab and temsirolimus in accordance with the Dutch policy for expensive inpatient drugs. However, little evidence became available during the four years of conditional funding, because only a few patients were treated with bevacizumab or temsirolimus in clinical practice. As a consequence, no assessment took place about the drugs' real-world values, given the drugs' limited budget impact for this indication.

The PERCEPTION renal cell cancer registry was disease-oriented, thereby making it possible to assess the real-world (cost-)effectiveness of other drugs, such as sunitinib. A reassessment of the effectiveness and cost-effectiveness of sunitinib in clinical practice was not required by the Dutch National Health Care Institute, since this only applied to expensive *in*patient drugs. Nevertheless, the PERCEPTION registry provided evidence on outcomes of sunitinib in clinical practice. As shown in chapter 2, median overall survival (OS) of patients treated with sunitinib in Dutch clinical practice was much shorter than the OS of patients treated with sunitinib in the pivotal phase III clinical trial. In clinical practice, the patient population was heterogeneous, whereas the outcomes of the clinical trial were based on a homogeneous and relatively young and healthy population (due to strict in- and exclusion criteria). Chapter 2 also provides important insight into the outcomes of sunitinib in subgroups, for example in patients who were not eligible to participate in

the pivotal phase III trial of sunitinib, such as patients with a worse performance status, brain metastases or a non-clear-cell subtype.

The relative effectiveness and cost-effectiveness of sunitinib versus other active treatments could not (reliably) be evaluated. In the pivotal phase III clinical trial, sunitinib was compared with IFN-α. Since sunitinib largely replaced IFN-α as first-line standard of care, a comparison between the two could not be made. Also a reliable comparison with other active treatments, such as bevacizumab plus IFN-α and pazopanib, could not be made, due to the small number of patients treated with these drugs. Nevertheless, since many treatment-eligible patients did not receive sunitinib (or any other targeted therapy) in clinical practice, a comparison between sunitinib and no treatment with targeted therapy could be made. Compared to a scenario in which none of the patients receives a targeted therapy, clinical practice, which is dominated by treatment with sunitinib, results in an ICER of €105,011 per QALY gained. As discussed in chapter 5, health can be gained if more treatment-eligible patients with mRCC receive targeted therapies (like sunitinib). However, the ICER is far beyond the upper limit of the cost-effectiveness threshold of €80,000 per QALY, as proposed by the Dutch Council for Public Health.[166]

Patient registries can provide important information about uptake and use in clinical practice

As illustrated with the PERCEPTION renal cell cancer registry, patient registries can provide important information about uptake and use of new therapies in clinical practice. If data on the use of new therapies in different time periods are collected, it is possible to see how rapidly their use increases after new (clinical) guidelines are published. Additionally, if data on patient and disease characteristics are captured in the registry, registries can provide insight into the use of therapies during end of life or in patient subgroups. For example, the PERCEPTION registry showed that targeted therapies are frequently used in patients with a non-clear cell subtype, even though the evidence base for targeted therapies in this subgroup is very limited. This information enables discussions about possible ineffective and/or unnecessary care. Lastly, if a registry is disease-oriented, it can provide additional information, for example on the frequency of untreated patients and how often these patients would have been eligible for treatment. This information can be used to study possible variation in the use of (new) therapies within or between hospitals, and facilitate discussions if therapies are used to their full potential.

Besides questions about uptake and use of new therapies in clinical practice, additional research questions can be answered using data from patient registries, especially registries that are disease-oriented. For example, in

chapter 3 large variation in clinical practice was observed in the use of a cytoreductive nephrectomy (CN) prior to treatment with sunitinib in patients presenting with mRCC. Results from this chapter suggest that a CN may be effective, but the absolute benefit may be modest after correction for time from diagnosis to sunitinib.

Although patient registries can provide important information about uptake and use in clinical practice, timely information is needed to enable a quick response if, for example, not all patients who are eligible to receive a certain drug actually receive the drug. Timely information about uptake and use can more easily be provided if continuous registries would exist.

Patient registries are seldom able to assess the relative effectiveness and cost-effectiveness of drugs

Although patient registries are an important source to provide insight into the uptake and use of new therapies in clinical practice, registries are seldom able to assess the relative effectiveness and cost-effectiveness of a drug. First, physicians may have profound reasons to choose a particular treatment over another, and these reasons will often relate to a patient's prognosis. As a consequence, it is not only treatment that causes a difference in outcome, but also the difference in prognosis. In general, there are two ways to reduce confounding in observational studies: prevention in the design phase (e.g. restriction or matching) and adjustment using statistical techniques (e.g. stratification or multivariable techniques).[167] Two methods have been applied in this thesis. In chapter 3, the effect of cytoreductive nephrectomy on overall survival in primary mRCC patients treated with first-line sunitinib was evaluated using propensity score matching, and in chapter 5, the real-world cost-effectiveness of several mRCC treatment strategies was estimated using a patient level simulation model and several multivariable techniques. Besides regression on a matched sample and multivariable regression as applied in this thesis, the NICE DSU technical support document on the use of observational data recommends some additional methods to estimate treatment effectiveness, such as regression adjustment, inverse probability weighting and doubly robust methods.[168] All these methods assume no selection on unobserved confounders, while it is almost impossible to assume that all possible confounders are observed (and measured), and have been adjusted for. In addition, these statistical methods are often complicated by missing values in patients' baseline characteristics, which is inherent to observational studies.

The NICE DSU technical support document also describes methods that can adjust for unobserved confounders, such as instrumental variable methods.[168]

With instrumental variable methods, the problem of selection on unobserved confounders is addressed by an instrumental variable, which is correlated with the treatment but only correlated with the outcome through its effect on treatment.[168] In practice, it might be difficult to find such a variable, because treatment choice often depends on a patient's prognosis, which also influences the outcome of treatment.

A second reason why patient registries are seldom able to assess the relative effectiveness and cost-effectiveness of a drug is the difficulty to include the minimum number of patients needed to detect true and clinically important differences between treatments. In chapter 3, the relatively small sample (i.e. 74 patients underwent cytoreductive nephrectomy prior to sunitinib) resulted in fairly wide confidence intervals for the hazard ratios for cytoreductive nephrectomy. In chapter 5 (about real-world cost-effectiveness), the cost-effectiveness of first-line sunitinib compared to first-line pazopanib (or bevacizumab plus IFN-α) could not be estimated, while this would have been relevant because these therapies are both recommended for mRCC patients with a favourable or intermediate prognosis. Pazopanib has only been recommended since 2010, which (partly) explains why an increase in its use was only seen from 2012. Since data collection for the PERCEPTION registry stopped at the end of 2013, only limited data were available on patients treated with pazopanib. Nevertheless, we expect the number of patients treated with first-line pazopanib to rise, given the results of a recent study that shows that pazopanib has a more favourable safety and quality-of-life profile.[88] Although sunitinib and pazopanib (and bevacizumab plus IFN-α) are both recommended for the same patient group, we expect the problem of insufficient patient numbers to grow in the near future, since new therapies often target small patient subgroups.

Lastly, observational studies are seldom able to assess drug interactions, where the effectiveness of a second-line therapy depends on the type of first-line therapy. The demand for such evidence increases with the advent of targeted agents intended for various lines of therapy. In chapter 5, a patient level simulation model was built, and time from diagnosis to second-line therapy (either obtained by first-line sunitinib or first-line bevacizumab plus IFN-α [the latter therapy was not evaluated in this thesis]) was taken into account when estimating the effectiveness of second-line therapy. In this way, the effectiveness of previous therapy was considered. However, this method assumes that the effectiveness of second-line therapy is not influenced by the type of previous therapy. This assumption most likely does not hold; for example, the effect of second-line sorafenib has been demonstrated in patients treated with first-line cytokines (e.g. IFN-α). Whether the effect of second-line sorafenib

is similar for patients treated with first-line sunitinib is unknown, but the effect might be less favourable, since sorafenib and sunitinib have a similar mechanism of action, i.e. both drugs primarily target tumour angiogenesis by inhibiting a variety of tyrosine kinases.[169] As a consequence, even larger patient numbers are needed to have sufficient statistical power to be able to detect a true difference between treatment strategies comprising one or more sequentially administered drugs.

Thus, although registries are able to provide valuable information about the uptake and use of new (expensive) therapies, registries should not be used to assess the relative effectiveness and cost-effectiveness of new therapies, unless patient numbers are sufficiently large and sufficient data on confounding variables (or instrumental variables) are available to reduce confounding. Continuous registries (or registries with a long follow-up) might be able to provide sufficient data from patients treated with standard or usual treatment and patients treated with the new drug. If patient numbers are small, collaboration between countries is inevitable.

Nevertheless, evidence on outcomes in clinical practice are only worthwhile if it would affect the decision about future funding, or if it would affect the quality of patient care in a different way.

Decisions about future funding based on evidence on outcomes in clinical practice are difficult

Besides the challenges associated with the evaluation of a drug's relative effectiveness and cost-effectiveness in clinical practice, it appeared to be difficult to decide about future funding based on these outcomes, after four years of conditional funding. In accordance with the Dutch policy for expensive drugs, data were collected on patients treated with bevacizumab plus IFN-α or temsirolimus, regarding appropriate drug use, effectiveness and cost-effectiveness. After four years of initial funding, the Dutch National Health Care Institute decided not to reassess these drugs, based on the rule that a reassessment was only required if the budget impact was at least 2.5 million Euros. However, for many other expensive inpatient drugs in various disease areas, an assessment of the drugs' real-world value (including effectiveness and cost-effectiveness) after four years of initial funding did not (yet) take place either or did not lead to an alternative decision. Recent research showed that the decision to stop reimbursement is more difficult than the decision to not reimburse a therapy from the start.[170] Only in some cases outcomes research influenced the decision-making process, e.g. a pay-for-performance agreement prevented the exclusion of omalizumab from the basic health insurance package after four years of conditional funding.

Additionally, as described before, evaluating outcomes in clinical practice is associated with limitations, which impact the internal validity of estimates of effectiveness and cost-effectiveness. This was for example seen in the UK multiple sclerosis risk sharing scheme; patients treated with interferon beta or glatiramer acetate were closely monitored to study long-term progression of disease, with an agreement that prices of these drugs would be reduced if patient outcomes were worse than predicted.[171] A direct benefit of the scheme was that drug costs were reduced to achieve a cost-effectiveness of £36,000 per QALY.[172] The first interim analysis after two years showed that patient outcomes were much worse than predicted.[171] However, it was argued that further follow-up and analyses were needed, because important methodological issues needed to be addressed, including the need for additional data about the comparator and the potential bias from missing data. Results at six years showed that patient outcomes (as reported in the clinical trials) were maintained, and that interferon beta and glatiramer acetate represent value for money.[173] As these findings were based on data from 4,137 patients who started a disease-modifying therapy in the UK, and data from 898 (historical) untreated patients from British Columbia, the question arises if a similar observational study in the Netherlands would have been able to demonstrate this effect. Additionally, this approach is limited if treatment paradigms change over time.

The value of a patient registry must outweigh its costs

The design and implementation of a patient registry can take a long time, as was experienced with the initiation of the PERCEPTION renal cell cancer registry. After the approval of the research protocol by the medical ethics committee, every single hospital started an evaluation of the study's local feasibility (using varying procedures). As a consequence, the recruitment of patients for the prospective part of the PERCEPTION registry started later than originally planned, and in fewer hospitals than originally planned. The administrative burden complicated the provision of timely information about uptake, use and outcomes in clinical practice, and led to higher-than-expected start-up costs.

Continuous patient registries (in contrast to time-restricted registries) might more easily provide timely information about uptake and use of new drugs in clinical practice, if data entry is rapid. The existence of continuous registries would avoid the preparation required to start a patient registry to answer every new research question, which would reduce the time needed to obtain answers to those questions. Additionally, continuous registries might

be able to provide data from patients treated with the standard or usual treatment, data which are usually unavailable when a registry starts when a new drug has been introduced. Since the introduction of a new drug may mean that patients will no longer receive standard or usual treatment, a comparison between a new drug and standard or usual treatment is hindered.

Nationwide patient registries could increase patient numbers (in comparison with registries restricted to certain regions or hospitals). If patient numbers are small, collaboration between countries is inevitable.

However, continuous nationwide registries will not solve all problems. First, problems associated with confounding remain, which will present important challenges when evaluating the relative effectiveness and cost-effectiveness of new drugs in clinical practice. Although sufficient data on confounding variables (or instrumental variables) combined with statistical techniques could reduce confounding, it will never minimise imbalances in confounding variables to the same extent as randomisation would. Second, expanding registries in both length (i.e. from a time-restricted registry to a continuous registry) and width (i.e. from a registry restricted to certain regions or hospitals to a nationwide registry) will increase the time and costs of data collection. Collecting data on confounding variables (or instrumental variables) will further increase costs. It is, therefore, important to consider the overall costs of a registry in relation to its value.

Costs might be reduced if patient registries could be maintained digitally using electronic health records, instead of having data managers copying data from individual patient records into the registry. One type of electronic health record in all hospitals across the country (instead of different types) might ease this process.

General conclusion and policy implications

In line with the current policy on so-called specialist drugs (including expensive inpatient and outpatient drugs),[13] the present thesis argues to limit the use of conditional funding and put more emphasis on initial reimbursement decisions, because of the limitations of evidence on outcomes in clinical practice, and the difficulties with making decisions about future funding based on these outcomes. Results of cost-effectiveness analyses should be used to inform (initial) reimbursement decisions. Only in specific cases (e.g. orphan drugs) outcomes research is able to provide a more robust estimate of cost-effectiveness. These cases should be selected carefully in order to minimize the costs of extensive data collection.

A notable problem with the available evidence at the initial reimbursement decision is that new drugs are not always compared with standard or usual care. As a consequence, policy makers do not always have the information they need to decide about reimbursement. This problem might be solved using indirect comparisons to estimate the effectiveness and cost-effectiveness of a new drug compared with standard or usual treatment.

It is important to use the results of cost-effectiveness analyses to keep our healthcare system financially sustainable as the number of cancer patients increases and the number of innovative, expensive drugs grows. If cost-effectiveness ratios are above a given (equity weighted) threshold that reflects societal willingness to pay for health gains, price negotiations (or other measures) are essential to improve the relationship between costs and effects of a new drug, and allow reimbursement from the basic health insurance package.

For sunitinib, an estimate of its cost-effectiveness was not made at the time of the initial reimbursement decision. Moreover, an assessment of the drug's real-world value after four years was not required. The results of a cost-effectiveness analysis could have been used to negotiate with the manufacturer about the price of the drug, as was for example seen in the UK where the first cycle of sunitinib is free to the NHS. Although the budget impact of sunitinib is much larger than the budget impact of bevacizumab or temsirolimus for mRCC, it is still relatively small compared to new drugs for cancer types like non-small cell lung cancer or breast cancer. Nevertheless, combined savings in small disease areas could be substantial, and make money available for other uses in the healthcare system.

Recently, in the Netherlands, results of a cost-effectiveness analysis were used to negotiate about the price of nivolumab, a new drug for non-small cell lung cancer. Since the cost-effectiveness of nivolumab was estimated to be 134,000 Euros per QALY gained,[174] the Dutch National Health Care Institute recommended not to reimburse nivolumab unless the manufacturer was willing to reduce the price of nivolumab to improve the ICER. After price negotiations with the manufacturer led to a reduction that was kept confidential, the Dutch Minister of Health decided to include nivolumab for treatment of patients with non-small cell lung cancer in the basic health insurance package. Interestingly, the financial arrangements she made with the manufacturer also apply to all other (future) indications for treatment with nivolumab, including melanoma and mRCC. However, it is currently unknown whether the price reduction of nivolumab will lead to ICERs in melanoma and mRCC that are similar to the ICER of nivolumab for non-small cell lung cancer, since cost-effectiveness analyses in these indications are still to be conducted.

While there certainly are limitations to the use of data about outcomes in clinical practice in decisions about future funding, registries such as the PERCEPTION renal cell cancer registry can be of great value. Information about

uptake and use of new drugs, as derived from these registries, could more often be used to monitor whether these drugs are used in patients in whom the drug is known to be effective and cost-effective, and to evaluate if all patients who are eligible to receive the drug actually receive the drug. This could help to determine if the drug is being used to its full potential. The Dutch National Health Care Institute has recently started such a programme ('Meaningful care' [in Dutch: Zinnige Zorg]). The aim of this programme is to identify and discourage ineffective or unnecessary care in order to improve the quality of patient care, improve health outcomes and avoid unnecessary costs. Patient registries appeared to be a valuable source to fulfill this aim.

Limitations and research implications

Several limitations of the research presented in this thesis need to be mentioned, in addition to the limitations that are already mentioned in the separate chapters.

First, not all available statistical techniques to estimate treatment effectiveness based on observational data have been applied in this thesis, especially those, such as instrumental variable methods, that can adjust for unobserved confounders. In practice, it might be difficult to find an instrumental variable, which is correlated with the treatment but only correlated with the outcome through its effect on treatment. Nevertheless, it is worthwhile to test its usefulness in future analyses of outcomes in clinical practice.

Second, as this thesis showed, conditional funding (i.e. the condition to collect data on effectiveness and/or cost-effectiveness in clinical practice) has its limitations. As a consequence, in addition to the research by Claxton et al[175] further research might indicate under what circumstances conditional funding may be an appropriate measure to guarantee early access, while additional evidence on outcomes in clinical practice are to be collected. Furthermore, other policy measures should be studied more closely, such as a more structural application of the so-called 'sluice' for expensive inpatient drugs. Due to this sluice expensive drugs do not automatically enroll in the basic health insurance package. Instead, the Dutch Minister of Health first seeks advice from the Dutch National Health Care Institute, and then takes measures for a (financially) sound introduction of the new drug. This measure was first applied in 2015 with the introduction of nivolumab.

Final remarks

This research revealed that many treatment-eligible mRCC patients did not receive any targeted therapy. This was also seen during the evaluation of the Dutch clinical guideline in 2011/12. As a consequence of this evaluation, clinicians stressed the need to evaluate the use of targeted therapies in clinical practice using up-to-date data. The Dutch National Health Care Institute has recently started this evaluation ('Meaningful care' [in Dutch: Zinnige Zorg]) based on data from the PERCEPTION registry, complemented with data from other sources. Additionally, plans are currently underway to start a new renal cell cancer registry to systematically monitor the use of targeted therapies and thereby supplement the data currently captured in the Dutch Cancer Registry.

References

1 Uyl-de Groot CA, de Vries EGE, Verweij J, Sullivan R. Dispelling the myths around cancer care delivery: It's not all about costs. Journal of Cancer Policy 2014;2(1):22-29.

2 Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. Lancet Oncol 2013;14(12):1165-1174.

3 Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford: Oxford University Press 2015.

4 Jönsson B, Ramsey SD, Wilking N. Cost effectiveness in practice and its effect on clinical outcomes. Journal of Cancer Policy 2014;2(1):12-21.

5 Littlejohns P, Weale A, Kieslich K, Wilson J, Rumbold B, Max C, Rid A. Challenges for the new Cancer Drugs Fund. Lancet Oncol 2016;17(4):416-418.

6 Rijksinstituut voor Volksgezondheid en Milieu. Cost of Illness [in Dutch: Kosten van Ziekten]. Available at: https://www.volksgezondheidenzorg.info/cost-of-illness. Accessed January 11, 2016.

7 Signaleringscommissie Kanker van KWF Kankerbestrijding. Accessibility of expensive cancer drugs, Now and in the future [in Dutch: Toegankelijkheid van dure kankergeneesmiddelen, Nu en in de toekomst]. 2014.

8 Coalition agreement cabinet Rutte-Asscher, Building bridges [in Dutch: Bruggen slaan, Regeerakkoord VVD - PvdA]. 2012.

9 College voor zorgverzekeringen. Assessment procedure for the reimbursement of inpatient drugs [in Dutch: Procedure beoordeling intramurale geneesmiddelen]. 2006.

10 Nederlandse Zorgautoriteit. Policy rule BR/CU-2035, performance and rates for specialist medical care [in Dutch: Beleidsregel BR/CU-2035, prestaties en tarieven medisch specialistische zorg]. 2012.

11 Nederlandse Zorgautoriteit. Policy rule BR/CU-2045, performance and rates for specialist medical care [in Dutch: Beleidsregel BR/CU-2045, prestaties en tarieven medisch specialistische zorg]. 2012.

12 Nederlandse Zorgautoriteit. Policy rule BR/CU-2125, performance and rates for specialist medical care [in Dutch: Beleidsregel BR/CU-2125, prestaties en tarieven medisch specialistische zorg]. 2015.

13 College voor zorgverzekeringen. Specialist drugs package management [in Dutch: Pakketbeheer specialistische geneesmiddelen]. 2013.

14 College voor zorgverzekeringen. Conditional access/ funding of care [in Dutch: Voorwaardelijke toelating/ financiering van zorg]. 2012.

15 Ministerie van Volksgezondheid, Welzijn en Sport en College voor zorgverzekeringen. Dutch assessment procedures for the reimbursement of outpatient medicines [in Dutch: Nederlandse beoordelingsprocedures voor de vergoeding van extramurale geneesmiddelen]. 2010.

16 Administrative outline agreement 2012-2015 between the Dutch Hospital Association, the Netherlands Federation of University Medical Centres, Independent Clinics Netherlands, Health Insurers Netherlands and the Ministry of Health, Welfare and Sport [in Dutch: Bestuurlijk hoofdlijnenakkoord 2012-2015 tussen de Nederlandse Vereniging van Ziekenhuizen, de Nederlandse Federatie van Universitair Medisch Centra, Zelfstandige Klinieken Nederland, Zorgverzekeraars Nederland en het Ministerie van Volksgezondheid, Welzijn en Sport]. 2011.

17 Ljungberg B, Campbell SC, Choi HY, Jacqmin D, Lee JE, Weikert S, Kiemeney LA. The epidemiology of renal cell carcinoma. Eur Urol 2011;60(4):615-621.

18 Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013;49(6):1374-1403.

19 Nederlandse Kankerregistratie, beheerd door het Integraal Kankercentrum Nederland. Dutch cancer figures [in Dutch: Cijfers over kanker]. Available at: http://www.cijfersoverkanker.nl. Accessed January 11, 2016.

20 Lam JS, Leppert JT, Belldegrun AS, Figlin RA. Novel approaches in the therapy of metastatic renal cell carcinoma. World J Urol 2005;23(3):202-212.

21 Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, Eigl BJ, Ruether JD, Cheng T, North S, Venner P, Knox JJ, Chi KN, Kollmannsberger C, McDermott DF, Oh WK, Atkins MB, Bukowski RM, Rini BI, Choueiri TK. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol 2009;27(34):5794-5799.

22 Heng DY, Xie W, Regan MM, Harshman LC, Bjarnason GA, Vaishampayan UN, Mackenzie M, Wood L, Donskov F, Tan MH, Rha SY, Agarwal N, Kollmannsberger C, Rini BI, Choueiri TK. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013;14(2):141-148.

23 Motzer RJ, Bacik J, Murphy BA, Russo P, Mazumdar M. Interferon-alfa as a comparative treatment for clinical trials of new therapies against advanced renal cell carcinoma. J Clin Oncol 2002;20(1):289-296.

24 Motzer RJ, Bander NH, Nanus DM. Renal-cell carcinoma. N Engl J Med 1996;335(12):865-875.

25 Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, Oudard S, Negrier S, Szczylik C, Kim ST, Chen I, Bycott PW, Baum CM, Figlin RA. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. N Engl J Med 2007;356(2):115-124.

26 Escudier B, Pluzanska A, Koralewski P, Ravaud A, Bracarda S, Szczylik C, Chevreau C, Filipek M, Melichar B, Bajetta E, Gorbunova V, Bay JO, Bodrogi I, Jagiello-Gruszfeld A, Moore N, AVOREN Trial investigators. Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. Lancet 2007;370(9605):2103-2111.

27 Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, Ou SS, Archer L, Atkins JN, Picus J, Czaykowski P, Dutcher J, Small EJ. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. J Clin Oncol 2008;26(33):5422-5428.

28 Sternberg CN, Hawkins RE, Wagstaff J, Salman P, Mardiak J, Barrios CH, Zarba JJ, Gladkov OA, Lee E, Szczylik C, McCann L, Rubin SD, Chen M, Davis ID. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. Eur J Cancer 2013;49(6):1287-1296.

29 Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, Staroslawska E, Sosman J, McDermott D, Bodrogi I, Kovacevic Z, Lesovoy V, Schmidt-Wolf IG, Barbarash O, Gokmen E, O'Toole T, Lustgarten S, Moore L, Motzer RJ, Global ARCC Trial. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med 2007;356(22):2271-2281.

30 Escudier B, Kataja V, ESMO Guidelines Working Group. Renal cell carcinoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2009;20 Suppl 4:81-82.

31 Escudier B, Kataja V, ESMO Guidelines Working Group. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2010;21 Suppl 5:v137-v139.

32 Vereniging van Integrale Kankercentra. Dutch guideline renal cell carcinoma, Version: 2.0 [in Dutch: Niercelcarcinoom Landelijke richtlijn, Versie: 2.0]. 2010.

33 Commissie Farmaceutische Hulp. Committee of Pharmaceutical Aid Report on everolimus (Afinitor®) [in Dutch: CFH-rapport everolimus (Afinitor®)]. 2010.

34 Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Oudard S, Negrier S, Szczylik C, Pili R, Bjarnason GA, Garcia-del-Muro X, Sosman JA, Solska E, Wilding G, Thompson JA, Kim ST, Chen I, Huang X, Figlin RA. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol 2009;27(22):3584-3590.

35 Gore ME, Szczylik C, Porta C, Bracarda S, Bjarnason GA, Oudard S, Lee SH, Haanen J, Castellano D, Vrdoljak E, Schoffski P, Mainwaring P, Hawkins RE, Crino L, Kim TM, Carteni G, Eberhardt WE, Zhang K, Fly K, Matczak E, Lechuga MJ, Hariharan S, Bukowski R. Final results from the large sunitinib global expanded-access trial in metastatic renal cell carcinoma. Br J Cancer 2015;113(1):12-19.

36 Escudier B, Eisen T, Porta C, Patard JJ, Khoo V, Algaba F, Mulders P, Kataja V, ESMO Guidelines Working Group. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2012;23 Suppl 7:vii65-71.

37 Sun M, Shariat SF, Cheng C, Ficarra V, Murai M, Oudard S, Pantuck AJ, Zigeuner R, Karakiewicz PI. Prognostic factors and predictive models in renal cell carcinoma: a contemporary review. Eur Urol 2011;60(4):644-661.

38 Motzer RJ, Escudier B, Bukowski R, Rini BI, Hutson TE, Barrios CH, Lin X, Fly K, Matczak E, Gore ME. Prognostic factors for survival in 1059 patients treated with sunitinib for metastatic renal cell carcinoma. Br J Cancer 2013;108(12):2470-2477.

39 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30(4):377-399.

40 Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, Negrier S, Chevreau C, Solska E, Desai AA, Rolland F, Demkow T, Hutson TE, Gore M, Freeman S, Schwartz B, Shan M, Simantov R, Bukowski RM, TARGET Study Group. Sorafenib in advanced clear-cell renal-cell carcinoma. N Engl J Med 2007;356(2):125-134.

41 Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Staehler M, Negrier S, Chevreau C, Desai AA, Rolland F, Demkow T, Hutson TE, Gore M, Anderson S, Hofilena G, Shan M, Pena C, Lathia C, Bukowski RM. Sorafenib for treatment of renal cell carcinoma: Final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. J Clin Oncol 2009;27(20):3312-3318.

42 Escudier B, Bellmunt J, Negrier S, Bajetta E, Melichar B, Bracarda S, Ravaud A, Golding S, Jethwa S, Sneller V. Phase III trial of bevacizumab plus interferon alfa-2a in patients with metastatic renal cell carcinoma (AVOREN): final analysis of overall survival. J Clin Oncol 2010;28(13):2144-2150.

43 Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, Archer L, Atkins JN, Picus J, Czaykowski P, Dutcher J, Small EJ. Phase III trial of bevacizumab plus interferon alfa versus interferon alfa monotherapy in patients with metastatic renal cell carcinoma: final results of CALGB 90206. J Clin Oncol 2010;28(13):2137-2143.

44 Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, Grunwald V, Thompson JA, Figlin RA, Hollaender N, Urbanowitz G, Berg WJ, Kay A, Lebwohl D, Ravaud A, RECORD-1 Study Group. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial.

Lancet 2008;372(9637):449-456.

45 Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, Grunwald V, Thompson JA, Figlin RA, Hollaender N, Kay A, Ravaud A, RECORD-1 Study Group. Phase 3 trial of everolimus for metastatic renal cell carcinoma : final results and analysis of prognostic factors. Cancer 2010;116(18):4256-4265.

46 Sternberg CN, Davis ID, Mardiak J, Szczylik C, Lee E, Wagstaff J, Barrios CH, Salman P, Gladkov OA, Kavina A, Zarba JJ, Chen M, McCann L, Pandite L, Roychowdhury DF, Hawkins RE. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. J Clin Oncol 2010;28(6):1061-1068.

47 Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, Hutson TE, Michaelson MD, Gorbunova VA, Gore ME, Rusakov IG, Negrier S, Ou YC, Castellano D, Lim HY, Uemura H, Tarazi J, Cella D, Chen C, Rosbrook B, Kim S, Motzer RJ. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. Lancet 2011;378(9807):1931-1939.

48 Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier S, Oudard S, Gore ME, Tarazi J, Hariharan S, Chen C, Rosbrook B, Kim S, Rini BI. Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. Lancet Oncol 2013;14(6):552-562.

49 Day D, Kanjanapan Y, Kwan E, Yip D, Lawrentschuk N, Andrews M, Davis ID, Azad AA, Rosenthal M, Wong S, Johnstone A, Gibbs P, Tran B. Patterns of care for metastatic renal cell carcinoma in Australia. BJU Int 2015;116 Suppl 3:36-41.

50 Prescribing and Primary Care team, Health and Social Care Information Centre. Use of NICE appraised medicines in the NHS in England – 2012, experimental statistics. 2014.

51 Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. Cancer Treat Rev 2008;34(3):193-205.

52 Medical Research Council Renal Cancer Collaborators. Interferon-alpha and survival in metastatic renal carcinoma: early results of a randomised controlled trial. Lancet 1999;353(9146):14-17.

53 Flanigan RC, Salmon SE, Blumenstein BA, Bearman SI, Roy V, McGrath PC, Caton JR,Jr, Munshi N, Crawford ED. Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer. N Engl J Med 2001;345(23):1655-1659.

54 Mickisch GH, Garin A, van Poppel H, de Prijck L, Sylvester R, European Organisation for Research and Treatment of Cancer (EORTC) Genitourinary Group. Radical nephrectomy plus interferon-alfa-based immunotherapy compared with interferon alfa alone in metastatic renal-cell carcinoma: a randomised trial. Lancet 2001;358(9286):966-970.

55 Escudier B, Porta C, Schmidinger M, Algaba F, Patard JJ, Khoo V, Eisen T, Horwich A, ESMO Guidelines Working Group. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014;25 Suppl 3:iii49-56.

56 Ljungberg B, Bensalah K, Canfield S, Dabestani S, Hofmann F, Hora M, Kuczyk MA, Lam T, Marconi L, Merseburger AS, Mulders P, Powles T, Staehler M, Volpe A, Bex A. EAU guidelines on renal cell carcinoma: 2014 update. Eur Urol 2015;67(5):913-924.

57 Heng DY, Wells JC, Rini BI, Beuselinck B, Lee JL, Knox JJ, Bjarnason GA, Pal SK, Kollmannsberger CK, Yuasa T, Srinivas S, Donskov F, Bamias A, Wood LA, Ernst DS, Agarwal N, Vaishampayan UN, Rha SY, Kim JJ, Choueiri TK. Cytoreductive nephrectomy in patients with synchronous metastases from renal cell carcinoma: results from the International Metastatic Renal Cell Carcinoma Database Consortium. Eur Urol 2014;66(4):704-710.

58 Choueiri TK, Xie W, Kollmannsberger C, North S, Knox JJ, Lampard JG, McDermott DF, Rini BI, Heng DY. The impact of cytoreductive nephrectomy on survival of patients with metastatic renal cell carcinoma receiving vascular endothelial growth factor targeted therapy. J Urol 2011;185(1):60-66.

59 Bamias A, Tzannis K, Papatsoris A, Oudard S, Beuselinck B, Escudier B, Liontos M, Elaidi TR, Chrisofos M, Stravodimos K, Anastasiou I, Mitropoulos D, Deliveliotis C, Constantinides C, Dimopoulos MA, Bamia C. Prognostic significance of cytoreductive nephrectomy in patients with synchronous metastases from renal cell carcinoma treated with first-line sunitinib: a European multiinstitutional study. Clin Genitourin Cancer 2014;12(5):373-383.

60 Warren M, Venner PM, North S, Cheng T, Venner C, Ghosh S, Venner AA, Finch D. A population-based study examining the effect of tyrosine kinase inhibitors on survival in metastatic renal cell carcinoma in Alberta and the role of nephrectomy prior to treatment. Can Urol Assoc J 2009;3(4):281-289.

61 Stuart EA. Matching methods for causal inference: A review and a look forward. Stat Sci 2010;25(1):1-21.

62 Becker SO, Ichino A. Estimation of average treatment effects based on propensity scores. The Stata Journal 2002;2(4):358-377.

63 Mitra R, Reiter JP. A comparison of two methods of estimating propensity scores after multiple imputation. Stat Methods Med Res 2016;25(1):188-204.

64 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res 2011;46(3):399-424.

65 Goossens LMA. Underestimated uncertainties, Hospital-at-home for COPD exacerbations and methodological issues in the economic evaluation of healthcare. Erasmus University Rotterdam 2013.

66 Jonasch E, Wood CG, Matin SF, Tu SM, Pagliaro LC, Corn PG, Aparicio A, Tamboli P, Millikan RE, Wang X, Araujo JC, Arap W, Tannir N. Phase II presurgical feasibility study of bevacizumab in untreated patients with metastatic renal cell carcinoma. J Clin Oncol 2009;27(25):4076-4081.

67 Cowey CL, Amin C, Pruthi RS, Wallen EM, Nielsen ME, Grigson G, Watkins C, Nance KV, Crane J, Jalkut M, Moore DT, Kim WY, Godley PA, Whang YE, Fielding JR, Rathmell WK. Neoadjuvant clinical trial with sorafenib for patients with stage II or higher renal cell carcinoma. J Clin Oncol 2010;28(9):1502-1507.

68 Hellenthal NJ, Underwood W, Penetrante R, Litwin A, Zhang S, Wilding GE, Teh BT, Kim HL. Prospective clinical trial of preoperative sunitinib in patients with renal cell carcinoma. J Urol 2010;184(3):859-864.

69 Bex A, Blank C, Meinhardt W, van Tinteren H, Horenblas S, Haanen J. A phase II study of presurgical sunitinib in patients with metastatic clear-cell renal carcinoma and the primary tumor in situ. Urology 2011;78(4):832-837.

70 Powles T, Blank C, Chowdhury S, Horenblas S, Peters J, Shamash J, Sarwar N, Boleti E, Sahdev A, O'Brien T, Berney D, Beltran L, Nathan P, Haanen J, Bex A. The outcome of patients treated with sunitinib prior to planned nephrectomy in metastatic clear cell renal cancer. Eur Urol 2011;60(3):448-454.

71 Flanigan RC, Mickisch G, Sylvester R, Tangen C, van Poppel H, Crawford ED. Cytoreductive nephrectomy in patients with metastatic renal cancer: a combined analysis. J Urol 2004;171(3):1071-1076.

72 U.S. National Institutes of Health. Clinical Trial to Assess the Importance of Nephrectomy (CARMENA). Available at: https://clinicaltrials.gov. Last updated August 7, 2014.

73 U.S. National Institutes of Health. Immediate Surgery or Surgery After Sunitinib Malate in Treating Patients With Metastatic Kidney Cancer (SURTIME). Available at: https://clinicaltrials.gov. Last updated February 18, 2015.

74 Lambea J, Hinojo C, Lainez N, Lazaro M, Leon L, Rodriguez A, Soto de Prado D, Esteban E. Quality of life and supportive care for patients with metastatic renal cell carcinoma. Cancer Metastasis Rev 2012;31 Suppl 1:S33-S39.

75 Hotte SJ, Bjarnason GA, Heng DY, Jewett MA, Kapoor A, Kollmannsberger C, Maroun J, Mayhew LA, North S, Reaume MN, Ruether JD, Soulieres D, Venner PM, Winquist EW, Wood L, Yong JH, Saad F. Progression-free survival as a clinical trial endpoint in advanced renal cell carcinoma. Curr Oncol 2011;18 Suppl 2:S11-S19.

76 Szende A, Oppe M, Devlin N, eds. EQ-5D Value Sets: Inventory, Comparative Review and User Guide. 1st ed. Dordrecht, the Netherlands: Springer Netherlands 2007.

77 Zorginstituut Nederland. Guidelines for conducting economic evaluations in healthcare [in Dutch: Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg]. 2016.

78 National Institute for Care and Health Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013.

79 de Groot S, van der Linden N, Franken MG, Blommestein HM, Leeneman B, van Rooijen E, van der Hoeven JJM, Wouters MW, Westgeest HM, Uyl-de Groot CA. Balancing the optimal and the feasible: A practical guide for setting up patient registries for the collection of real-world data for health care decision making based on Dutch experiences. Value Health 2016; In press.

80 Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365-376.

81 Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd ed.). 2001.

82 Versteegh MM, Vermeulen KM, Evers SM, de Wit GA, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health 2016;19(4):343-352.

83 Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA 1995;273(1):59-65.

84 Cella D, Michaelson MD, Bushmakin AG, Cappelleri JC, Charbonneau C, Kim ST, Li JZ, Motzer RJ. Health-related quality of life in patients with metastatic renal cell carcinoma treated with sunitinib vs interferon-alpha in a phase III trial: final results and geographical analysis. Br J Cancer 2010;102(4):658-664.

85 Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, Stein K. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. Health Technol Assess 2010;14(2):1-184.

86 Remak E, Charbonneau C, Negrier S, Kim ST, Motzer RJ. Economic evaluation of sunitinib malate for the first-line treatment of metastatic renal cell carcinoma. J Clin Oncol 2008;26(24):3995-4000.

87 Motzer RJ, Michaelson MD, Redman BG, Hudes GR, Wilding G, Figlin RA, Ginsberg MS, Kim ST, Baum CM, DePrimo SE, Li JZ, Bello CL, Theuer CP, George DJ, Rini BI. Activity of SU11248, a multitargeted inhibitor of vascular endothelial growth factor receptor and platelet-derived growth factor receptor, in patients with metastatic renal cell carcinoma. J Clin Oncol 2006;24(1):16-24.

88 Motzer RJ, Hutson TE, Cella D, Reeves J, Hawkins R, Guo J, Nathan P, Staehler M, de Souza P, Merchan JR, Boleti E, Fife K, Jin J, Jones R, Uemura H, De Giorgi U, Harmenberg U, Wang J, Sternberg CN, Deen K, McCann L, Hackshaw MD, Crescenzo R, Pandite LN, Choueiri TK. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N Engl J Med 2013;369(8):722-731.

89 Sleijfer S, Verweij J. The price of success: cost-effectiveness of molecularly targeted agents. Clin Pharmacol Ther 2009;85(2):136-138.

90 Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. Eur J Cancer 2010;46(4):765-781.

91 De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, Trama A, Visser O, Brenner H, Ardanaz E, Bielska-Lasota M, Engholm G, Nennecke A, Siesling S, Berrino F, Capocaccia R, EUROCARE-5 Working Group. Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE--5-a population-based study. Lancet Oncol 2014;15(1):23-34.

92 de Groot S, Sleijfer S, Redekop WK, Oosterwijk E, Haanen JB, Kiemeney LA, Uyl-de Groot CA. Variation in use of targeted therapies for metastatic renal cell carcinoma: Results from a Dutch population-based registry. BMC Cancer 2016;16:364.

93 Centraal Bureau voor de Statistiek. Life expectancy; gender and age, from 1950 (per year) [in Dutch: Levensverwachting; geslacht en leeftijd, vanaf 1950 (per jaar)]. Available at: http://statline.cbs.nl. Accessed May 6, 2014.

94 Cella D, Pickard AS, Duh MS, Guerin A, Mishagina N, Antras L, Neary MP, McCann L, Hodge R, Sternberg CN. Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomised phase III trial. Eur J Cancer 2012;48(3):311-323.

95 Yang S, de Souza P, Alemao E, Purvis J. Quality of life in patients with advanced renal cell carcinoma treated with temsirolimus or interferon-alpha. Br J Cancer 2010;102(10):1456-1460.

96. Cella D, Escudier B, Rini B, Chen C, Bhattacharyya H, Tarazi J, Rosbrook B, Kim S, Motzer R. Patient-reported outcomes for axitinib vs sorafenib in metastatic renal cell carcinoma: phase III (AXIS) trial. Br J Cancer 2013;108(8):1571-1578.

97 Centraal Bureau voor de Statistiek. Consumer prices; price index 2006 = 100 [in Dutch: Consumentenprijzen; prijsindex 2006 = 100]. Available at: http://statline.cbs.nl. Accessed May 6, 2014.

98 College voor zorgverzekeringen. Guidelines for pharmacoeconomic research [in Dutch: Richtlijnen voor farmaco-economisch onderzoek]. 2006.

99 Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, ISPOR-SMDM Modeling Good Research Practices Task Force. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health 2012;15(6):843-850.

100 Hoyle M, Green C, Thompson-Coon J, Liu Z, Welch K, Moxham T, Stein K. Cost-effectiveness of temsirolimus for first line treatment of advanced renal cell carcinoma. Value Health 2010;13(1):61-68.

101 Hoyle M, Green C, Thompson-Coon J, Liu Z, Welch K, Moxham T, Stein K. Cost-effectiveness of sorafenib for second-line treatment of advanced renal cell carcinoma. Value Health 2010;13(1):55-60.

102 Benedict A, Figlin RA, Sandstrom P, Harmenberg U, Ullen A, Charbonneau C, Sandin R, Remak E, Hariharan S, Negrier S. Economic evaluation of new targeted therapies for the first-line treatment of patients with metastatic renal cell carcinoma. BJU Int 2011;108(5):665-672.

103 Casciano R, Chulikavit M, Di Lorenzo G, Liu Z, Baladi JF, Wang X, Robertson J, Garrison L. Economic evaluation of everolimus versus sorafenib for the treatment of metastatic renal cell carcinoma after failure of first-line sunitinib. Value Health 2011;14(6):846-851.

104 Lam JS, Leppert JT, Figlin RA, Belldegrun AS. Role of molecular markers in the diagnosis and therapy of renal cell carcinoma. Urology 2005;66 Suppl 5:1-9.

105 Dutch reimbursement prices (Z-index) [in Dutch: Officiële vergoedingsprijzen (Z-index)]. 2015.

106 Tan SS, Bouwmans CA, Rutten FF, Hakkaart-van Roijen L. Update of the Dutch Manual for Costing in Economic Evaluations. Int J Technol Assess Health Care 2012;28(2):152-158.

107 Nederlandse Zorgautoriteit. Tariffs as issued by the Dutch Healthcare Authority [in Dutch: DBC zorgproducten tariefapplicatie]. Available at: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief. Last updated January 2015.

108 Gliklich RE, Dreyer NA, eds. Registries for Evaluating Patient Outcomes: A User's Guide. 2nd ed. Rockville, MD: Agency for Healthcare Research and Quality 2010.

109 Lis Y, Roberts MH, Kamble S, J Guo J, Raisch DW. Comparisons of Food and Drug Administration and European Medicines Agency risk management implementation for recent pharmaceutical approvals: report of the International Society for Pharmacoeconomics and outcomes research risk benefit management working group. Value Health 2012;15(8):1108-1118.

110 College voor zorgverzekeringen. Guidance for outcomes research [in Dutch: Leidraad voor Uitkomstenonderzoek 'ten behoeve van de beoordeling doelmatigheid intramurale geneesmiddelen']. 2008.

111 Baxter SL, Wormald RP, Musa JM, Patel D. Blindness Registers as Epidemiological Tools for Public Health Planning: A Case Study in Belize. Epidemiology Research International 2014.

112 Blommestein HM, Issa DE, Pompen M, Ten Hoor G, Hogendoorn M, Joosten P, Zweegman S, Huijgens PC, Uyl-de Groot CA. Cost-effectiveness of rituximab as maintenance treatment for relapsed follicular lymphoma: results of a population-based study. Eur J Haematol 2014;92(5):398-406.

113 Oerlemans S, Issa DE, van den Broek EC, Nijziel MR, Coebergh JW, Mols F, van de Poll-Franse LV. Impact of therapy and disease-related symptoms on health-related quality of life in patients with follicular lymphoma: results of the population-based PHAROS-registry. Eur J Haematol 2014;93(3):229-238.

114 Oerlemans S, Issa DE, van den Broek EC, Nijziel MR, Coebergh JW, Huijgens PC, Mols F, van de Poll-Franse LV. Health-related quality of life and persistent symptoms in relation to (R-)CHOP14, (R-)CHOP21, and other therapies among patients with diffuse large B-cell lymphoma: results of the population-based PHAROS-registry. Ann Hematol 2014;93(10):1705-1715.

115 van den Broek EC, Oerlemans S, Nijziel MR, Posthuma EF, Coebergh JW, van de Poll-Franse LV. Impact of active surveillance, chlorambucil, and other therapy on health-related quality of life in patients with CLL/SLL in the Netherlands. Ann Hematol 2015;94(1):45-56.

116 Blommestein HM, Verelst SG, de Groot S, Huijgens PC, Sonneveld P, Uyl-de Groot CA. A cost-effectiveness analysis of real-world treatment for elderly patients with multiple myeloma using a full disease model. Eur J Haematol 2016;96(2):198-208.

117 van der Linden N, van Gils CW, Pescott CP, Buter J, Uyl-de Groot CA. Cetuximab in locally advanced squamous cell carcinoma of the head and neck: generalizability of EMR 062202-006 trial results. Eur Arch Otorhinolaryngol 2014;271(6):1673-1678.

118 van der Linden N, van Gils CW, Pescott CP, Buter J, Vergeer MR, Groot CA. Real-world cost-effectiveness of cetuximab in locally advanced squamous cell carcinoma of the head and neck. Eur Arch Otorhinolaryngol 2015;272(8):2007-2016.

119 van der Linden N, Buter J, Pescott CP, Lalisang RI, de Boer JP, de Graeff A, van Herpen CM, Baatenburg de Jong RJ, Uyl-de Groot CA. Treatments and costs for recurrent and/or metastatic squamous cell carcinoma of the head and neck in the Netherlands. Eur Arch Otorhinolaryngol 2016;273(2):455-464.

120 van de Poll-Franse LV, Horevoorts N, van Eenbergen M, Denollet J, Roukema JA, Aaronson NK, Vingerhoets A, Coebergh JW, de Vries J, Essink-Bot ML, Mols F, The Profiles Registry Group. The Patient Reported Outcomes Following Initial treatment and Long term Evaluation of Survivorship registry: scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. Eur J Cancer 2011;47(14):2188-2194.

121 Bellgard MI, Macgregor A, Janon F, Harvey A, O'Leary P, Hunter A, Dawkins H. A modular approach to disease registry design: successful adoption of an internet-based rare disease registry. Human Mutation 2012;33(10):E2356-2366.

122 Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol 2013;42(4):1012-1014.

123 Saczynski JS, McManus DD, Goldberg RJ. Commonly used data-collection approaches in clinical research. Am J Med 2013;126(11):946-950.

124 Wattigney WA, Croft JB, Mensah GA, Alberts MJ, Shephard TJ, Gorelick PB, Nilasena DS, Hess DC, Walker MD, Hanley DF Jr, Shwayder P, Girgus M, Neff LJ, Williams JE, LaBarthe DR, Collins JL. Establishing data elements for the Paul Coverdell National Acute Stroke Registry: Part 1: proceedings of an expert panel. Stroke 2003;34(1):151-156.

125 Willis CD, McNeil JJ, Cameron PA, Phillips LE. Monitoring drug safety with registries: useful components of postmarketing pharmacovigilance systems. Journal of Clinical Epidemiology 2012;65(2):121-125.

126 European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP), Module VI - Management and reporting of adverse reactions to medicinal products (Rev 1). 2014.

127 Molsen E, Trotter J, Smith MD. Use of patient registries: results of the ISPOR patient registry special interest group survey. ISPOR Connections 2005;11(6).

128 Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. British Journal of General Practice 2010;60(572):e128-136.

129 Reeves MJ, Nickles AV, Roberts S, Hurst R, Lyon-Callo S. Assessment of the completeness and accuracy of case ascertainment in the Michigan Stroke Registry. Circulation: Cardiovascular Quality and Outcomes 2014;7(5):757-763.

130 Korngut L, Johnston M, Pringsheim T, Jetté N. The future of neurological patient registries. Clinical Practice 2014;11(5):509-516.

131 Arthursson AJ, Furnes O, Espehaug B, Havelin LI, Söreide JA. Validation of data in the Norwegian Arthroplasty Register and the Norwegian Patient Register: 5,134 primary total hip arthroplasties and revisions operated at a single hospital between 1987 and 2003. Acta Orthop 2005;76(6):823-828.

132 van Steenbergen LN, Denissen GAW, Spooren A, van Rooden SM, van Oosterhout FJ, Morrenhof JW, Nelissen RGHH. More than 95% completeness of reported procedures in the population-based Dutch Arthroplasty Register: External validation of 311,890 procedures. Acta orthopaedica 2015;86(4):498-505.

133 Kearney TM, Donnelly C, Kelly JM, O'Callaghan EP, Fox CR, Gavin AT. Validation of the completeness and accuracy of the Northern Ireland Cancer Registry. Cancer epidemiology 2015;39(3):401-404.

134 Londero SC, Mathiesen JS, Krogdahl A, Bastholt L, Overgaard J, Bentsen J, Hahn CH, Schytte S, Pedersen HB, Christiansen P, Godballe C. Completeness and validity in a national clinical thyroid cancer database: DATHYRCA. Cancer epidemiology 2014;38(5):633-637.

135 Engelberg Center for Health Care Reform at Brookings. How registries can help performance measurement improve care. 2010.

136 PARENT, Cross-border PAtient REgistries iNiTiative. Available at: http://patientregistries.eu. Accessed January 4, 2014.

137 Richesson RL. Data standards in diabetes patient registries. Journal of Diabetes Science and Technology 2011;5(3):476-485.

138 Catarinella F, Stavast I, Wittens C. A European venous registry: pitfalls and opportunities. Phlebology 2014;29 Suppl 1:188-192.

139 van der Velde ET, Vriend JWJ, Mannens MMAM, Uiterwaal CSPM, Brand R, Mulder BJM. CONCOR, an initiative towards a national registry and DNA-bank of patients with congenital heart disease in the Netherlands: rationale, design, and first results. European Journal of Epidemiology 2005;20(6):549-557.

140 Tosh J, Stevenson M, Akehurst R. Health economic modelling of treatment sequences for rheumatoid arthritis: a systematic review. Curr Rheumatol Rep 2014;16(10):447.

141 Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices-overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Value Health 2012;15(6):796-803.

142 de Groot S, Blommestein HM, Redekop WK, Sleijfer S, Kiemeney LA, Oosterwijk E, Uyl-de Groot CA. Potential health gains for patients with metastatic renal cell carcinoma in daily clinical practice: A real-world cost-effectiveness analysis of sequential first- and second-line treatments. Submitted.

143 Kyle RA. Prognostic factors in multiple myeloma. Stem Cells 1995;13 Suppl 2:56-63.

144 Ludwig H, Bolejack V, Crowley J, Blade J, Miguel JS, Kyle RA, Rajkumar SV, Shimizu K, Turesson I, Westin J, Sonneveld P, Cavo M, Boccadoro M, Palumbo A, Tosi P, Harousseau JL, Attal M, Barlogie B, Stewart AK, Durie B. Survival and years of life lost in different age cohorts of patients with multiple myeloma. J Clin Oncol 2010;28(9):1599-1605.

145 Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J, Boccadoro M, Child JA, Avet-Loiseau H, Kyle RA, Lahuerta JJ, Ludwig H, Morgan G, Powles R, Shimizu K, Shustik C, Sonneveld P, Tosi P, Turesson I, Westin J. International staging system for multiple myeloma. J Clin Oncol 2005;23(15):3412-3420.

146 San Miguel JF, Garcia-Sanz R. Prognostic features of multiple myeloma. Best Pract Res Clin Haematol 2005;18(4):569-583.

147 Fonseca R, Stewart AK. Targeted therapeutics for multiple myeloma: the arrival of a risk-stratified approach. Molecular Cancer Therapeutics 2007;6:802-810.

148 Huijgens PC, Posthuma EFM, Coebergh JWW, van de Poll-Franse LV, Uyl-de Groot CA, Sonneveld P. A 'population based registry' for hemato-oncology. Nederlands Tijdschrift voor Hematologie 2010;7:321-325.

149 Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health 2012;15(6):812-820.

150 Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Value Health 2012;15(6):821-827.

151 Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. Pharmacoeconomics 2008;26(2):131-148.

152 Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011.

153 Leunis A, Redekop WK, van Montfort KA, Lowenberg B, Uyl-de Groot CA. The development and validation of a decision-analytic model representing the full disease course of acute myeloid leukemia. Pharmacoeconomics 2013;31(7):605-621.

154 Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text. 2nd ed. New York: Springer 2005.

155 Davis S, Stevenson M, Tappenden P, Wailoo A. NICE DSU Technical Support Document 15: Cost-effectiveness modelling using patient-level simulation. 2014.

156 Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, Cook J, Glick H, Liljas B, Petitti D, Reed S. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health 2005;8(5):521-533.

157 Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacoot N, Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess 2004;8(36):1-158.

158 Heeg BM, Damen J, Buskens E, Caleo S, de Charro F, van Hout BA. Modelling approaches: the case of schizophrenia. Pharmacoeconomics 2008;26(8):633-648.

159 Commissie Farmaceutische Hulp. Committee of Pharmaceutical Aid Report on bevacizumab (Avastin®) [in Dutch: CFH-rapport bevacizumab (Avastin®)]. 2008.

160 Commissie Farmaceutische Hulp. Committee of Pharmaceutical Aid Report on temsirolimus (Torisel®) [in Dutch: CFH-rapport temsirolimus (Torisel®)]. 2008.

161 Commissie Farmaceutische Hulp. Committee of Pharmaceutical Aid Report on sunitinib (Sutent®) [in Dutch: CFH-rapport sunitinib (Sutent®)]. 2007.

162 Commissie Farmaceutische Hulp. Committee of Pharmaceutical Aid Report on pazopanib (Votrient®) [in Dutch: CFH-rapport pazopanib (Votrient®)]. 2010.

163 National Institute for Health and Clinical Excellence (NICE). Bevacizumab (first-line), sorafenib (first- and second-line), sunitinib (second-line) and temsirolimus (first-line) for the treatment of advanced and/or metastatic renal cell carcinoma. 2009.

164 National Institute for Health and Clinical Excellence (NICE). Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. 2009.

165 Jonsson B. Being NICE is not the problem! Eur J Cancer 2009;45(7):1100-1102.

166 Raad voor de Volksgezondheid & Zorg. Justified and sustainable care [in Dutch: Rechtvaardige en duurzame zorg]. 2006.

167 Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration 2011.

168 Faria R, Hernandez Alava M, Manca A, Wailoo A. NICE DSU Technical support document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data. 2015.

169 Gotink KJ, Verheul HM. Anti-angiogenic tyrosine kinase inhibitors: what is their mechanism of action? Angiogenesis 2010;13(1):1-14.

170 Lawerman-van de Wetering EJ. Valuing and Refining Outcome Measures for Economic Evaluations in Health Care. Erasmus University Rotterdam 2015.

171 Boggild M, Palace J, Barton P, Ben-Shlomo Y, Bregenzer T, Dobson C, Gray R. Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator. BMJ 2009;339:b4677.

172 Duddy M, Palace J. The UK Risk-Sharing Scheme for interferon-beta and glatiramer acetate in multiple sclerosis. Outcome of the year-6 analysis. Pract Neurol 2016;16(1):4-6.

173 Palace J, Duddy M, Bregenzer T, Lawton M, Zhu F, Boggild M, Piske B, Robertson NP, Oger J, Tremlett H, Tilling K, Ben-Shlomo Y, Dobson C. Effectiveness and cost-effectiveness of interferon beta and glatiramer acetate in the UK Multiple Sclerosis Risk Sharing Scheme at 6 years: a clinical cohort study with natural history comparator. Lancet Neurol 2015;14(5):497-505.

174 Zorginstituut Nederland. Pharmaco-Economic report on nivolumab (Opdivo®) for treatment of patients with squamous-cell NSCLC [in Dutch: Farmaco-Economisch rapport voor nivolumab (Opdivo®) bij de behandeling van plaveiselcel NSCLC]. 2015.

175 McKenna C, Soares M, Claxton K, Bojke L, Griffin S, Palmer S, Spackman E. Unifying Research and Reimbursement Decisions: Case Studies Demonstrating the Sequence of Assessment and Judgments Required. Value Health 2015;18(6):865-875.

Summary

Summary

The growing pressure on healthcare budgets creates great tension between financial sustainability of healthcare systems and accessibility to (new) treatments. Healthcare authorities need to make choices between reimbursement of healthcare procedures, services and programs, but these choices have to be made under considerable uncertainty about (long-term) costs and effects. To ensure equal and undelayed access to promising but expensive inpatient drugs, conditional funding was introduced in the Netherlands in 2006. At that time, conditional funding implied the additional funding of innovative drugs (i.e. reimbursing hospitals [most of] the drug costs) for a period of three years on the condition that data regarding uptake, use and outcomes of these drugs in clinical practice were to be collected. These data were used to decide whether or not additional funding continued to exist. Two drugs for metastatic renal cell carcinoma (mRCC), that is, bevacizumab and temsirolimus, were eligible for conditional funding. The collection and analysis of data about uptake, use, and outcomes of these new drugs in clinical practice form the focus of the present thesis. This thesis also illustrates the possibilities and impossibilities of evaluating these outcomes. Additionally, it reflects on the difficulties of making decisions about future funding (based on evidence from clinical practice).

To evaluate the uptake and use of the new drugs for mRCC in Dutch clinical practice, as well as treatment outcomes (in terms of effectiveness and cost-effectiveness), a population-based registry was created (i.e. PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry). Since the registry was disease-oriented (i.e. all patients diagnosed with [m]RCC were eligible for inclusion, regardless of the type of treatment they received), it not only provided data on patients treated with bevacizumab or temsirolimus, but also on patients treated with other so-called targeted therapies.

Chapter 2 of this thesis shows that targeted therapies, sunitinib in particular, have largely replaced interferon-alfa (IFN-α) as first-line standard of care. Furthermore, the results from the PERCEPTION registry revealed that few patients were treated with bevacizumab (combined with IFN-α) or temsirolimus, even though these therapies were added to European guidelines in 2009, and to Dutch guidelines in 2010. Chapter 2 also shows that approximately one-third of treatment-eligible patients did not receive targeted therapies despite their ability to improve survival. Patients aged 65+ years were less likely to receive targeted therapy; other factors like number of metastatic sites were of influence in some subgroups.

The new targeted therapies for mRCC have been shown to increase progression-free survival (PFS) and overall survival (OS). However, it is unknown whether or not cytoreductive nephrectomy (CN) (i.e. CN aims to remove as much of the tumour as possible) further prolongs the survival of patients with primary mRCC treated with sunitinib (or another targeted agent), given the lack of evidence from randomised controlled trials (RCTs). This has led to a large variation in clinical practice, reflected by data from the PERCEPTION registry showing that 33% of the patients presenting with mRCC underwent CN prior to sunitinib. In **chapter 3** it is suggested that CN may be effective in this patient group, based on data from the PERCEPTION registry and propensity score matching. However, the benefit appeared to be modest after correction for the time from diagnosis to the start of treatment with sunitinib.

As previously indicated, the new targeted therapies for mRCC have been shown to increase PFS. It is assumed that one of the associated benefits is a delay in health-related quality of life (HRQoL) deterioration, as a result of a delay in progression of disease. However, little data are available supporting this relationship. **Chapter 4** shows that disease symptoms, such as fatigue and pain, are the key drivers for reduced HRQoL in mRCC (measured by a cancer-specific questionnaire [EORTC QLQ-C30] and a generic questionnaire [EQ-5D-5L]). Since symptoms increase with progression of disease (as shown in chapter 4), targeted therapies that increase PFS can postpone reductions in HRQoL.

Despite the reimbursement of the targeted therapies for mRCC, little is known about their cost-effectiveness in clinical practice. A cost-effectiveness analysis identifies the ratio between the costs of an intervention and its health benefits. In **chapter 5** it is shown that current treatment practice, which is dominated by sunitinib, led to an increase of 0.230 quality-adjusted life years (QALYs) compared to a scenario in which none of the patients receives a targeted therapy. The incremental costs are €24,179, resulting in an incremental cost-effectiveness ratio (ICER) of €105,011 per QALY gained. As discussed in chapter 5, health can be gained (in comparison with current treatment practice) if more treatment-eligible patients with mRCC receive targeted therapies (like sunitinib). However, the ICER is far beyond the upper limit of the cost-effectiveness threshold of €80,000 per QALY, as proposed by the Dutch Council for Public Health.

The PERCEPTION renal cell cancer registry provided information about uptake, use and outcomes of the new drugs for mRCC in clinical practice. It also provided evidence on the effect of CN prior to treatment with sunitinib. Chapter 5 and 6 provide recommendations for future research. **Chapter 6** provides practical guidance in setting up patient registries to facilitate re-

al-world data collection for health care decision-making. In this chapter it is argued that it is crucial to set up an efficient registry that serves its aims, by collecting the right data of the right patient in the right way. **Chapter 7** provides practical recommendations to conduct cost-effectiveness analyses of treatment strategies spanning multiple treatment lines, using data from clinical practice and a discrete event simulation model. As more treatments become available (within and beyond treatment lines), full disease models are expected to increase in importance.

Chapter 8 provides a discussion of the findings and explores the implications and limitations of this thesis. As illustrated with the PERCEPTION renal cell cancer registry, patient registries can provide important information about uptake and use of new therapies in clinical practice. However, data from registries are seldom able to assess the effectiveness and cost-effectiveness of a drug. First, it is often not only treatment that causes a difference in outcome, because physicians may have profound reasons to choose a particular treatment over another, and these reasons will often relate to a patient's prognosis. There are several ways to reduce confounding in observational studies, but all these methods assume no selection on unobserved confounders, while it is almost impossible to assume that all possible confounders are observed (and measured) and have been adjusted for. A second reason why patient registries are seldom able to assess the effectiveness and cost-effectiveness of a drug is the difficulty to include the minimum number of patients (in the intervention and the control group) needed to detect true and clinically important differences between treatments. The problem of insufficient patient numbers is expected to grow in the near future, since new therapies often target small patient subgroups.

Continuous registries might be able to provide data from patients treated with the standard or usual treatment (control group), data that are usually unavailable when a registry starts when a new drug has been introduced. The existence of continuous patient registries might therefore overcome some of the limitations of time-restricted registries, such as the PERCEPTION registry. Nationwide patient registries could increase patient numbers (in comparison with registries restricted to certain regions or hospitals). However, expanding registries in both length (i.e. from a time-restricted registry to a continuous registry) and width (i.e. from a registry restricted to certain regions or hospitals to a nationwide registry) will increase the time and costs of data collection. Collecting data on confounding variables (or instrumental variables) will further increase costs. It is therefore important to consider the overall costs of a registry in relation to its value.

Besides the challenges associated with the evaluation of a drug's effectiveness and cost-effectiveness in clinical practice, it appeared to be difficult to decide about future funding based on these outcomes (after a period of conditional funding). Only in some cases outcomes research (i.e. the collection and analysis of data from clinical practice) influenced the decision-making process in the Netherlands.

In line with the current policy on so-called specialist drugs (including expensive inpatient and outpatient drugs), the present thesis argues to limit the use of conditional funding and put more emphasis on initial reimbursement decisions. Results of cost-effectiveness analyses should be used to inform (initial) reimbursement decisions, in order to keep our healthcare system financially sustainable as the number of cancer patients increases and the number of innovative, expensive drugs grows. If cost-effectiveness ratios are above a given (equity weighted) threshold that reflects societal willingness to pay for health gains, price negotiations (or other measures) are essential to improve the relationship between costs and effects of a new drug, and allow reimbursement from the basic health insurance package. Only in specific cases (e.g. orphan drugs) is outcomes research able to provide a more robust estimate of cost-effectiveness. These cases should be selected carefully in order to minimize the costs of extensive data collection.

Samenvatting

Samenvatting

De toenemende druk op zorgbudgetten veroorzaakt een grote spanning tussen de betaalbaarheid van het zorgsysteem enerzijds en de toegankelijkheid tot (nieuwe) behandelingen anderzijds. Zorgautoriteiten moeten keuzen maken tussen de vergoeding van zorgprocedures, diensten en programma's, terwijl er op dat moment vaak onzekerheid is over de (lange termijn) kosten en effecten van deze interventies. Om gelijke en tijdige toegang tot veelbelovende maar kostbare intramurale medicijnen te garanderen, werd in Nederland in 2006 voorwaardelijke financiering geïntroduceerd. Voorwaardelijke financiering stond op dat moment voor de aanvullende financiering van innovatieve medicijnen (i.e. de vergoeding van ziekenhuizen voor [het merendeel van] de kosten van deze medicijnen) voor een periode van drie jaar. Gedurende deze periode dienden gegevens met betrekking tot de opname, het gebruik en de uitkomsten van deze medicijnen in de klinische praktijk te worden verzameld. Deze gegevens werden gebruikt om te bepalen of de aanvullende financiering werd gecontinueerd. Twee medicijnen voor gemetastaseerd niercelcarcinoom (metastatic renal cell carcinoma, mRCC), te weten bevacizumab en temsirolimus, kwamen in aanmerking voor deze voorwaardelijke financiering. Het verzamelen en analyseren van gegevens omtrent de opname, het gebruik en de uitkomsten van deze nieuwe medicijnen in de klinische praktijk vormen de focus van dit proefschrift. Dit proefschrift gaat eveneens in op de mogelijkheden en onmogelijkheden van het evalueren van deze uitkomsten. Ook reflecteert dit proefschrift op de moeilijkheden bij het maken van beslissingen over toekomstige financiering (op basis van bewijs uit de klinische praktijk).

Om de opname, het gebruik en de uitkomsten (in termen van effectiviteit en kosteneffectiviteit) van de nieuwe medicijnen voor mRCC in de klinische praktijk te evalueren, werd een patiëntenregister opgezet (i.e. PERCEPTION, PharmacoEconomics in Renal CEll carcinoma: a PopulaTION-based registry). Omdat dit register ziekte-georiënteerd was (i.e. alle patiënten met de diagnose [m]RCC kwamen in aanmerking voor inclusie, ongeacht het type behandeling dat zij kregen) leverde het niet alleen gegevens op over patienten die behandeld werden met bevacizumab of temsirolimus, maar ook over patiënten die behandeld werden met andere zogenaamde doelgerichte (targeted) therapieën.

Hoofdstuk 2 van dit proefschrift laat zien dat de traditionele behandeling met interferon-alfa (IFN-α) grotendeels vervangen is door behandeling met doelgerichte therapieën, voornamelijk sunitinib. Ook toonden de resultaten van het PERCEPTION register aan dat weinig patiënten werden behandeld met bevacizumab (gecombineerd met IFN-α) of temsirolimus, terwijl deze behandelingen worden aanbevolen in Europese richtlijnen sinds 2009, en in Nederlandse richtlijnen sinds 2010. Hoofdstuk 2 laat bovendien zien dat ongeveer een derde van de patiënten die in aanmerking kwamen voor een doelgerichte therapie niet behandeld werd met één van deze therapieën, terwijl deze behandelingen overleving kunnen verbeteren. Patiënten ouder dan 65 jaar kregen minder vaak een doelgerichte therapie; andere factoren, zoals het aantal locaties met metastasen, waren van invloed in sommige subgroepen.

De nieuwe doelgerichte therapieën voor mRCC hebben aangetoond progressie-vrije en totale overleving te verlengen. Het is echter niet bekend of een cytoreductieve nefrectomie (CN) (i.e. een chirurgische ingreep met als doel de omvang van de tumor te verminderen) de levensduur van patiënten met primair mRCC die behandeld worden met sunitinib (of een andere doelgerichte therapie) verder kan verlengen, vanwege het gebrek aan bewijs uit gerandomiseerde gecontroleerde trials (randomised controlled trials, RCTs). Dit heeft geleid tot grote behandelvariatie in de klinische praktijk. De gegevens uit het PERCEPTION register ondersteunen dit, en laten zien dat 33% van de patiënten met primair mRCC een CN onderging voorafgaand aan behandeling met sunitinib. **Hoofdstuk 3** toont aan dat CN mogelijk effectief is in deze patiëntengroep, gebaseerd op gegevens uit het PERCEPTION register en 'propensity score matching'. Het voordeel bleek echter bescheiden na correctie voor de tijd tussen diagnose en de start van de behandeling met sunitinib.

Zoals eerder gesteld verbeteren de nieuwe doelgerichte therapieën voor mRCC progressie-vrije overleving. Er wordt verondersteld dat dit een gunstig effect heeft op gezondheidsgerelateerde kwaliteit van leven (health-related quality of life, HRQoL), maar er zijn weinig gegevens beschikbaar die deze relatie ondersteunen. **Hoofdstuk 4** laat zien dat ziektesymptomen, zoals vermoeidheid en pijn, bepalende factoren zijn voor een achteruitgang in kwaliteit van leven van patiënten met mRCC (gemeten aan de hand van een kanker-specifieke vragenlijst [EORTC QLQ-C30] en een generieke vragenlijst [EQ-5D-5L]). Omdat symptomen toenemen bij progressie van ziekte (zoals aangetoond in hoofdstuk 4), kunnen doelgerichte therapieën die progressie-vrije overleving verlengen, achteruitgang in kwaliteit van leven uitstellen.

Ondanks de vergoeding van de doelgerichte therapieën voor mRCC, is er weinig bekend over de kosteneffectiviteit van deze behandelingen in de klinische praktijk. Een kosteneffectiviteitsanalyse laat zien hoe de kosten van een interventie zich verhouden tot de gezondheidsopbrengsten. **Hoofdstuk 5** toont aan dat de huidige inzet van behandelingen, die gedomineerd wordt door sunitinib, leidt tot een toename van voor kwaliteit-gecorrigeerde levensjaren (quality-adjusted life years, QALYs) met 0,230 in vergelijking met een scenario waarin patiënten niet behandeld worden met doelgerichte

therapieën. De additionele kosten zijn €24.179, resulterend in een incrementele kosteneffectiviteitsratio (incremental cost-effectiveness ratio, ICER) van €105.011 per gewonnen QALY. Zoals besproken in hoofdstuk 5 kan er meer gezondheid gewonnen worden (in vergelijking met de huidige inzet van behandelingen) wanneer meer patiënten met mRCC behandeld zouden worden met doelgerichte therapieën (zoals sunitinib). De ICER ligt echter ver boven de drempel voor kosteneffectiviteit van €80.000 per QALY, zoals voorgesteld door de Nederlandse Raad voor de Volksgezondheid & Zorg.

Het PERCEPTION register bood inzicht in de opname, het gebruik en de uitkomsten van nieuwe medicijnen voor mRCC in de klinische praktijk. Het register bood ook bewijs over het effect van CN voorafgaand aan de behandeling met sunitinib. Hoofdstuk 5 en 6 bieden handvatten voor toekomstig onderzoek. **Hoofdstuk 6** bevat praktische aanbevelingen voor het opzetten van patiëntenregisters voor het verzamelen van gegevens uit de klinische praktijk ter ondersteuning van besluitvorming in de gezondheidszorg. In dit hoofdstuk wordt betoogd dat het cruciaal is om een efficiënt register op te zetten dat de beoogde doelen dient, door de juiste gegevens te verzamelen van de juiste patiënt op de juiste manier. **Hoofdstuk 7** biedt praktische aanbevelingen voor het uitvoeren van kosteneffectiviteitsanalyses van behandelstrategieën bestaande uit één of meer behandellijnen, waarbij gebruik gemaakt wordt van gegevens uit de klinische praktijk en een zogenaamd 'discrete event simulation' model. Nu er steeds meer behandelingen beschikbaar komen (in verschillende behandellijnen) zullen volledige ziektemodellen in belang toenemen.

Hoofdstuk 8 bevat een discussie over de bevindingen en verkent de implicaties en beperkingen van dit onderzoek. Zoals aangetoond met het PERCEPTION register, kunnen patiëntenregisters belangrijke informatie bieden ten aanzien van de opname en het gebruik van nieuwe behandelingen in de klinische praktijk. Het is echter zelden mogelijk om de effectiviteit en kosteneffectiviteit van een medicijn vast te stellen aan de hand van gegevens uit registers. Ten eerste wordt een verschil in uitkomsten vaak niet alleen door de behandeling veroorzaakt. Artsen hebben veelal gegronde redenen om de voorkeur te geven aan een bepaalde behandeling, en deze redenen zullen vaak gerelateerd zijn aan de prognose van de patiënt. Er zijn diverse manieren om deze verstoring (confounding) in observationele studies te reduceren, maar al deze methoden gaan er vanuit dat selectie enkel plaatsvindt op basis van geobserveerde variabelen. Het is echter bijna onmogelijk te veronderstellen dat alle mogelijke verstorende variabelen (confounders) geobserveerd (en gemeten) zijn en dat daarvoor voldoende gecorrigeerd is. Een tweede complexiteit bij het vaststellen van de effectiviteit en kosteneffectiviteit van een

medicijn op basis van gegevens uit patiëntenregisters is de inclusie van het minimum benodigde aantal patiënten (in de interventie- en controlegroep) om ware en klinisch van belang zijnde verschillen tussen behandelingen aan te tonen. Dit probleem zal naar verwachting groter worden in de nabije toekomst, omdat nieuwe behandelingen zich steeds vaker richten op kleine subgroepen van patiënten.

Continue registers zijn mogelijk in staat om informatie te bieden over patiënten die behandeld werden met de standaard of gebruikelijke behandelmethode (controlegroep). Deze gegevens zijn normaal gesproken niet beschikbaar als een register start op het moment dat een nieuw medicijn wordt geïntroduceerd. Continue patiëntenregisters kunnen daarom sommige beperkingen van tijdgebonden registers, zoals het PERCEPTION register, voorkomen. Landelijke patiëntenregisters kunnen patiëntenaantallen verhogen (in vergelijking met registers die beperkt zijn tot bepaalde regio's of ziekenhuizen). Echter, het uitbreiden van registers in zowel lengte (i.e. van tijdgebonden registers naar continue registers) als breedte (i.e. van een register beperkt tot bepaalde regio's of ziekenhuizen naar een landelijk register) verhoogt de tijd en kosten die gepaard gaan met de gegevensverzameling. Het verzamelen van gegevens over verstorende variabelen (of instrumentele variabelen) zal de kosten nog verder doen toenemen. Het is daarom belangrijk om de totale kosten van een register te bezien in relatie tot de waarde van het register.

Naast de uitdagingen rond de evaluatie van de effectiviteit en kosteneffectiviteit van medicijnen in de klinische praktijk, blijkt het ook lastig om beslissingen te maken over toekomstige financiering op basis van deze uitkomsten (na een periode van voorwaardelijke financiering). Slechts in sommige gevallen beïnvloedde uitkomstenonderzoek (i.e. het verzamelen en analyseren van gegevens uit de klinische praktijk) het besluitvormingsproces in Nederland.

In lijn met het huidige beleid voor zogenaamde specialistische medicijnen (inclusief kostbare intramurale en extramurale medicijnen), betoogt dit proefschrift om het gebruik van voorwaardelijke financiering te beperken en meer de nadruk te leggen op initiële vergoedingsbeslissingen. Resultaten van kosteneffectiviteitsanalyses moeten gebruikt worden ter ondersteuning van deze (initiële) beslissingen, voor het behoud van een financieel duurzaam zorgsysteem, nu het aantal kankerpatiënten toeneemt en het aantal innovatieve en kostbare medicijnen groeit. Als kosteneffectiviteitsratio's boven een bepaalde ('equity'-gewogen) drempel, die de maatschappelijke bereidheid om voor gezondheidswinsten te betalen reflecteert, uitkomen, dan zijn prijsonderhandelingen (of andere maatregelen) essentieel om de relatie tussen kosten en effecten van een nieuw medicijn te verbeteren, en vergoeding van-

uit de basisverzekering mogelijk te maken. Alleen in specifieke gevallen (zoals bij medicijnen voor zeldzame ziekten) is uitkomstenonderzoek in staat om een meer robuuste schatting te geven van de kosteneffectiviteit. Deze gevallen dienen zorgvuldig te worden geselecteerd om de kosten van omvangrijke gegevensverzameling te beperken.

PhD portfolio

List of publications

About the author

PhD Portfolio

PhD candidate: Saskia de Groot
Institute: institute of Health Policy & Management, Erasmus University Rotterdam, the Netherlands
PhD period: 2009–2016
Promotors: Prof.dr. C.A. Uyl-de Groot
Prof.dr. L.A.L.M. Kiemeney
Copromotor: Dr. W.K. Redekop

PhD Training

Propensity Scores and Observational Studies of Treatment Effect. ISPOR short course: Berlin, Germany. 2012.

MSc in Health Sciences, specialisation Clinical Epidemiology.

Core Curriculum

- *Study Design*
- *Classical Methods for Data-analysis*
- *Clinical Epidemiology*
- *Methodologic Topics in Epidemiological Research*
- *Modern Statistical Methods*

Netherlands Institute for Health Sciences: Rotterdam, the Netherlands. 2009–2011.

Patient registries. ISPOR short course. Prague, Czech Republic. 2010.

Advanced Modelling Methods for Health Economic Evaluation. University of York: York, United Kingdom. 2010.

Academic Writing in English. Language & Training Centre, Erasmus University Rotterdam: Rotterdam, the Netherlands. 2010.

Teaching

Health Technology Assessment, master programme Health Economics Policy & Law, Erasmus University Rotterdam. Instructor computer lab. 2014–2016.

Health Economics, Erasmus Summer Programme, Erasmus MC - Netherlands Institute for Health Sciences. Instructor computer lab. 2014–2015.

Methods & Techniques 1, bachelor programme Health Policy & Management, Erasmus University Rotterdam. Tutor. 2014–2015.

Bachelor and master theses, bachelor programme Health Policy & Management and Master programme Health Economics Policy & Law, Erasmus University Rotterdam. Supervisor and co-evaluator. 2011–2014.

Statistics A, pre-master programme Health Policy & Management, Erasmus University Rotterdam. Tutor. 2013–2014.

Socio-medical sciences, bachelor programme Health Policy & Management and Pre-master programme Health Policy & Management, Erasmus University Rotterdam. Tutor. 2011–2014.

Podium presentations

Challenges and solutions to successfully determine real-world cost-effectiveness [Workshop]. ISPOR 18th annual European congress: Milan, Italy. 2015.

As real as it gets: challenges in setting up patient registries for the collection of real-world data on behalf of policy making [Workshop]. ISPOR 17th annual European congress: Amsterdam, the Netherlands. 2014.

Potential health gains for patients with metastatic renal cell carcinoma in daily clinical practice: A cost-effectiveness analysis of sequential first- and second line treatments. EuroTARGET meeting: Cambridge, United Kingdom. 2014.

First results PERCEPTION registry [presentation in Dutch: Eerste resultaten PERCEPTION-studie]. Symposium Wat stroomt er door de nier?: Tilburg, the Netherlands. 2013.

When is lower level evidence of effectiveness acceptable in reimbursement decisions? [Discussed]. LolaHESG 2012: Almen, the Netherlands. 2012.

PERCEPTION, a patient registry to inform outcomes research [presentation in Dutch: PERCEPTION, een patiëntenregistratie ten behoeve van uitkomstenonderzoek]. Symposium Nieuwe ontwikkelingen in de behandeling van niercelcarcinoom: Amsterdam, the Netherlands. 2012.

Poster presentations

Patient and disease characteristics are important determinants of health-related quality of life of patients with metastatic renal cell carcinoma: results from a population-based registry. ISPOR 17th annual European congress: Amsterdam, the Netherlands. 2014.

The cost-effectiveness of sequential first- and second-line treatments in metastatic renal cell carcinoma using real-world data and a patient-level simulation model. ISPOR 16th annual European congress: Dublin, Ireland. 2013.

The evaluation of the use and effectiveness of bevacizumab for patients with metastatic renal cell carcinoma in daily practice. ISPOR 15th annual European congress: Berlin, Germany. 2012.

Are population-based registries a suitable tool for outcomes research in cancer? Experiences from four registries. ISPOR 15th annual European congress: Berlin, Germany. 2012. *(nominated)*

A population-based registry for the evaluation of new treatment options for patients with metastatic renal cell carcinoma. ISPOR 14th annual European congress: Madrid, Spain. 2011.

When is lower level evidence of effectiveness acceptable in reimbursement decisions?: a decision algorithm to guide policy makers. ISPOR 14th annual European congress: Madrid, Spain. 2011.

List of scientific publications

Blommestein, H.M., **de Groot, S.**, Aarts, M.J., Vemer, P., de Vries, R., van Abeelen, A.F.M., Posthuma, E.F., Uyl-de Groot, C.A. Cost-effectiveness of obinutuzumab for chronic lymphocytic leukaemia in The Netherlands. Leukemia research 2016; *In Press*. [shared first authorship]

de Groot, S., van der Linden, N., Franken, M.G., Blommestein, H.M., Leeneman, B., van Rooijen, E., van der Hoeven, J.J.M., Wouters, M.W., Westgeest, H.M., Uyl-de Groot, C.A. Balancing the optimal and the feasible: A practical guide for setting up patient registries for the collection of real-world data for health care decision making based on Dutch experiences. Value in Health 2016; *In Press*. [shared first authorship]

de Groot, S., Sleijfer, S., Redekop, W.K., Oosterwijk, E., Haanen, J.B., Kiemeney, L.A.L.M., Uyl-de Groot, C.A. Variation in use of targeted therapies for metastatic renal cell carcinoma: Results from a Dutch population-based registry. BMC Cancer 2016;16:364.

de Groot, S., Redekop, W.K., Sleijfer, S., Oosterwijk, E., Bex, A., Kiemeney, L.A.L.M., Uyl-de Groot, C.A. Survival in patients with primary metastatic renal cell carcinoma treated with sunitinib with or without previous cytoreductive nephrectomy: Results from a population-based registry. Urology 2016;95:121-127.

Blommestein, H.M., Verelst, S.G., **de Groot, S.**, Huijgens, P.C., Sonneveld, P., Uyl-de Groot, C.A. A cost-effectiveness analysis of real-world treatment for elderly patients with multiple myeloma using a full disease model. Eur J Haematol 2016;96(2):198-208.

de Groot, S., Rijnsburger, A.J., Versteegh, M.M., Heymans, J.M., Kleijnen, S., Redekop, W.K., Verstijnen, I.M. Which factors may determine the necessary and feasible type of effectiveness evidence? A mixed methods approach to develop an instrument to help coverage decision-makers. BMJ Open 2015;5(7).

van Gils, C.W., **de Groot, S.**, Tan, S.S., Redekop, W.K., Koopman, M., Punt, C.J., Uyl-de Groot, C.A. Real-world resource use and costs of adjuvant treatment for stage III colon cancer. Eur J Cancer Care (Engl) 2015;24(3):321-332.

van Gils, C.W., **de Groot, S.**, Redekop, W.K., Koopman, M., Punt, C.J., Uyl-de Groot, C.A. Real-world cost-effectiveness of oxaliplatin in stage III colon cancer: a synthesis of clinical trial and daily practice evidence. Pharmacoeconomics 2013;31(8):703-718.

Uyl-de Groot, C.A., **de Groot, S.**, Steenhoek, A. The economics of improved cancer survival rates: better outcomes, higher costs. Expert Rev Pharmacoecon Outcomes Res 2010;10(3):283-292.

List of policy-related publications

Fayter, D., Büyükkaramikli, N., Birnie, R., **de Groot, S.**, Wolff, R., Armstrong, N., Worthy, G., Albuquerque de Almeida, F., Stirk, L., Al, M., Kleijnen, J. Pomalidomide with dexamethasone for treating relapsed and refractory multiple myeloma after at least two regimens including lenalidomide and bortezomib (review of TA338): a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd.

Advice to The National Institute for Health and Care Excellence, 2016.

de Groot, S., Aben, K., Sleijfer, S., Haanen, J.B.A.G., Bex, A., Kiemeney, L.A.L.M., Uyl-de Groot, C.A. Meaningful care: The use of expensive drugs in renal cell carcinoma [in Dutch: Zinnige zorg: Inzet van dure oncolytica bij niercelcarcinoom]

Advice to the Dutch National Health Care Institute, 2016.

Versteegh, M.M., Weistra, K., **de Groot, S.**, Redekop, W.K., Davis, P., Rutten, F.F.H., Oortwijn, W. Literature-based approach to defining the concept of healthcare which requires "highly specialised and cost-intensive medical infrastructure or medical equipment"

Advice to the European Commission, 2014.

de Groot, S., Rijnsburger, A.J., Redekop, W.K., Versteegh, M.M., Heymans, J.M., Link, A., Verstijnen, I.M., Boksteijn, B., Kleijnen, S., Staal, P. Academia II: Methodology Development, Sub-project 'Effectiveness and Evidence-Based Medicine [in Dutch: Academia II: Methodiekontwikkeling, Deelproject 'Effectiviteit en Evidence-Based Medicine']

Advice to the Dutch health care insurance board, 2012.

van Gils, C.W.M., Redekop, W.K., Mol, L., **de Groot, S.**, Tan, S.S., Koopman, M., Gaultney, J., Franken, M.G., Punt, C.J.A., Uyl-de Groot, C.A. Pilot outcomes research: Effects and costs of oxaliplatin in stage III colon and metastatic colorectal cancer

Advice to the Dutch health care insurance board, 2010.

About the author

Saskia de Groot was born in Gouda on December 6th 1985. In 2004 she started the bachelor programme Health Policy & Management at the Erasmus University Rotterdam (2004–2007). In 2009, she obtained her master's degree in Health Economics, Policy & Law. She continued her studies by enrolling in a research master programme in Health Sciences with a specialisation in Clinical Epidemiology at the Netherlands Institute for Health Sciences (2009–2011). During her studies at the Erasmus University Rotterdam, she started working as a research assistant at the institute of Health Policy & Management. In 2009 she started a PhD project on the evaluation of clinical and economic outcomes of new therapies for metastatic renal cell carcinoma, which resulted in this dissertation. Besides this research, she worked on various advisory projects, including different projects for the Dutch National Health Care Institute.

Dankwoord

Dankwoord

Dit proefschrift zou niet tot stand zijn gekomen zonder de hulp van vele mensen. Een aantal van hen wil ik hier in het bijzonder noemen.

Allereerst wil ik graag mijn promotoren bedanken. Carin, jij motiveerde me tot het schrijven van dit proefschrift. Bedankt voor deze kans en het vertrouwen! Je was altijd enorm betrokken, ik wil je daarom ook bedanken voor de prettige en persoonlijke samenwerking. Jouw toewijding om de kwaliteit van zorg te verbeteren is onvoorstelbaar, en daarmee weet jij veel mensen te inspireren! Bart, het was fijn om jou als tweede promotor te hebben. Jouw bijdrage aan het opzetten en onderhouden van het PERCEPTION register, evenals jouw aandeel bij de publicaties die hieruit voortkwamen vormden een belangrijk onderdeel van het welslagen van mijn promotieonderzoek. Bedankt daarvoor!

Ook wil ik graag mijn copromotor bedanken. Ken, jouw bijdrage aan mijn proefschrift is van onschatbare waarde. Onze gesprekken vormden veelal aanleiding tot aanvullende analyses, je wist me altijd te motiveren om nog net één stapje verder te gaan. Ik heb hier ontzettend veel van geleerd. Bedankt ook voor jouw geduld en alle kansen die je me de afgelopen jaren bood op het gebied van onderwijs en onderzoek.

Beste Stefan en John, bedankt voor jullie steun bij het opzetten van het PERCEPTION register. Zonder jullie vertrouwen was het niet gelukt dit register te laten slagen. Stefan, ik wil jou ook graag hartelijk danken voor je hulp bij de interpretatie van de onderzoeksresultaten. Beste Egbert, ook jou wil ik graag bedanken voor je hulp bij het opzetten van het PERCEPTION register, en de waardevolle feedback op de hieruit voortgekomen publicaties.

Mijn dank gaat ook uit naar de registratiemedewerkers en datamanagers van het Integraal Kankercentrum Nederland, in het bijzonder Paulien, Kees, Aysun, Anita, Anita en Karin. Ook al was de data-verzameling voor het PERCEPTION register omvangrijk en niet altijd eenvoudig, jullie zorgden ervoor dat de data tijdig beschikbaar was. Bedankt voor jullie doorzettingsvermogen en zorgvuldigheid. Daarnaast wil ik jullie graag bedanken voor het fijne contact. Ook wil ik de andere personen van het Integraal Kankercentrum Nederland bedanken die het tot stand komen en het onderhouden van het PERCEPTION register hebben ondersteund, in het bijzonder Marjorie, Katja, Sigrid, Wilma, Erica en Rob.

Apostolos, Hedwig, Igna, Laura, Matthijs, Pieter en Tim; zonder jullie aanwezigheid op de universiteit waren de afgelopen jaren lang zo leuk niet geweest. Lieve Hedwig, vanwege de vele raakvlakken van onze onderzoeken werkten we veel samen en dat was altijd een feestje. Samen ontwierpen we onze ziektemodellen, en ontwikkelden we onze vaardigheden. Maar er was zoveel meer, zoals de congressen en de vele kilometers die we samen hebben hardgelopen. Bedankt voor de fijne samenwerking en al het plezier dat we samen hadden, en daarnaast bedankt voor jouw zorgzaamheid! Lieve Igna, toen we nog studeerden werkten we samen in verschillende commissies. Na ons afstuderen, werkten we tegelijk, weliswaar op uiteenlopende onderzoeksterreinen, aan onze promotieonderzoeken. Het is fijn om jou in de buurt te hebben. Ik wil je graag bedanken voor jouw oprechte interesse in de voortgang van mijn proefschrift, en mijn leven buiten de universiteit. Lieve Laura, je bent niet weg te denken uit de afgelopen jaren. Wat heb ik genoten van onze tripjes naar onder andere Praag, Madrid, Berlijn en Dublin, maar ook Lola! Bedankt voor alle afleiding gedurende de afgelopen jaren en onze fijne gesprekken! Matthijs, bedankt dat je er vaak voor me bent en naar me luistert. Daarnaast dank voor de kansen die je me biedt binnen iMTA. Tim, bedankt voor zoveel meer dan je hulp bij het opmaken van mijn tabellen. Bedankt voor al je adviezen en steun in de afgelopen jaren. Het delen van een kamer met jou is één groot geluksmomentje ;-).

Lieve Jac, ook al zijn we al tijden geen kamergenoten meer, jij kan niet ontbreken in dit dankwoord. Jouw steun in de eerste jaren van mijn promotietraject was voor mij van hele grote waarde. Bedankt voor al jouw goede raad en de gezelligheid! Lieve Annemieke, samen starten we de dag vaak met een theetje (hoe groter, hoe beter!), en bespreken we wat ons bezighoudt. Bedankt voor jouw fijne aanwezigheid! Lieve Margreet, regelmatig hielp je mij bij het beantwoorden van vragen over de data-verzameling voor het PERCEPTION register. Ik had dat alleen niet gekund. Ook met andere vragen kwam ik graag naar jou, en jij nam altijd de tijd. Bedankt daarvoor!

Ook wil ik graag alle andere (oud-)collega's van het iBMG bedanken voor de fijne werksfeer en de gezelligheid tijdens de jaarlijkse uitjes.

Lieve Eva en Melissa, de allereerste jaren bij het iBMG bracht ik met jullie door. Samen studeerden we voor tentamens, bezochten we feestjes en waren we actief voor de FBMG. Wat was het een fijne tijd! Deze activiteiten zijn intussen vervangen door etentjes en borrels. En ook al zien we elkaar niet meer dagelijks, het blijft fijn om met jullie te praten, nu onder andere over onze ervaringen op de werkvloer. Dank jullie wel dat jullie er altijd voor me zijn, en ik mijn ervaringen met het geven van onderwijs en presentaties, en het schrijven van mijn proefschrift met jullie kon delen!

Lieve Eline, ook jij bent niet weg te denken uit mijn eerste jaren bij het iBMG. Wat maakte je me vaak aan het lachen met jouw rake opmerkingen. Tegelijkertijd was je zorgzaam en geïnteresseerd, en daarmee een hele fijne vriendin. Ook na ons afstuderen, toen we elkaar minder vaak zagen, toonde jij altijd veel interesse in mijn promotieonderzoek en alles daarbuiten. Ik mis je.

Lieve Marije, samen studeerden we in onze zolderkamers aan de Aegidiusstraat. Daarna volgden er vele experimenten in de keuken. Tegenwoordig kiezen we vaak voor wat meer gemak en gaan we lekker uit eten (tenzij we erop uit trekken op de racefiets ;-)) Bedankt voor al je interesse en alle gezelligheid in de afgelopen jaren. Ik weet zeker dat je een glansrijke carrière tegemoet gaat!

Lieve Daphne, nauwgezet volgde je de voortgang van mijn proefschrift, en de status van mijn artikelen. Ook al zijn we al jaren geen collega's meer en gaan we al jaren niet meer samen skaten, we voeren nog steeds eindeloze gesprekken en ik hoop dat we dat blijven doen. Bedankt dat je zo'n lieve vriendin voor me bent!

Lieve Anne, Charlotte, Jannet, Leontien, Manouk, Marieke, Marrit en Tara, bedankt voor alle gezellige momenten in de afgelopen jaren, in wisselende samenstellingen, en soms opeens weer compleet! Lieve Lot, intussen hebben we zoveel mooie plekken met elkaar bezocht, toen we nog studeerden, maar ook daarna. Ook al worden de tripjes steeds korter, het is altijd fijn om met jou op pad te gaan! Lieve Manouk, samen werkten we de afgelopen jaren aan onze proefschriften. We deelden tegen- en meevallers. Ook was er de nodige ontspanning, van samen eten (bij voorkeur kort ;-)) en festivalletjes bezoeken tot mooie reizen. Dank je voor te veel om op te noemen, maar in het bijzonder je enorme steun en lieve gebaren! Lieve Taar, wat bof ik dat we al zoveel jaren vriendinnetjes zijn. Bedankt voor de vele fijne momenten in de afgelopen jaren in de vorm van etentjes, festivalletjes en sauna-bezoeken. Daarnaast kan ik je niet genoeg bedanken voor je meeleven tijdens mijn promotieonderzoek.

Lieve Lin, pap en mam, zoals jullie weten bezorgde het schrijven van dit proefschrift me soms best wel eens kopzorgen. Ik was daarom soms gehaast en niet altijd ontspannen. Bedankt voor alle liefde die ik desondanks van jullie kreeg. Lieve Lin, ik kan me geen betere zus wensen. Dank je dat je altijd voor me klaarstaat, en het me altijd gemakkelijk probeert te maken. Ook wil ik je bedanken voor je enorme hulp bij de laatste loodjes van dit proefschrift. Lieve pap en mam, jullie hebben me altijd alle ruimte en mogelijkheden gegeven om mijzelf te ontwikkelen. Ook staat de deur bij jullie altijd open, het is heel erg fijn om bij jullie weer even op te laden. Bedankt voor jullie grenzeloze steun en het vertrouwen.

Colofon

Author
Saskia de Groot

Design
Joost Dekker Grafisch ontwerper

Published & Printed by
Drukkerij Mostert en van Onderen!

ISBN/EAN 978 94 90858 48 3

The PERCEPTION registry was financially supported by the Netherlands Organisation for Health Research and Development (grant number: 152001014).
Additional funding for the registry was provided by Pfizer B.V. (formerly Wyeth Pharmaceuticals B.V.) and Roche Nederland B.V.

All rights reserved. No part of this thesis may be reproduced or transmitted in any form by any means without written permission from the author.