General discussion

Cancer is the second leading cause of death after cardiovascular diseases in Europe, accounting for 26% of all death in 2016.[2] During the last three decades, the worldwide incidence of cancer increased by 50%, while mortality due to cancer increased with 20% during the same time frame.[2]

Novel treatment options such as chimeric antigen receptor (CAR) T cells are on the verge of revolutionising the field of haemato-oncology as they have the potential of actually curing certain types of cancer.[409] Nevertheless, prices for novel cancer drugs in general and for haematologic malignancies in particular, are high and increasing throughout the last decades.[45] To keep healthcare systems affordable, decision makers have adopted several measures to control the price at which novel treatments are reimbursed. Through a formal Health Technology Assessment (HTA), the value that patients and health systems perceive for the particular treatment can be assessed, and a price for reimbursement can be set accordingly.[7] Although the concept of HTA is already used since the 1980's, various challenges persist to this day.

Recently, the *European Network for Health Technology Assessment* (EUnetHTA) defined nine domains of HTA. This dissertation mainly focussed on identifying and addressing challenges in the domain of *costs and economic evaluation*. In addition, the domains of *safety*, *clinical effectiveness*, as well as *patient and social aspects* were taken into account. More specifically, this dissertation explored several challenges in assessing costs and cost-effectiveness of treatments in haemato-oncology. Several aims were defined in **Chapter 1** which were analysed in three parts. PART I explored challenges in the evidence synthesis for HTA. PART II aimed at providing evidence on the cost-utility of novel and expensive treatments in haemato-oncology. PART III described implications of these cost-utility analyses (CUAs) on healthcare decision making and investigated the impact of expensive immunotherapies for the treatment of cancer on the (future) healthcare expenditure in Europe.

This final chapter discusses various aspects and findings of this dissertation and ends with recommendations for future research and healthcare policy.

PART I: CHALLENGES IN THE EVIDENCE SYNTHESIS FOR HTA

Three distinct challenges in synthesising evidence for HTA were identified and addressed in this dissertation.

Synthesising evidence from existing economic evaluations

In the context of evidence-based medicine (EBM), health decision making should be based on so-called "evidence to decision frameworks", meaning that all relevant factors for a decision are assessed and considered both systematically and transparently.[410] To answer research

questions in the *costs and cost-effectiveness* domain of HTA, previously published economic evidence should therefore be reviewed systematically.[17]

To this end, researchers and decision makers could make use of specialised databases that focus on indexing published economic evaluations such as the U.K. *National Health Service Economic Evaluation Database* (NHS EED) or the *Health Economic Evaluation Database* (HEED). However, since funding had ended, the HEED is no longer accessible for searches since the end of 2014, and the NHS EED is no longer updated since March 2015.[71,72] Hence, finding all relevant economic information on a specific health topic has become a challenging and (more) time consuming task. Today, researchers and policy makers need to search economic evidence through databases that primarily index biomedical literature. Ideally, the resumption of the former specialised databases on economic evaluations could solve this issue. However, the reasons that led to their cessation might probably still persist and therefore these solutions seem unlikely. Simultaneously, no authoritative guidance on how to conduct systematic reviews of economic evaluations in biomedical databases was available.

To fill this gap, **Chapter 2** aimed at supporting researchers to prepare systematic literature reviews of economic evaluations for informing evidence-based healthcare decisions. As such, **Chapter 2** details the second step of a five-stepped approach of this process. The full five steps include (1) initiating a systematic review of economic evaluations, (2) identifying full economic evaluations (see Chapter 2), (3) data extraction, risk of bias and transferability (see Wijnen et al., 2016[70]), (4) reporting of results, and (5) discussion and interpretation of results. The entire approach including a brief summary of all steps is described elsewhere.[30]

Generally, this guidance was well received in the scientific community which is reflected in several citations of this work. More specifically, the WHO-INTEGRATE (INTEGRATe Evidence) framework version 1.0, advised to gather evidence related to the review of economic analyses following the guidance provided in **Chapter 2**.[410] This can be viewed as some sort of validation of the guidance.

Nevertheless, the guidance may become (partly) futile or outdated in the future. This may be due to several reasons. For instance, the way databases can be searched is constantly refined. In 2016, the biomedical research database *Embase* integrated a new search form enabling the database searcher to enter search terms separately for the different aspects of the PICO (Patient, Intervention, Comparator, Outcome) scheme.[411] The user interface then automatically suggests synonyms and combines those into a full search query.[412] This feature facilitates an instant conceptualisation of the PICO scheme for the respective research question. While this might save time for the user, important synonyms could still be missed, and truncated search terms are not added. Our guidance offers suggestions to incorporate a variety of search terms and is therefore still relevant. Also, other databases do not yet provide such elaborate user interfaces. Searching *Embase* with the integrated PICO search tool might thus be a good starting point in designing a new search strategy. Subsequently, the query could be critically assessed with the guidance presented in **Chapter 2**. However, such a

process is not covered by the current version of the guidance and an update in this regard is warranted. For all other databases this guidance still offers useful resources for building a search query with a desired level of specificity and sensitivity.

The evolution of biomedical research databases may call for another update of the guidance soon. Not only because user interfaces are increasingly refined and tailored to the researchers' needs. Novel, intelligent and automated search algorithms are at the verge of changing the way literature is searched in general. Machine learning algorithms and tools are promising approaches to reduce the workload of systematic literature reviews and can be applied to inform evidence synthesis already today.[413,414] However, such tools are still under development and may require an extensive background in information technology or biomedical informatics.[414,415] Until such approaches have become fully mature and available to a broader audience, **Chapter 2** may serve as a practical guidance to undertake "handcrafted" systematic searches for economic evaluations.

Synthesising evidence on effects

Data from randomised controlled trials (RCTs) are often referred to as the "golden standard" of collecting clinical evidence on safety and efficacy of a drug.[36] Consequently, evidence on effects for HTA is preferably collected through such studies.[17,19,416] However, these studies are not *per se* designed to inform economic evaluations. RCTs are for instance primarily powered to detect differences in clinical outcomes. Since variables related to costs have higher variance, the required sample size to detect differences in costs is considerably higher when compared to the needed sample size to detect differences in clinical effects.[417,418] Furthermore, clinical studies have limited follow-up times. Consequently, desired clinical benefits such as overall survival (OS) can often not fully be captured. Therefore, so-called surrogate endpoints are often chosen as primary trial endpoints. In the field of haemato-oncology, these often include event-free survival, freedom of treatment failure, or progression-free survival (PFS). Surrogate endpoints such as PFS do not usually reflect better health-related quality of life (HRQoL) or overall longevity.[45] Outcomes of the study in **Chapter 6** seem to confirm the former, since HRQoL for patients in progression-free disease did not differ from patients in progressive disease. This aspect is discussed in more detail in **Part II** of this Chapter.

Patient selection in accordance with the decision problem

The decision to reimburse a novel treatment is usually taken for a (sub-)group of patients with a specific disease. Therefore, it is important that economic evaluations clearly define the target patient population including all relevant subgroups. When data from RCTs are used to inform efficacy parameters of economic evaluation, challenges may arise. This is because RCTs are conducted under strictly controlled and idealised conditions. The selection of a narrowly defined patient population can hence potentially affect external validity, which is need when (reimbursement) decisions need to made on a more general level.[419] A recent

literature review showed that the majority (71.2%) of included RCTs in cardiology, mental health, and oncology reported that their samples were not broadly representative of real-world patients.[420] More specifically, patients enrolled in oncological studies were found to be often younger, less likely to be female, have better performance status, and better disease prognosis than real-world cancer patients.[420]

This issue poses a general challenge to the generalisability of all economic evaluations making use of RCT data, especially when the (reimbursement) decision problem focusses on specific patient subgroups of the gathered evidence.

In **Chapter 7** for instance, evidence on effects were used for a subgroup of the trial population in the final analysis. In this way the decision uncertainty for the reimbursement authority could be reduced to an extent that the treatment could be accepted for reimbursement. However, initially, the trial was not designed to detect differences in efficacy between the novel treatment and its comparator for the relevant subgroup. Consequently, *post hoc* subgroup analyses were necessary for the economic evaluation.

While *post hoc* subgroup analyses are possible for clinical trial data, the *Consolidated Standards of Reporting Trials* (CONSORT) initiative strongly criticised such approaches.[421] To allow for more robust estimates, subgroups analyses should be pre-specified and it is suggested that such specifications could be made mandatory, at least for publications.[422]

In **Chapter 6**, the relevant subgroup for the reimbursement decision problem were pre-stratified. Therefore, using evidence from this subgroup for the CUA did not introduce additional uncertainty. This was however only possible since patient-level data were available. Usually, such subgroups analyses cannot be perform since the relevant empirical survival data is often not published.[318]

To tackle the issue of low external validity of RCT data, adapting trial designs to include more representative patient samples or supplementing RCT data with evidence from supportive studies could be a solution.[420] Regarding evidence of specific subgroups, clearly defined and pre-specified subgroups could offer more valid and reliable outcomes, relevant for economic evaluations. In addition, such evidence should be published so that relevant cost-effectiveness studies can be conducted without having access to patient-level data.

Estimating long-term efficacy

Since clinical studies usually have restricted follow-up times and most jurisdiction prefer a lifetime horizon for economic evaluations,[19,26] there is a need to extrapolate the empirical data to a longer time horizon. Researchers can choose between several techniques, and parametric survival models are often used for this purpose. The literature refers to the exponential, Weibull, Gompertz, log-normal, and log-logistic models as "standard" parametric models.[306] These models are frequently used in economic evaluations and their employed methodology including their strength and weaknesses is well documented.[211,306,423] When parametric models are used to extrapolate patient data of clinical studies, both internal and external

validity need to be verified. Internal validity is commonly assessed through visual inspection of the fitted curves to the observed survival. In addition, model fit criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are used for this purpose. External validity is subsequently assessed by comparing the extrapolated survival to data from other studies (when available) or the general population mortality (extrapolated data should not exceed this mortality). In addition, the plausibility of the extrapolated long-term survival should be validated by clinical experts.[211]

Chapters 5 and 6 employed such standard parametric survival models to extrapolate the empirical survival data of the respective clinical studies. The results showed that the internal validity of the extrapolated curves was high since all parametric survival curves visually fit the empirical data well and both AIC and BIC values did not differ to a great extent. While internal validity could be established, long-term estimates differed considerably. This aspect of survival extrapolation is long recognised in the literature.[306] To ensure external validity in **Chapters 5 and 6**, clinical experts were consulted to validate the long-term estimates. However, as of yet, standard methods for the elicitations of expert opinion for HTA are scarce,[424] and although some tools exist to aid in this endeavour, none of them focus of the external validation of parametric survival models.[425–427] This may introduce some uncertainty in the elicitation process. For the studies presented in **Chapters 5 and 6**, clinical experts were asked to validate long-term survival by means of a semi-structured questionnaire in combination with subsequent telephone interviews. This may be a possible solution in the absence of a clear methodological guidance. However, depending on the complexity of the disease, renowned clinical experts may be difficult to find. For **Chapter 5** for instance, only one clinical expert could be included. This was because paediatric and young adult patients with relapsed or refractory B-cell acute lymphoblastic leukaemia (pALL) are only treated in one specialised centre in the Netherlands. Consequently, the number of clinical experts available to consult about this particular indication is limited. Although for **Chapter 6** more clinical experts could be included, expert opinion in general may be biased with regard to the experts' experience, training, mood or motivation.[428,429]

Due to these limitations, there is an expressed need to improve parametric survival extrapolations to incorporate external (long-term) data.[430–433] Recently, Vickers (2019) evaluated several survival curve extrapolation techniques using long-term observational cancer data.[434] The author generally concluded that long-term survival predictions can be improved by directly using such mature long-term data and recommends different methodological approaches to extrapolate survival data and to integrate long-term data, depending on the observed treatment benefit. Nevertheless, several limitations where stated with regard to the data used in the study. Most importantly, the different techniques were assessed using data from patients ≥80 years old, which might have led to biased results.[434] This is particularly important to acknowledge when the suggested methodology is used to extrapolate data of patients with haematological malignancies, since in Europe, their average age at initial diag-

nosis is considerably lower (i.e. approximately 64 years; median: 69 years; standard deviation [SD] or range not reported).[435]

Extrapolating empirical survival of cancer patients remains a challenging task. And although guidance exist to support researchers in employing "standard" parametric models,[211] the plethora of alternative modelling approaches such as for instance model-averaging techniques,[436,437] hybrid models,[438,439] or cure models[440] show that this is an evolving field of research. Therefore, it seems that a definitive guidance cannot yet be issued. Future research should focus on validating promising modelling techniques to enable robust long-term estimates of survival data for the use in economic evaluations. Simultaneously, existing guidance for such methodologies should frequently be updated to incorporate newest findings and recommendations.

Choosing an appropriate treatment comparator

Since economic evaluations are comparative studies, any novel treatment that will be assessed for its cost-utility needs to be compared to at least one other course of action.

The importance of choosing an appropriate comparator for any economic evaluation is highlighted by the Drummond checklist for assessing economic evaluations.[19] The second question of this checklist focusses on whether a comprehensive description of the competing alternatives was given and whether *relevant* alternatives were omitted. Already in 1998, Mullins and Ogilvie, concluded that most pharmacoeconomic guidelines agreed on an appropriate comparator being either a therapy currently used in standard practice, or a therapy that is most likely to be replaced by the novel treatment.[441] More recently, reviews of country-specific economic evaluation guidelines came to similar conclusions.[442,443] Consequently, the EUnetHTA stated in its 2015 guideline that comparators should "reflect most relevant alternative intervention(s) used in clinical practice".[443]

Despite such clear and longstanding preferences for comparators in economic evaluations, choosing a relevant comparator was challenging for **Chapters 5 and 6**. Although several treatment alternatives existed for the patient population studied in **Chapter 5**, no standard of care was yet defined. Therefore, all commonly administered treatment alternatives in the reference country, for which survival data were available, were selected for the evaluation. Similarly, no clear standard of care was defined for patients studied in **Chapter 6**. And although clinical experts indicated several treatment options for the reference country, reliable survival data of these treatments were not available.

Using phase II clinical data

Conducting pivotal phase III trials is a time-consuming effort and it can take decades until such studies are concluded. To improve a timely access for patients to new medicines, the *European Medicine Agency* (EMA) established the so-called PRIME scheme to "enhance support for the development of medicines that target an unmet medical need".[444] Consequently,

evidence from phase II clinical trials is increasingly used by the EMA to decide on the marketing approval of new (cancer) treatments.[445] The question arises to what extent phase II clinical data can be used to inform early HTAs.

Following the *ISPOR Good Research Practices in Modelling Task Force*, conceptualising (decision) models for economic evaluations is an iterative process, involving several steps and stakeholders.[210] To keep the time between marketing approval and the possible reimbursement of novel and expensive treatments as short as possible, a timely beginning of conceptualising the model is warranted. **Chapter 4** demonstrated that evidence from previously published phase II studies can be used to conceptualise such a model and to simulate long-term survival outcomes. Such simulations make a *de novo* decision model transparent and discussable within the scientific community. In addition, the simulation results already indicate the magnitude of the efficacy that can be expected from the novel treatment. The proposed model in **Chapter 4**, together with its results was generally accepted by scientific peers, supporting its credibility.[446,447] Complemented with clinical evidence from a comparator treatment and input parameters on costs, this model could be used for a model-based CUA. In such a case, an early HTA could be conceivable, accelerating later reimbursement procedures. In case results of such an analysis show that the novel treatment would result in a significantly unfavourable cost-effective ratio when compared to the pertinent willingness-to-pay (WTP) threshold, the use or development of the novel treatment could be halted. Nevertheless, the analysis of **Chapter 4** needs to be interpreted as an *indication* for the long-term survival of the studied patient population which needs to be validated through phase III clinical data. Currently, the multi-centre, international phase III clinical trial RETHRIM aims at creating such needed evidence. However, issues with patient accrual have delayed the end of the study extensively. This is due to several reasons such as improved preventive measures to develop the disease (acute graft-versus-host-disease [aGvHD] in this case) and the relatively small indication (aGvHD can be considered as rare[448]).

In **Chapter 5**, phase II clinical data was used to perform a formal CUA for a Dutch reimbursement dossier. Previously, the studied therapy (tisagenlecleucel) had received marketing approval following the EMA's PRIME scheme based on phase II clinical data.[50,449] Consequently, evidence from an RCT were not available at the time the economic evaluation was conducted. With the evidence available, CUAs could be conducted in several jurisdictions, including the one of **Chapter 5**.[254,256–258,389,450,451] Nevertheless, most European Member States did not opt for a "classic" reimbursement of the therapy. All (former) EU-5 Member States (i.e. France, Germany, Italy, Spain, and the UK) have adopted some kind of outcomes-based reimbursement scheme (OBR).[407] Reasons for this included the high costs of the treatment but also the considerable amount of decision uncertainty stemming from restricted efficacy data.[407] In the Netherlands, the decision to reimburse the treatment was based on a simple budget impact analysis (BIA) instead of making use of the available results of the CUA in **Chapter 5** (more about this in Part II of this discussion). Consequently, the

novel treatment was reimbursed following a standard procedure in the Netherlands and no OBR was negotiated.

Similar to the findings of **Chapter 4**, the results of **Chapter 5** need to be validated with long(er)-term efficacy data, preferably from phase III clinical data.

Synthesising evidence on costs

When economic evidence is not available or transferrable to the setting of interest, new evidence needs to be generated. **Chapter 3** can be seen as a case study to estimate healthcare costs based on two distinct approaches.

First, healthcare costs of paediatric patients with sickle cell disease were based on the pertinent clinical practice guideline (CPG). These costs could quickly be estimated since the expected resource use frequency was described in sufficient detail and could be valued with reference prices. However, such an approach only considers the *standard* resource use of patients. Consequently, it does not provide insights into real-word resource use. As such, it neglects any additional or emergency visits to the hospital for instance. Depending on the studied disease, such visits can be relatively frequent (e.g. in the case of sickle cell disease[452]) and should therefore be considered when costs are estimated.

Therefore, a second approach was explored. For this, available patient level data of a hospital financial claims database was used to gather information of real-world resource use. This approach was significantly more challenging when compared to the first one due to several reasons. For instance, receiving access to the financial claims database was complicated. Since hospital claims data are generated for billing purposes,[453] they are usually managed by the financial department of a healthcare institution whose primary aim is not the support of scientific research. Also, claims data hold less information on patients than for instance medical records. While this makes them easily de-identifiable,[453] it is not always clear which of the recorded measures are most useful to represent utilisation.[454] This can only be resolved in a dialogue with the healthcare professionals and the data administrator.

In conclusion, hospital claims databases are not designed to support research endeavours on costs. In addition, such databases are subject to frequent updates and changes. The respective Dutch reimbursement system for instance was introduced in 2005 and profoundly revised in 2012.[455] It can therefore still be considered as a system under development. Especially at the time data for the initial analysis of **Chapter 3** was requested (2017). Enhancements and updates of hospital information systems under development are necessary and important. However, such changes may render previous or newly collected data incompatible and hence not useful for data analyses. Nevertheless, claim-based studies are conducted at least since the 1980's for manifold purposes in the US.[456] A review of such studies between 2000 and 2005 in five healthcare journals found that the majority used claims data to study aspects of access to healthcare (49%), followed by quality of healthcare (24%), and interventions,

therapies, or treatments (13%).[457] Studies on healthcare costs were not mentioned, although some studies exist.[458,459]

Generally, the literature suggests that using claims data to estimate the cost of illnesses is feasible and may provide access to a relatively large sample size while avoiding selection bias.[458–461] This is in line with the findings reported in **Chapter 3**. In addition, **Chapter 3** demonstrated that hospital financial claims data can be used for the estimation of real-world healthcare costs. Supplemented with patient characteristics such as age, sex, and diagnosis, such databases can be a powerful and reliable source of information that is readily available and frequently updated. Alternatively, as mentioned in **Chapter 1**, information on resource use could be synthesised from patient questionnaires, although this may be related to some bias.[43] Future research could compare either approach (i.e. collecting resource use data from financial claims databases versus patient questionnaires) to establish the comparative evidence on the validity of either method.

PART II: COST-UTILITY OF NOVEL TREATMENTS IN HAEMATO-ONCOLOGY

Costs of novel cancer treatments for haematological conditions are high and can put the affordability of other new treatments at risk. Therefore, reimbursement decisions need to be made in a transparent and systematic way. CUAs can provide the needed information to make evidence-based decisions in healthcare and are therefore a vital part of the reimbursement decision process. This dissertation assessed the cost-utility of two novel and expensive treatments that entered the European market in 2019.

The cost-utility of tisagenlecleucel when compared to clofarabine monotherapy (Clo-M), clofarabine combination therapy (Clo-C), and blinatumomab (Blina) in paediatric patients with relapsed/refractory acute lymphoblastic leukeamia (pALL) was assessed in **Chapter 5**. Based on the disease burden and applicable willingness-to-pay (WTP) threshold, tisagenlecleucel could be considered cost-effective from all perspectives studied. Several other CUA assessed the same therapy in Canada,[389] the US,[254,257,451] and the UK.[258] All of these studies came to the same conclusion of tisagenlecleucel being cost-effective in the base-case analysis.

Similarly, **Chapter 6** assessed the cost-effectiveness of lenalidomide plus rituximab versus rituximab monotherapy for previously treated follicular lymphoma (FL). Based on the selected WTP-threshold for the base-case analysis, lenalidomide could be considered cost-effective. However, the probabilistic sensitivity analyses indicated some uncertainty in the results. In addition, some scenarios exceeded the chosen WTP-threshold for the base-case analyses, rendering the treatment not cost-effective. To this date, two other studies are available that assessed lenalidomide in previously treated FL in the UK and China.[316,318] While the treatment was considered cost-effective in a UK setting, the Chinese study concluded

that lenalidomide was not cost-effective. Several reasons could have led to the conclusion of the latter. Most importantly, Zhang and colleagues assumed a rather restrictive time horizon of ten years.[318] In such a short time frame, differences in both effects and costs of the two studied treatments could not fully be captured. Assuming a similar time horizon in our study would also lead to an ICER above the assumed WTP-threshold.

Including future non-medical costs in a societal perspective

As described in **Chapter 1**, two main perspectives used in CUAs stand out: the healthcare perspective and the societal perspective.[21] For either perspective, most pharmacoeconomic guidelines prefer a lifetime horizon on costs.[26] Consequently, the inclusion of future costs should be considered as well. A systematic review published in 2015, found that the number of publications to include future costs in general increased by 40% (to 70.8%) between 2008 and 2013.[27] While most of the studies detected (i.e. 49%) incorporated future medical costs related to the studied disease (also referred to as related future costs), only 4.2% included future medical costs not related to the studied disease (also referred to as future unrelated medical costs). None of the studies included in the review considered future non-medical (i.e. productivity or consumption) costs. A reason to not include the latter may be that most pharmacoeconomic guidelines do not yet explicitly mentioned their inclusion.[26,27] Some guidelines (e.g. in the US) however already do so and it seems that the scientific debate on whether to include these costs settles in favour of the US approach.[26,27,259,287] Regardless this ongoing debate, the impact of including future non-medical costs in CUAs remains understudied.[26,27]

Both **Chapters 5 and 6** assessed the cost-utility of a novel treatment in haemato-oncology from a healthcare and a societal perspective. In addition, future non-medical costs were included to determine the impact of this component on the ICER. The inclusion of the latter costs was possible due to a recent update of the iMTA *Practical Application to Include Disease Costs* (PAID) tool.[276] Hence, the studies presented in **Chapters 5 and 6** were the first to incorporate these costs in their assessment and can therefore be seen as case studies to examine the impact on the ICER when these costs are added to a societal perspective.

Including future non-medical consumption costs in a societal perspective, lead to an increase in the ICER between 17% and 21% in **Chapter 5,** when compared to the societal perspective alone. When future non-medical costs were included in a societal perspective in **Chapter 6**, the ICER increased by approximately 22% when compared to the societal perspective alone. However, despite these rather large increases in the ICER, the results of these two studies showed that considering future non-medical consumption costs within a societal perspective does not necessarily influence the decision to consider a treatment cost-effective. Whether or not this holds true for other studies as well depends on the modelled efficacy of a treatment, its costs, and the assumed WTP-threshold. When incremental effects (i.e. life years or QALYs) are relatively high for the novel treatment, considering

future non-medical costs might not influence a positive reimbursement (as seen in **Chapter 5**). However, when the ICER from a societal perspective is already close to the assumed WTP-threshold, considering future non-medical costs might lead to an ICER exceeding the WTP-threshold (as seen in **Chapter 6**).

It needs to be noted that especially the results of the study presented in **Chapter 5** are related to some uncertainty. Since the patient population of interest were children and young adults, the modelled time horizon was substantially longer when compared to patients included in **Chapter 6**. This not only introduces uncertainty in the modelling of long-term efficacy but also in modelling cost items such as productivity losses or gains. Typically, economic evaluations from a societal perspective consider future production costs. These are usually referred to as productivity losses when adults are the patient population of interest. When children and young adults are the patient group of interest however (as in **Chapter 5**), they are usually not yet part of the workforce and hence productivity *losses* cannot not be considered. Instead, disease survivors might be able to work in the future. Hence productivity *gains* should be considered for this group. Modelling potential productivity gains in this population is challenging since little is known about both educational and employment prospects of long-term survivors of childhood cancer. Also, the estimation of potential productivity gains should account for future fluctuations on the job market for the jurisdiction of interest which introduces additional sources of uncertainty.

The decision as to whether or not future non-medical costs should be considered in economic evaluations should not be driven by the impact of these costs on the ICER. It is rather a fundamental methodological question within health economics that needs to be answered through scientific discourse. While this debate is currently ongoing, clear recommendations put forward through pharmacoeconomic guidelines have probably the biggest potential to resolve this uncertainty. The recommendations issued by the US *Second Panel on Cost Effectiveness in Health and Medicine* in 2018 were a first step in this direction.[287] Whether other countries will follow is not yet clear. However, given the potential of exceeding the WTP-threshold when future non-medical costs are considered, their impact should at least be explored in sensitivity analyses of economic evaluation already now. Especially, since the recent update of the iMTA PAID tool facilitates the inclusion of these costs, at least for the Netherlands.[259] Arguably, this tool should be extended to also include cost estimates from other countries to allow for a more seamless implementation in different jurisdictions. Efforts to do this are already underway as a recent publication of a separate PAID version to include future unrelated medical costs for economic evaluations in England and Wales show.[462] It is unfortunate that this version does not include future non-medical costs altogether.

The need for accurate health state utility values

Health state utility (HSU) values are an integral part of CUAs as they allow the calculation of quality-adjusted life years (QALYs). Since most contemporary clinical trials collect health-

related quality of life (HRQoL) data, they can be an important source for synthesising such evidence.[463] Nevertheless, HSU values are regarded as one of the most uncertain input parameters of cost-utility models which simultaneously have the potential to heavily impact both accuracy and precision of the model results.[464]

Despite their relative importance to economic modelling studies, HSU estimates of haematological malignancies seem to be rarely reported in the literature. For the indication of FL for instance, utilities from a study of Wild et al.[340] are commonly used for CUAs in this field.[465–472] Similarly, studies examining HRQoL of patients diagnosed with multiple myeloma were found to be scarce.[473]

Results from the widely cited study of Wild et al.[340] are only available in form of a conference abstract and therefore the employed methodology is not fully disclosed. For the study in **Chapter 6**, patient level HRQoL data were available and therefore HSUs for the economic model were estimated based on the most recent ISPOR guideline.[464] Outcomes of this analysis suggested no statistically significant difference between either administered treatment in the clinical study. This was in line with the analysis of the same data conducted by the principle investigator of the trial.[300] More importantly however, our analysis also did not find any differences in HRQoL between the health state of progression-free and progressive disease. This finding is at least counterintuitive as a recent cohort study including more than 2,000 patient with metastatic breast, pancreatic, lung, or colorectal cancer found that disease progression is related to worse outcomes in many HRQoL scales.[474] It needs to be noted that this study did not include patients with haematological cancers, nor did the study include a generic HRQoL questionnaire such as the EQ-5D.[299] For economic evaluations in healthcare, such generic questionnaires are generally preferred as they allow a comparison of outcomes across different diseases.[19] Currently, a comprehensive overview relevant HSUs for haematological diseases is lacking. A systematic summary in the form a literature review could shed light on this. Alternatively, a cohort study including patients with haematological malignancies could reveal HSUs for this patient group. In this way, findings on the HSU used in **Chapter 6** could be validated.

PART III: IMPLICATIONS OF CUAS ON HEALTHCARE DECISION-MAKING

Implications of CUAs on reimbursement decisions in haemato-oncology

As stated in **Chapter 1**, economic evidence generated from CUAs can guide reimbursement decision-making in healthcare. In this dissertation, the policy implications were studied for the three CUAs described in **Chapters 5 to 7**. All three analyses assessed the cost-utility of novel, expensive treatments in haemato-oncology and provided the first economic evidence

on the treatment of interest for the pertinent jurisdiction. To understand the implications of the CUA results on policy and decision making, **Chapters 5 to 7** need to be interpreted in the context of the respective jurisdiction. While **Chapters 5 and 6** were conducted to primarily inform Dutch reimbursement decision, **Chapter 7** was performed for a UK perspective.

The relevance of CUAs when the estimated budget impact is low in the Netherlands

In the Netherlands, CUAs are required to inform reimbursement decisions for novel and expensive treatments. To keep the time between marketing authorization and the decision for reimbursement as short as possible, manufacturers often commission independent research institutes to perform the CUA for the eventual submission of evidence to the *National Health Care Institute* (Zorginstituut Nederland, ZIN). Studies presented in **Chapters 5 and 6** are both examples of this process as they were both submitted for the reimbursement decision process of the treatment studied. Nevertheless, results of neither Chapter seemed to have played a role in the decision to reimburse the novel treatment.

Despite its relatively high list price of 320,000 EUR per patient, the budget impact of tisagenlecleucel (**Chapter 5**) was seen as "low" and below the threshold that requires a full cost-effectiveness analysis. After all, the ZIN estimated that approximately 9 patients per year would receive the drug. When compared to the current standard treatment of blinatumomab, the ZIN estimated that this would lead to incremental costs of 2.1 million EUR per year, considering treatment, administration and monitoring costs. Nevertheless, results of **Chapter 8** show that these costs are likely to increase in the future due to an increase in eligible patients.

Although the Dutch EE guideline prefers discounted, societal costs from a lifetime perspective to support evidence-based decisions in healthcare,[314] the estimates from the budget impact analysis (BIA) for tisagenlecleucel only considered undiscounted direct medical healthcare costs for one year. Such an analysis neglects several aspects that are specific to CAR T-cell therapies and influence both costs and patient outcomes.

CAR T-cell therapy is a particularly complex treatment requiring an extraction of the patient's own T cells.[475] These cells are then transported to a specialised facility where they are genetically engineered to become CAR T-cells that target the desired cancer cells.[476] After the CAR T-cells are amplified by several million-fold, they can be transfused back into the patient.[477] This is a time consuming process with a median manufacturing time of 23 days (range, 21-37 days) from receipt of the material at the manufacturing facility to return to the clinical facility.[478] It needs to be noted that these figures are only valid for US which holds at least two centralised manufacturing facilities.[478] For the European market the first manufacturing facility is currently built in Switzerland and estimated to be functional by the year 2021.[479] Until then, shipment of patients' own T cell and CAR T-cells between

Europe and the US is necessary, adding to both costs and wait time. Only recently a study on the impact of increased wait times on overall mortality of CAR T-cells in large B-cell lymphoma (DLBCL) found that even a modest delay in the therapy significantly impacts its effectiveness negatively.[409] There is little reason to believe that this impact might be different in other indications. In addition to the high costs of pre- and post-treatment, some patient might not survive such wait times and decease before receiving the final product. All such (potential) negative impacts on both effect and costs of the novel treatment were not and cannot be considered in a BIA and hence its value of information to the reimbursement decision can be questioned.

In a similar case, the results of the CUA presented in **Chapter 6** showed a considerable level of uncertainty of the novel treatment (R-LEN) being cost-effective in a Dutch setting. Nevertheless, the Dutch Minister of Health described the reimbursement negotiations with the manufacturer as an "exceptional" case, for which the ZIN was not asked for advice.[321] He argued that the treatment was originally placed in the lock due expected high costs for the treatment of another indication (multiple myeloma). Once in the lock, the treatment would automatically be exempted from the basic insurance package for all new indications. For the recent extension of marketing authorisation, a previously conducted horizon scan of the ZIN had expected no more than 10 to 15 patients on a yearly basis for the indication at hand.[321] Therefore, total healthcare costs for the new indication were expected to not exceed the amount of 1 million EUR per year.[321] Specifics of this calculation were not disclosed but it can be assumed that these estimates were based on a BIA.

Although the budget impact for novel treatments studied in **Chapters 5 and 6** were seen as too low to require a full HTA for the reimbursement assessment by the Dutch authorities, it needs to be noted that both treatments currently have marketing authorisation for several indications in the EU. Tisagenlecleucel (**Chapter 5**) holds central marketing authorisation by the EMA for children and young adults with B-cell acute lymphoblastic leukaemia (pALL), and diffuse large B-cell lymphoma (DLBCL).[50] Lenalidomide (**Chapter 6**) is authorised by the EMA for the treatment of multiple myeloma, myelodysplastic syndrome, mantle cell lymphoma, and FL.[480] And although the argumentation of a "low budget impact" might hold true when considering one particular indication, the budgetary implications for these two treatments across all indications might differ significantly. Indeed, the results of **Chapter 8** show that the costs for CAR T-cell therapy alone could be as high as 54.5 million EUR for the treatment of patients with pALL and DLBCL in the Netherlands in 2020. Similarly, a recently published investigation showed that lenalidomide alone generates a global annual revenue of approximately 6.2 million EUR, making it the most successful treatment in terms of total revenue.[481]

While BIAs can inform reimbursement decisions, they are also associated with considerable uncertainty.[482] Due to limited data availability for most input parameters, standard methods used in economic evaluations such as one-way or probabilistic sensitivity analyses

cannot be conducted.[483] A systematic review of BIAs on pharmaceutical drugs in the EU found that most of these analyses were conducted with poor methodological quality.[484] In addition, 76% of the BIAs considered were not accompanied by a full economic evaluation.[484] Another review of BIAs for the US market found that the ratio of predicted versus actual budget impact ranged between 0.2 and 37.5 with a mean value of 5.5.[485]

Admittedly, for the particular case of tisagenlecleucel, outcomes of the economic evaluation in **Chapter 5** and results from the BIA would essentially lead to the same positive recommendation to reimburse tisagenlecleucel. However, this is not necessarily the case for **Chapter 6**, as the probability of R-LEN being cost-effective heavily depends on the chosen perspective and the several other assumptions.

Basing reimbursement decisions of novel and expensive treatment options on BIAs alone may be a way to quickly derive conclusions. Nevertheless, such an approach does not allow for optimal resource allocation or to generate evidence on the health gains of a particular intervention. In 2015, the ZIN warned that if no such information was gathered, "society may end up spending money on interventions that result in relatively few health gains for patients".[486] Using information from BIAs to derive reimbursement decisions undermines the goal to derive such decisions on the basis of a solid evidence-based framework. CUAs *per contra* are better suited to support such decisions.

The impact of immature RCT data on reimbursement decisions

Chapter 7 demonstrated that, in the case of immature RCT data, reimbursement decisions can become increasingly uncertain. Although the estimated ICER in the study presented in **Chapter 7** was below a WTP-threshold of 30,000 GBP/QALY gained, the NICE appraisal committee initially issued a negative reimbursement decision. This was because the committee found that the estimates of the cost-utility analysis were not robust enough, due to the immaturity of the clinical trial data. After narrowing down the patient population and providing a higher financial discount on the price for the intervention, the degree of decision uncertainty was reduced to an extent that the appraisal committee could issue a positive recommendation. This demonstrates that the choice of the patient population is of crucial importance to determine whether a treatment should be reimbursed. This finding correlates with the issues of *post hoc* subgroup analysis discussed earlier in this chapter. Hence, involving reimbursement authorities early in the discussion of which subgroups are of eventual interest for reimbursing novel treatments might help in setting-up RCTs with relevant subgroups.

The financial impact of novel immunotherapies in haemato-oncology

The chimeric antigen receptor (CAR) technology is seen a biologically and economically powerful tool.[487] Biologically powerful because CAR T cell therapy has been able to cure cancer in some patients for whom chemotherapy had failed.[487] Economically powerful because the CAR T cell therapy tisagenlecleucel was considered the most expensive oncological

therapy available at the time of its marketing approval by the US *Food and Drug Administration* (FDA) in 2017.[451,488]

Based on HTA evidence, all countries of the former EU-5 have adopted outcomes-based reimbursement scheme to safeguard patients' access to the therapy.[407] As mentioned earlier, in the Netherlands, "standard reimbursement" was negotiated on the basis of a BIA.

Nonetheless, the *European Haematology Association* (EHA) was concerned about the sustainability of the pricing of novel oncological treatments in haematology. Especially, since CAR T-cell therapies can be used to treat many more haematological malignancies than the ones for which they currently hold market approval (i.e. ALL and DLBCL). The potential financial impact of this therapy on European health systems was unknown, and **Chapter 8** aimed to shed light on this issue.

Based on ongoing clinical studies, it was estimated that the use for CAR T-cell therapies might be expanded to include other haematological indications such as mantle cell lymphoma (MCL), FL, acute myeloid leukemia (AML), multiple myeloma (MM), and lymphocytic leukemia (CLL). **Chapter 8** found that the future cost of CAR T-cell therapies in the field of haematology will be substantial and increasing. This increase will not only be caused by an extension of indication but also because the underlying population forecasts predicted an increase in the eligible patient population. Population growth and an increase in (haematological) cancer incidences were drivers for this trend. In addition, CAR T-cell therapies were approved only for third or later treatment lines. The aim of several RCTs was to determine the use of this therapy in second-line treatment.[489] If this proves to be effective, the eligible patient population might once more increase drastically. Consequently, the healthcare costs associated with CAR T-cell therapies will again increase.

These findings indicate that the trend of increasing costs of cancer treatments (introduced in the **Chapter 1**), is unlikely to be a temporary phenomenon. Indeed, novel cancer treatments such a CAR T-cell therapies showed to demand prices that were unheard of a couple of years before their introduction. With an extension of indication and a growing cancer incidence, healthcare systems might soon not able to afford any novel treatments. Paired with unequal access to novel treatments that lead to tremendous losses of life years in itself,[490] this trend can have devastating effects on the health of populations.

Decision makers have several tools at their disposal to control prices of medical treatments such as value-based pricing through HTA. This dissertation identified and addressed some of the current challenges researchers and reimbursement decision makers face when conducting or interpreting economic evaluations as part of formal HTA. Ultimately, it is in the hand of the decision makers whether or not economic evaluations should be used to decide on the reimbursement of novel and expensive treatments in haemato-oncology.

CONCLUDING REMARKS & RECOMMENDATIONS

The aim of this dissertation was threefold. First, challenges in the evidence synthesis for HTA were explored and addressed. Second, the cost-utility of novel and expensive treatments in the field of haemato-oncology was assessed. Third, several implications of CUAs on the healthcare decision-making were discussed.

More specifically, in the first part of this dissertation the current challenge of finding previously published economic evaluations was addressed by proposing a practical guidance aiding in systematically searching biomedical databases for such studies. In addition, several challenges were identified in synthesising evidence on effects of clinical studies, and evidence on costs from financial claims databases. The second part presented missing evidence on the cost-utility of two novel treatments in haemato-oncology and discussed the inclusion of future non-medical costs in a societal perspective of CUAs and the choice of health state utilities for these analyses. Finally, the third part placed the findings of the two CUAs of the second part into the context of healthcare decision making and provided an outlook on future healthcare expenditure in the field of haemato-oncology.

The findings of this dissertation lead to several implications and recommendations that are listed and explained below. It needs to be noted that the identified challenges in this dissertation do not represent an exhaustive list of all possible challenges in HTA. Nor are the here examined domains the sole domains of HTA. The following implications and recommendations should therefore be read in this context. In addition, their order does not reflect preference.

One, future research should aim at enhancing and updating current guidelines to conduct systematic literature reviews of economic evaluations, taking into account improved and simplified user interfaces of biomedical databases. Simultaneously, novel approaches such as a machine learning algorithm should be studied for their suitability to automate (parts of) such reviews. In case such tools are deemed appropriate their use should actively be recommended in up-to-date guidelines.

Two, future research needs to validate promising novel modelling techniques to extrapolate short-term empirical survival of clinical studies to a longer time horizon while considering external long-term data.

Three, results from cost-utility analyses that were based on efficacy data from phase II clinical studies need to be validated with long-term efficacy data from pivotal trials. Furthermore, clinical studies to collect the needed data should become mandatory when reimbursement decisions are based in phase II studies to reduce long-term uncertainties of the reimbursement decision.

Four, future research should determine strengths and weaknesses of synthesising evidence on resource use from financial claims databases compared to patient questionnaires. In addi-

tion, such research should elucidate the possibilities to enhance financial claims databases in a way so that they are better suitable for scientific research.

Five, national pharmacoeconomic guidelines should be updated to specifically recommend or discourage the inclusion of future non-medical costs in CUAs. It seems that the literature has a preference for including such costs since only then estimates from a societal perspective are complete. However, a lack in recommendations from said guidelines did not encourage their inclusion thus far. Clear recommendations such as in the US would help to generate comprehensive results from a societal perspective while ensuring a relative degree of comparability between studies that include such costs.

Six, future research should establish reliable methods to estimate potential future productivity gains for children and young adult cancer survivors. This is because CUAs typically consider future effects (e.g. long-term survival and HRQoL). Consequently, future costs such as those related to productivity, should be considered as well. While standard methodological approaches to include productivity losses are well documented in the literature, similar techniques are lacking to estimate productivity gains of young cancer survivors. This is problematic because other than adult patients, paediatric and young adult patients do not usually incur productivity losses due to their age. Hence future cost estimates in this young group of patients are often incomplete.

Seven, further research is needed to determine reliable estimates of health state utility values for patients with haematological malignancies. A systematic literature review in this field could help to summarise available evidence and to identify gaps.

Eight, the use of budget impact analyses to guide reimbursement decision in healthcare, especially in the field of haemato-oncology should be reconsidered. Other than budget impact analyses, economic evaluations can address a broad range of uncertainty and should therefore be seen as the superior methodology to address the reimbursement decision problem.

In essence, this dissertation adds to the growing body of literature aiming at updating and enhancing (methods of) economic evaluations through identifying and addressing challenges in HTA, to make them an even more robust and reliable tool in assessing the value of novel treatments.