Broadening Health Technology Assessment

To support setting boundaries to the basic benefit package

Joost Enzing

ISBN: 978949042066-6
Cover by Blastboys
Printed by Gildeprint

Broadening Health Technology Assessment

To support setting boundaries to the basic benefit package

Verbreding van Health Technology Assessment

Om begrenzing van het basis pakket te ondersteunen

Proefschrift

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

Prof.dr. A.L. Bredenoord

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

vrijdag 22 september 2023 om 10:30 uur

door

Joost Jan Enzing
geboren te 's-Gravenhage.

Erasmus University Rotterdam

Promotiecommissie:

Promotoren:

Prof.dr. W.B.F. Brouwer

Prof.dr. A. Boer

Overige leden:

Prof.dr. D.M.J. Delnoij

Prof.dr. G.A. de Wit

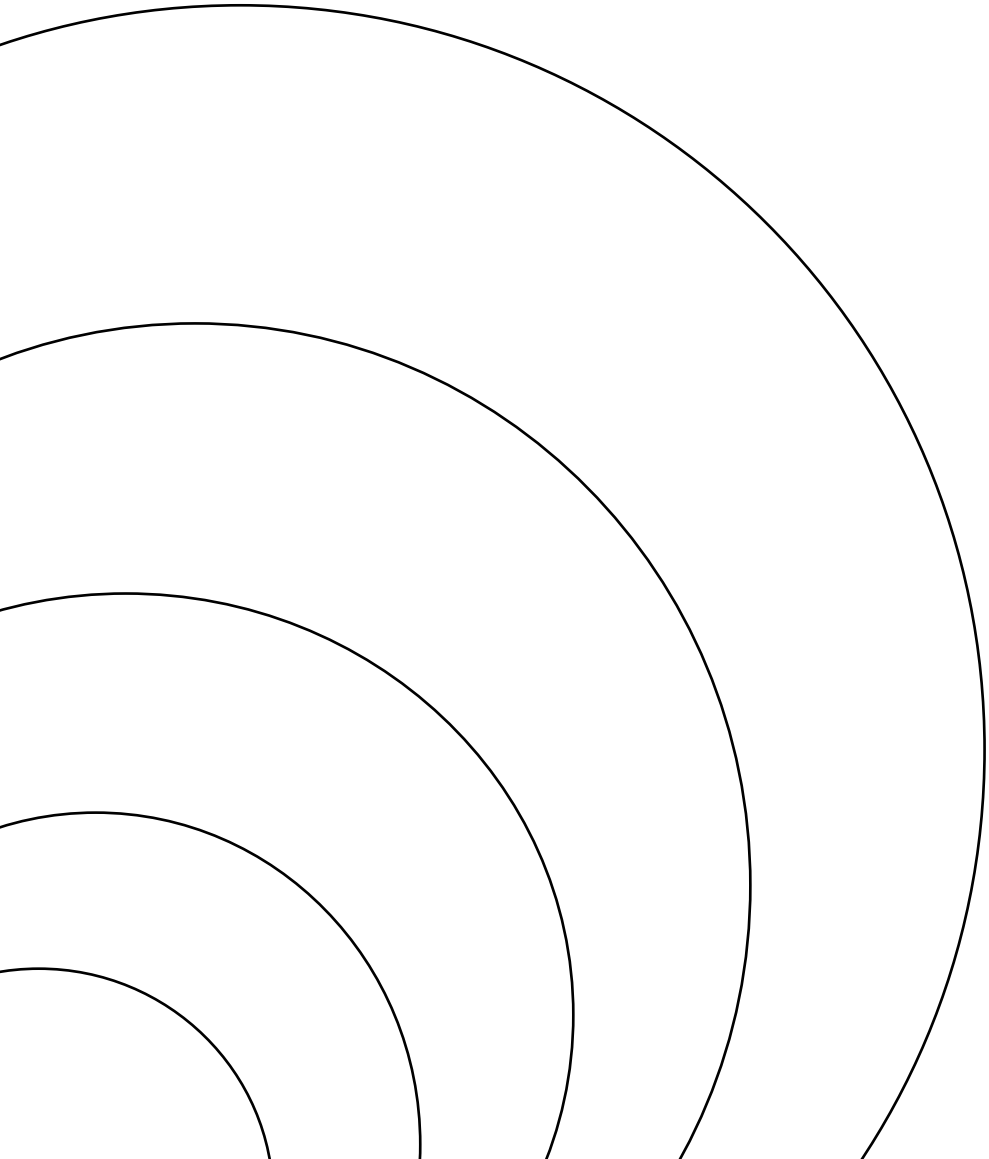
Prof.dr. R.M.P.M. Baltussen

Copromotor:

Dr. S. Knies

Table of contents

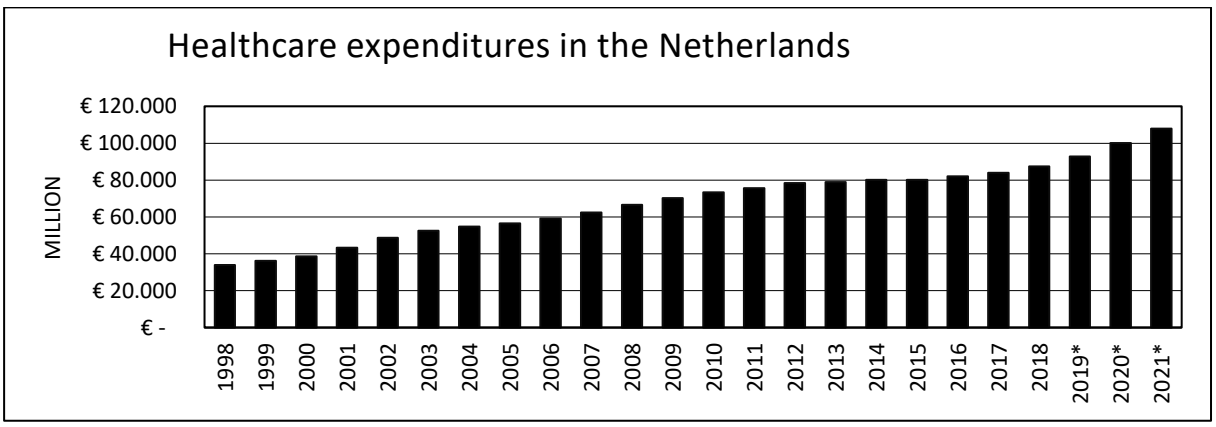
Chapter 1
Introduction



1.1 Introduction

In the Netherlands, all citizens are legally obliged to take out health insurance, covering the costs of an elaborate set of healthcare interventions, known as the basic benefit package (BBP). This BBP includes both curative care and long-term care, that is, care covered under the Health Insurance Act (Zorgverzekeringswet; Zvw) or the Long-Term Care Act (Wet Langdurende Zorg; WLZ) respectively. The Dutch government strives to ensure that this care is of good quality, that it is accessible to those who need it, and that it is affordable for both patients and society. The latter remains a continuous challenge, considering the almost continuously increasing total expenditures on healthcare in the Netherlands, as shown in Figure 1.1 Importantly driven by technological developments, increased wealth and expectations, and demographics [1], these expenditures exceeded €107 billion in 2021 and grew to make up 11.2% of the gross domestic product of the Netherlands [2]. This growing share of total wealth spent on healthcare, questions the sustainable affordability of healthcare and stresses the need to justify the expenditures – as they represent opportunity costs also in other societal sectors such as education or social security. Controlling healthcare expenditure is called for, therefore. [3] [4]

Figure 1.1 Dutch healthcare expenditures on curative and long-term care 1998-2021 (source: CBS) (=preliminary figures)*



There are multiple policy instruments available to control healthcare expenditures. [5] These instruments include those with an exclusive or primary focus on limiting expenditures, which comes at the risk of having unwanted effects on the quality and accessibility of care and, indeed, ultimately, the health and welfare of the population. Another way is to set deliberate boundaries to the content of the BBP, which may help to control expenditures while considering the overall aims of the healthcare sector rather than solely focusing on affordability. Setting such boundaries, as well as procedures to safeguard them, allows the explicit balancing of costs and benefits of healthcare interventions that are or could be part of the BBP, hence balancing affordability with quality and accessibility. Ideally this provides the opportunity to safeguard that insured care optimally contributes to the public goals for and public values within healthcare. For example, by deciding on whether or not to include a surgical procedure into the BBP, aspects like its relative effectiveness, costs, and the severity of the treated disease can be explicitly considered. This policy instrument of setting boundaries to the content of the BBP based on criteria in line with the overall healthcare goals, is used in many jurisdictions with collectively financed healthcare, such as England, Australia, Canada and the Netherlands. To support and structure the decisions necessary to set boundaries to the content of their BBP, decision-making frameworks have been developed and implemented within these jurisdictions. These frameworks are based on a process typically known as health technology assessment (HTA). The international HTA community recently defined HTA as:

"*a multidisciplinary process that uses explicit methods to determine the value of a health technology at different points in its lifecycle. The purpose is to inform decision-making in order to promote an equitable, efficient, and high-quality health system.*" [6]

HTA and the decision-making frameworks based upon it are well-established, especially in the domain of pharmaceuticals, but not necessarily complete or optimal. Some of the common decision-criteria may require attention in their operationalization (e.g. equity [7] and efficiency [8]), and the set of considered criteria may be broadened (e.g. by adding environmental aspects [9] or price fairness [10]). Furthermore, when considering the optimization of these frameworks, and their relatively clear methodological approaches, it should be underlined that their use is currently mostly limited to (new) pharmaceuticals. For other types of care, a systematic consideration of aspects like their effectiveness, cost-effectiveness and severity of disease before they are included in the covered BBP is mostly lacking. Rather, these types of care enter the healthcare system and the BBP without an explicit evaluation. As expenditures on pharmaceuticals only represent a relative small part of total healthcare expenditures, a more systematic consideration of other healthcare interventions could help to control expenditures and to ensure that care included in the BBP sufficiently contributes to the goals of the healthcare sector. This will require a broader use of HTA and is an ambition in jurisdictions as the Netherlands.[11] However, current frameworks may not be optimal for use in these other healthcare interventions.

This thesis has as its central aim to contribute to the optimization of the decision-making framework for making decisions on including healthcare interventions into the BBP based on explicitly stated criteria, as well as of the underlying HTA methodology. Therefore, two important research areas will be covered. The first concerns investigating the broadening of the *scope* of HTA by adding "profitability to the manufacturer" as a considered criterion. The second concerns investigating broadening of the *use* of HTA, in the context of healthcare interventions other than pharmaceuticals.

1.2 Health Technology Assessment in The Netherlands

In the Netherlands, setting boundaries to the content of the BBP is the ultimate responsibility of the Minister of Health, Welfare and Sport (MoH). However, for this task the MoH is advised by the National Health Care Institute (Zorginstituut Nederland or ZIN). In practice, this mostly concerns advising decisions regarding the inclusion or exclusion of interventions into or from the BBP; most often relating to curative care covered under the Health Insurance Act and less often care covered under the Long-Term Care Act.

Especially when a 'negative decision' is reached, i.e., a decision to not include a new technology in the BBP or to exclude a currently covered intervention from the BBP, this may cause result in a strong public and political opposition and a call for accountability. [12] [13] This is often related to the fact that not being included in the BBP typically leads to a situation that an intervention will not be available, accessible or affordable for most patients. To prepare their advice on these matters to the MoH in an accountable way, ZIN uses the principles of HTA in a standardized way, both in terms of the content of what is evaluated as well as in terms of the process to reach a final advice. In terms of process, for ZIN this involves selecting healthcare interventions for assessment, subsequently systematically assessing (i.e., mapping out) the relevant characteristics of the selected interventions (the "assessment phase"; also based on input of third parties such as professionals or companies) as highlighted below, appraising the results of this assessment (the "appraisal phase"), and formulating final recommendations to the MoH. These recommendations may concern the inclusion or exclusion of an intervention into the BBP, but can also highlight associated conditions regarding inclusion. Together with the results of the assessment and the appraisal these recommendations are made public in a reimbursement report.

1.3 Assessment criteria

In the process of HTA, the assessment phase is a highly important one. There, evidence is gathered and evaluated, in order to determine whether or not an intervention meets the criteria that need to be fulfilled in order to be allowed into the BBP. Among organisations applying HTA, multiple sets of assessment criteria are used, although there appears to be some practical consensus that the criteria safety, effectiveness, cost-effectiveness, and budget impact, are pivotal. [14] In the Netherlands, the four main assessment criteria are: necessity, effectiveness, cost-effectiveness and feasibility. [15] These criteria, which ideally should be fulfilled by all care covered in the BBP, can be traced back to a landmark publication on the issue of delineating the BBP. [16] ZIN has applied and further refined these criteria over the years and uses them in the assessment of curative interventions.

The criterion of **necessity** consists of two sub-criteria: the financial necessity of insurance coverage, and the medical necessity of the treatment. The financial necessity of insurance coverage importantly depends on the costs of the intervention and whether an individual can bear these costs. The medical necessity of treatment importantly depends on the severity of the disease. This severity or burden of disease can be quantified in different ways. [17] ZIN first calculates the so-called absolute shortfall: the number of Quality Adjusted Life Years (QALYs – a measure that combines length and health-related quality of life) lost due to the disease. This loss may occur due to a reduction in quality of life due to the disease, a reduction in length of life due to the disease, or both. Secondly, ZIN divides the absolute shortfall by the number of QALYs that a comparable person would experience without the disease. Hence, the proportion of otherwise lived health lost due to the disease is calculated, which is known as proportional shortfall. [18] The latter plays an important role in the Dutch

decision making framework, with higher levels of proportional shortfall being associated with higher reference values (i.e., thresholds) for incremental cost-effectiveness ratios. [19]

The criterion of **effectiveness** focuses on the extent to which the intervention does what it intends to do, in the context of the disease in question. The main question to be answered is whether the intervention, given its desired and undesired consequences for the designated patient group, has (added) value compared to the usual treatment. This criterion is formulated as that an intervention should be ”established medical science and medical practice“. To determine effectiveness, ZIN uses the principles of Evidence Based Medicine (EBM), and specifically the GRADE system. This system makes it possible to weigh different types of evidence, from expert opinion to evidence from RCTs, systematically and in a hierarchy. In contrast to the other three criteria, the criterion of effectiveness also has a legal status. [20]

The third criterion of **cost-effectiveness** (or “value-for-money”) establishes whether the balance between costs and benefits (including health effects) of an intervention is acceptable. To assess this an economic evaluation can be performed. In the context of healthcare interventions, these economic evaluations typically take the form of a cost-utility analysis (CUA). A CUA relates the incremental health benefits, expressed in terms of QALYs, that a healthcare intervention brings to the additional incremental costs, both compared to an relevant comparator (such as usual care). This results in an incremental cost-effectiveness ratio (ICER), which highlights how many euros one has to spend additionally to produce one additional unit of QALY. This ICER thus highlights a ‘price per QALY’ one pays by using the new intervention. A necessary subsequent step is to determine whether this price is acceptable, i.e., whether the ICER is not too high. In order to

do that the calculated ICER is compared to a relevant reference value. [21] As indicated above, the relevant reference value is determined in relation to the burden of disease of the treated patients, expressed as proportional shortfall. This reference value can vary from €20,000/QALY for mild diseases, via €50,000/QALY for moderately severe diseases, to €80,000/QALY for severe diseases, with very low levels of severity in theory associated with a reference value of €0 per QALY. [11]

Finally, the criterion of **feasibility** focuses on whether it is feasible to include the intervention in the BBP, even when the other criteria are fulfilled. Feasibility is assessed by mapping out pragmatic issues that can hamper or promote the successful coverage and implementation of an intervention in practice. For instance, it may be explored whether society can bear the total costs of inclusion of the intervention in the BBP. For this, an estimate is made of the financial consequences of inclusion: the budget impact.
The evidence gathered on these four main assessment criteria is used during the appraisal of the intervention. While performing this appraisal, the main criteria are weighed in combination with several other aspects considered relevant for the decision, also in a societal context. [15] As examples, aspects as the availability of alternatives, orphan status of disease, patient vulnerability and palliative versus curative interventions could be involved in weighing.

1.4 Broadening the scope of HTA

The set of societally relevant criteria currently considered in the context of HTA and delineation of the BBP deserves scrutiny and may need changes or refinement. Various aspects of healthcare interventions could be considered as additional criteria to be explicitly evaluated within an HTA process. For instance, systematically collected information on the environmental impact of healthcare interventions, the labor intensity of

interventions, or their broad impact on socio-economic health inequalities may be considered next to or incorporated in the currently used criteria of necessity, effectiveness, cost-effectiveness and feasibility. In this thesis, we will focus on one specific aspect, i.e. the profitability of an intervention for its manufacturer. This topic has been discussed, also through strong normative statements on assumed high profit margins especially in the context of expensive pharmaceuticals, in public debates both within and outside ZIN. [10, 22] It also has received attention in the scientific literature, for instance in the context of theories regarding fair pricing of pharmaceuticals. [10, 23-27] Despite the attention in the public domain and scientific literature for this issue, profitability is not a criterion within the current HTA frameworks and therefore not systematically considered in decisions on inclusion to the BBP. Here, we hope to provide insights which stimulate ZIN and other HTA organisations to reflect on the current scope of their assessments. We do note that our focus on profitability does not imply that other aspects of healthcare interventions (like the ones mentioned above) are less important in any way.

1.5 Broadening the use of HTA

HTA in principle can be applied to all types of healthcare interventions, including tests, medical devices, pharmaceuticals and procedures. In practice, most cases of assessments by ZIN concern (outpatient) pharmaceuticals. This may be somewhat surprising, given the relatively limited share of expenditures on outpatient pharmaceuticals in the total healthcare expenditures (i.e., less than 5%). [28] Medical devices, surgical procedures, mental healthcare, long-term care and other types of care are assessed relatively infrequently. As a result, these interventions are not systematically evaluated in an HTA process to ensure that they meet the BBP criteria and therefore that their coverage is indeed justified given the broad goals of the healthcare system. This has been recognized as a missed

opportunity, especially in light of the continuously increasing healthcare expenditures and the perceived need to set limits. ZIN also recognizes this and has repeatedly expressed the ambition to broaden the use of HTA, going beyond pharmaceuticals to (all) other types of healthcare interventions. [11] Following the example of many, an influential Dutch advisory body (Wetenschappelijke Raad voor het Regeringsbeleid), recently emphasized that such a broadening the use of HTA should be part of the development towards a more sustainable healthcare system. [4] However, broadening the use of HTA comes with many challenges, which makes it a difficult goal to attain. This may also be derived from the fact that the Dutch situation is not unique. Indeed, the bias towards evaluating (outpatient) pharmaceuticals and the ambition to broaden the use of HTA beyond pharmaceuticals is not only observed in the Netherlands but rather is an international phenomenon. [29, 30] In several jurisdictions, steps towards a broader use of HTA are foreseen or taken. For example, the Public Health and Social Care Centre in England assesses social care interventions, while the Health Technology Expert Review Panel in Canada assesses non-drug healthcare interventions as does the European HTA network EUnetHTA. Notwithstanding these developments, the bias towards pharmaceuticals still exists and HTA in these other domains comes with distinct challenges, related to methodological issues when performing HTA in these other contexts, as well as to policy issues regarding how to organise the HTA process and how to implement outcomes.

This thesis aims to support researchers and policy makers in moving forward towards a broader use of HTA. We try to do this by providing more insight into the challenges involved in the use of HTA in non-pharmaceutical healthcare interventions, as well as by contributing to solution for some of the identified challenges. To this end, we look at the challenges of broadening the use of HTA, not only for HTA organisations like ZIN but also

for people involved in performing CUAs in these broader contexts, as well as for researchers collecting outcome data. Various challenges in this area have already been discussed and studied, and for some solutions have been put forward in the scientific literature. [31, 32] For example, when the use of HTA is broadened towards the assessment of medical devices this brings specific methodological challenges. Such issues need to be addressed in an HTA process informing decisions on inclusion to the BBP. Another example relates to the fact that it has been disputed whether generic health-related quality of life (QoL) measures, such as the EQ-5D (which is the prescribed outcome measure for CUAs of pharmaceutical interventions), are valid when the use of HTA is broadened to the assessment of interventions aimed at improving wellbeing or mental health. [33, 34] This problem can, in theory, be overcome by using alternative outcome measures, although this solution also creates new challenges (e.g., in decision making). [35, 36]

The goal to broaden the use of HTA may be ambitious, but systematically weighing the relevant criteria for *all* types of care interventions, is necessary to ensure their contribution to optimal quality, accessibility and affordability of healthcare. This thesis will study the broad challenges of performing HTA for other than pharmaceutical interventions, in different contexts and from different angles. This should provide both a general overview of the issues as well as more detailed insights in specific contexts, i.e., medical devices and mental health.

1.6 Central aim of this thesis

The central aim of this thesis is to contribute to the optimization of the decision-making framework for making decisions on including healthcare interventions into the BBP based on explicitly stated criteria, as well as of the underlying HTA methodology, by exploring a broadening of the scope and use of HTA.

To do so, we will address three main research questions:

1. Would broadening the scope of HTA with "profitability" as an assessment criterion influence decisions on the inclusion of interventions into the BBP?
2. What are challenges when broadening the use of HTA towards non-pharmaceutical healthcare interventions?
3. Can specific methodological challenges be resolved when broadening the use of HTA towards non-pharmaceutical healthcare interventions?

These questions will be addressed in three distinct parts of this thesis, each with its own focus:

Part 1. Broadening the scope of HTA with profitability as an additional assessment criterion

Chapter 2. The presence of information on profitability within current HTA reports

The aim of this chapter is to determine whether manufacturers' costs in relation to price (or "profitability"), although outside the traditional set of assessment criteria, are currently explicitly considered by HTA organisations as reflected in reimbursement reports of expensive drugs.

Chapter 3. The role of information on profitability in hypothetical reimbursement decisions

The aim of this chapter is to explore whether and to what extent actual healthcare policy makers take information on profit margins into account in hypothetical reimbursement decisions, when presented alongside common information on pharmaceutical products.

Part 2. Challenges for HTA organisations when broadening the use of HTA towards healthcare interventions other than outpatient pharmaceuticals

Chapter 4. Challenges for reimbursement decision-making in the Netherlands

The aim of this chapter is to explore important challenges that ZIN will need to address when broadening the use of HTA and the decision-making process based upon it towards healthcare interventions other than outpatient pharmaceuticals.

Part 3. Case studies regarding methodological challenges when broadening the use of HTA towards healthcare interventions other than outpatient pharmaceuticals

Chapter 5. Challenges for economic evaluations of medical devices

Broadening the use of HTA will require the assessment of interventions with characteristics deviating from those of pharmaceuticals. The aim of this chapter is to explore whether specific characteristics of medical devices are addressed in peer reviewed, full economic evaluations using trans catheter aortic valve implantation (TAVI) as a case study.

Chapter 6. Challenges for outcome measurement of mental healthcare

Broadening the use of HTA may require measuring outcomes beyond or other than health-related quality of life. The aim of this chapter is to explore the psychometric properties of an outcome measure which was specifically developed for the context of mental healthcare interventions, the Mental Health Quality of Life (MHQoL).

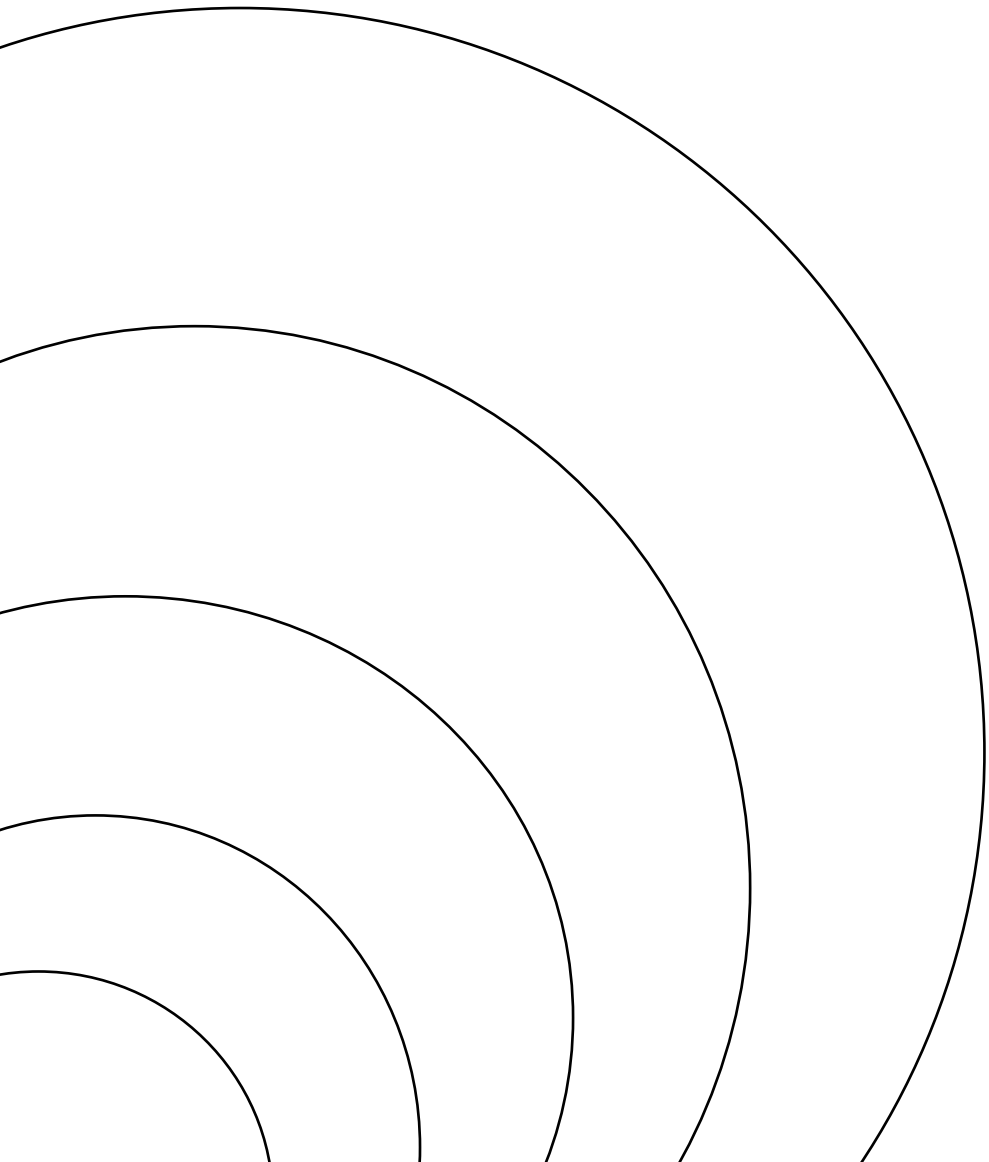
The findings presented in chapters 2 to 6 will be synthesized and discussed in chapter 7 to answer our research questions. Furthermore, in chapter 7 we reflect on the strengths and limitations of this thesis and formulate implications for HTA research and decision-making frameworks based upon it.

Finally, it needs noting that chapters 2 to 6 were written as separate scientific papers, which can therefore be read independently from each other, and may have some overlap.

Part 1
Broadening the scope of HTA with profitability as an additional assessment criterion

Chapter 2
The presence of information on profitability within current HTA reports

Based on: Enzing JJ, Knies S, Engel J, IJzerman MJ, Sander B, Vreman R, Boer B, Brouwer WBF. Do Health Technology Assessment organisations consider manufacturers' costs in relation to drug price? A study of reimbursement reports. Cost Eff Resour Alloc. 2022 Aug 31;20(1):46. doi: 10.1186/s12962-022-00383-y. PMID: 36045377; PMCID: PMC9434877.



Abstract
Drug reimbursement decisions are often made based on a price set by the manufacturer. In some cases, this price leads to public and scientific debates about whether its level can be justified in relation to its costs, including those related to research and development (R&D) and manufacturing. Such considerations could enter the decision process in collectively financed health care systems. This chapter investigates whether manufacturers' costs in relation to drug prices, or profit margins, are explicitly mentioned and considered by health technology assessment (HTA) organisations.

An analysis of reimbursement reports for cancer drugs was performed. All relevant Dutch HTA-reports, published between 2017 and 2019, were selected and matched with HTA-reports from three other jurisdictions (England, Canada, Australia). Information was extracted. Additionally, reimbursement reports for three cases of expensive non-oncolytic orphan drugs prominent in pricing debates in the Netherlands were investigated in depth to examine consideration of profit margins.

A total of 66 HTA-reports concerning 15 cancer drugs were included. None of these reports contained information on manufacturer's costs or profit margins. Some reports contained general considerations of the HTA organisation which related prices to manufacturers' costs: six contained a statement on the lack of price setting transparency, one mentioned recouping R&D costs as a potential argument to justify a high price. For the case studies, 21 HTA-reports were selected. One contained a cost-based price justification provided by the manufacturer. None of the other reports contained information on manufacturer's costs or profit margins. Six reports contained a discussion about lack of transparency. Reports from two jurisdictions contained invitations to justify high prices by demonstrating high costs.

Thus, despite the attention given to manufacturers' costs in relation to price in public debates and in the literature, this issue does not seem to get explicit systematic consideration in the reimbursement reports of expensive drugs.

2.1 Introduction

In collectively financed health systems reimbursement decisions regarding new pharmaceuticals, in a number of jurisdictions, are informed by health technology assessment reports and the result of systematic decision processes. Such reimbursement decisions regarding pharmaceuticals are often made based on a price set by the manufacturer of the drug. This price typically covers all costs relevant to the manufacturer as well as a profit margin. Often, the relative sizes of these components of the final (list) price are unclear. In some cases, the price a manufacturer sets for its product may be considered high (in absolute sense or given its effects), which can lead to public and scientific debates about whether this price is justified [25, 37]. Such debates are fuelled by the growth in new, highly priced drugs (and other technologies for that matter), leading to questions about the sustainability of current pricing and reimbursement models [38]. However, whether manufacturers' costs (including those for research and development (R&D), manufacturing, marketing and overheads) in relation to price, and therefore their profit margin, are available to HTA organisations and are explicitly considered by these organisations in current reimbursement decisions concerning drugs, to our knowledge, has not been examined.

In the context of reimbursement decisions it is important to distinguish between the assessment and the appraisal phase of the decision-making process. All available evidence, mostly clinical and economic, is collected and synthesised during the assessment phase. It is unlikely that information on manufacturers' costs or profits will be disclosed at the time of submission and assessment for reasons of confidentiality. Indeed, despite the increased attention for price setting, manufacturers' costs and profit margins are not part of the common assessment criteria considered during the HTA process. For example, in the EUnetHTA Core Model [39], a European HTA framework,

neither manufacturers' costs nor profit margins, are mentioned. These topics are also normally not covered in overviews of used or proposed decision criteria [40-42]. However, despite not being available at the assessment phase, the information may still be relevant in the appraisal phase supporting the decision-making process. In the appraisal phase, typically, a committee critically appraises the available scientific evidence but also can consider societal and ethical aspects deemed relevant in reaching a decision or making a recommendation. Prices and profit margins could be discussed and weighed in this phase, alongside other broader considerations regarding the evaluated technology, in reaching a final decision or recommendation. Moreover, the increased reliance on price negotiations in the reimbursement process (e.g. [43]) may suggest that decision makers expect that there at least *could* be room for price reductions, which in turn may suggest the expectation that it would be possible to negotiate towards an acceptable profit margin. Such negotiations can also be part of or the result of the appraisal phase. Hence, given the increased attention for price setting as well as the increased reliance on price negotiations, manufacturers' costs in relation to prices, or profit margins, could be an explicit part of the deliberations during the appraisal phase, also to justify certain decisions or recommendations. A recent discrete choice experiment among Dutch healthcare decision makers suggested that information on profit margin would influence their reimbursement recommendations when available [44].

Whether and how manufacturers' costs and profits are currently addressed in the appraisal phase of reimbursement decisions, is an important but understudied topic. How appraisal committees consider this issue may also be related to their views on price setting and the context in which a new intervention will be used. Regarding the latter, the need for active attention for and negotiations of prices may be affected by the competitiveness of the

market a technology enters into after a reimbursement decision. Regarding the former, views on 'desirable' price setting range from value-based approaches, relating prices more to added (therapeutic) value than to the costs of manufacturing the product [45] to cost-based approaches that take the manufacturers' costs as a starting point for determining 'reasonable' or 'fair' prices [10]. Implicit or explicit negotiations could help to achieve such 'reasonable' prices, also determining profit margins. In so, they determine the division of the generated surplus, i.e., the monetary difference between manufacturers' costs and maximum willingness to pay, between the manufacturer and society. Their success, however, will also depend on relative negotiating power [46].

Considerations of manufacturers' costs in relation to price and (expected) profit margins may be relevant for appraisal committees in formulating a decision or advice in relation to a specific reimbursement decision. Given the attention for and potential relevance of profits for reimbursement decisions in different contexts, this study therefore investigates, for selected jurisdictions, whether manufacturers' costs in relation to price are currently explicitly considered by HTA organisations as reflected in reimbursement reports of expensive drugs. In doing so, it is acknowledged that the phases of assessment, appraisal and price negotiations may be organised differently in different jurisdictions and not always be fully distinguishable. Such reports are publicly available and provide insight into the explicit deliberations of appraisal committees and the arguments used in this context. To our knowledge, this study is the first to address this issue. Although not all deliberations may be documented within these reports, the results may contribute to our understanding of the role of manufacturers' profits in current reimbursement decisions.

2.2 Methods

A study of HTA-reports, documents, or sets of documents reporting a reimbursement decision, was performed to investigate whether manufacturers' costs in relation to prices were explicitly considered by the HTA organisation. This study consisted of both an analysis of systematically selected cancer drugs reports and three case studies on expensive non-oncolytic orphan drugs. For pragmatic reasons, the study was limited to reports published in English or Dutch. To cover a wide geographical range but still a manageable amount of documents, HTA-organisations from four jurisdictions were selected, namely Zorginstituut Nederland (ZIN; the Netherlands), the National Institute for Health and Care Excellence (NICE; England), the Canadian Agency for Drugs and Technologies in Health (CADTH; Canada), and the Pharmaceutical Benefit Advisory Committee (PBAC) and the Medical Services Advisory Committee (MSAC) (Australia). HTA-reports were obtained from the respective websites in June 2020. These reports generally contain results from both the assessment and the appraisal phase.

The analysis of systematically selected reports was limited to decisions concerning cancer drugs, as these pharmaceuticals are generally expensive, and discussions about manufacturers' cost in relation to price may be expected to be relatively prominent for such products. As a starting point, relevant reports from ZIN, published between 2017 and 2019, were selected from the ZIN website, restricted to those containing the keyword oncology. The resulting reports were screened independently by two reviewers (JJE & JE) and included when these considered a cancer drug (excluding e.g. diagnostics). For the cancer drugs which were the subject of the included ZIN HTA-reports, corresponding HTA-reports in the other three jurisdictions were retrieved for the considered cancer drug and indication, accepting minor differences in indication or drug combinations.

This approach facilitated comparison across the four jurisdictions. For the Australian jurisdiction, where resubmissions are common, inclusion of oncology reports was limited to first submissions.

The analysis of cancer drug reports was supplemented with three in depth case studies of expensive non-oncolytic orphan drugs because of their prominence in the pricing debate in the Netherlands: lumacaftor/ivacaftor (Orkambi®) for the treatment of cystic fibrosis, eculizumab (Soliris®) for the treatment of paroxysmal nocturnal haemoglobinuria (PNH) and eculizumab (Soliris®) for the treatment of atypical haemolytic uraemic syndrome (aHUS). These cases, which represent 'extremes' in proposed prices, were purposely selected to increase the change of observing a discussion on price in relation to manufacturers' costs. Available HTA-reports on these products were retrieved for all four jurisdictions and investigated in terms of their discussion of prices in relation to costs, similarly to those on cancer drugs.

The collected reports were read independently by two reviewers (JJE & JE) who extracted data using a structured data extraction form implemented in Microsoft Excel (Microsoft, Redmond, WA). To provide insight in general characteristics of the assessed drugs and the reimbursement recommendations, additional information on cost-effectiveness, budget impact and price negotiations for the included products was collected, as well as information on the final reimbursement decision (JE, validated by JJE).

Firstly, information on the manufacturers' costs or the manufacturers' profits related to the evaluated drug was extracted. For example, this extraction would include the mentioning of specific investments needed for the development of the drug. During this extraction the reviewers used the

following broad definitions:

- Manufacturers' costs include past, present and future costs related to the product and borne by the manufacturer.

- Manufacturers' profits (or profit margins) are financial benefits for the manufacturer realized when revenues generated by the drug exceed the costs to the manufacturer.

Secondly, the reviewers extracted text fragments which contained considerations on manufacturers' costs in relation to price. Signal words used during this extraction were: price, costs, R&D, manufacturing, overhead, profits, profit margin, substantiation, fairness, fair, reward for innovation, recouping and transparency. This extraction would, for example, include discussions on the potential role of manufacturers' costs within the reimbursement decision process. Considerations relating prices to cost-effectiveness or budget impact, which may also contain the used signal words, were excluded as these considerations concern costs to the payer, and not costs to the manufacturer. Considerations solely present within external stakeholder comments included in the reports, without explicit reflection by the HTA organisation, were not included, as these were not interpreted as considerations of the HTA organisation. However, considerations presented by an HTA organisation as a result of their reviews were included. The reviewers combined their extraction results, and in case of disagreement, this was resolved by discussion with two additional authors (SK & RV).

2.3 Results

In this section we will describe the results of our analysis of cancer drugs reports and subsequently of our orphan drugs case studies.

2.3.1 Analysis of HTA-reports on cancer drugs

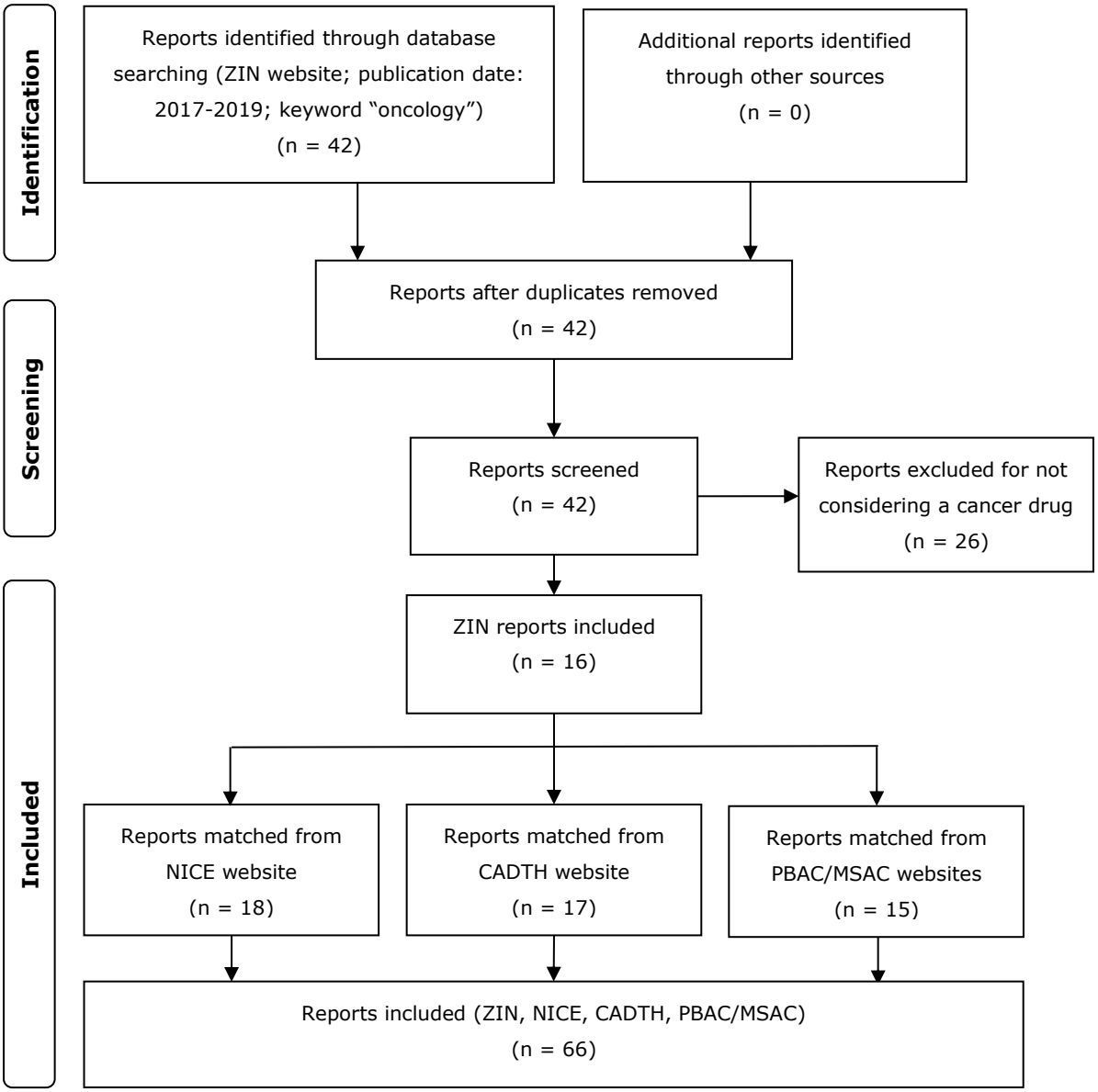
Of all relevant reports published by ZIN in the years 2017, 2018 and 2019 (n=42), 16 HTA reports of a cancer drug were included in the analysis (see Table 2.1). From the websites of the other HTA-organisations, 18 NICE reports, 17 CADTH reports and 15 PBAC/MSAC reports were retrieved. Some of the Dutch reports considered two indications for the same drug, while these indications were reported in separate documents by NICE and/or CADTH, which explains the higher number of included reports from these two institutions. Overall, 66 reports were included (see Table S1 in the supporting information for references; see Figure 2.1 for a PRISMA Flow Diagram).

Table 2.1 Included cancer drugs in the analysis and number of reports per jurisdiction

Active ingredient /generic name	**Brand name**	**Indication**	**ZIN**	**NICE**	**CADTH**	**PBAC/MSAC**
dabrafenib & trametinib	Tafinlar & mekenist	Melanoma	1	1	1	1
ipilimumab & nivolumab	Yervoy & Opdivo	Renal cell carcinoma	1	1	1	1
venetoclax & retuximab	Venclyxto & generic	Chronic lymphocytic leukaemia	1	1	1	1
Durvalumab	Imfinzi	Non-small cell lung cancer	1	1	1	1
Abemaciclib	Verzenios	Breast cancer	1	2	1	1
axicabtagene ciloleucel	Yeskarta	Large B-cell lymphoma	1	1	1	1
Tisagenlecleucel	Kymriah	Large B-cell lymphoma / Acute lymphocytic leukaemia	2	2	1	2
dinutuximab beta	Qarziba	Neuroblastoma	1	1	1	0
Osimertinib	Tagrisso	Non-small cell lung cancer	1	1	1	1
Atezolizumab	Tecentriq	Non-small cell lung cancer	1	1	1	1
Ribociclib	Kisqali	Breast cancer	1	1	1	1
Daratumumab	Darzalex	Myeloma	1	1	1	1
Cetuximab	Erbitux	Colorectal cancer	1	1	1	1
Ibrutinib	Imbruvica	Lymphocytic leukaemia	1	1	2	1
Palbociclib	Ibrance	Breast cancer	1	2	2	1

ZIN = Zorginstituut Nederland, NICE = National Institute for Health and Care Excellence, CADTH = Canadian Agency for Drugs and Technologies in Health, PBAC = Pharmaceutical Benefit scheme and medical services Advisory, MSAC = Medical Services Advisory Committee

Figure 2.1 PRISMA Flow Diagram: Inclusion of Cancer drug reports



None of the 66 reports contained information on manufacturers' costs or profit margins of the evaluated drugs. Seven reports contained a consideration which related manufacturers' costs to price: six discussed a perceived lack of pricing transparency and one mentioned recouping

development costs.

Widely used criteria in the context of reimbursement recommendations were found to be (the uncertainty around the) effectiveness, cost-effectiveness and budget impact. Often, reports included conditions under which final reimbursement could be allowed, including price reductions. Within the investigated reports, references to price negotiations were common: 10 out of 16 ZIN reports recommended to start price negotiations, 16 out of 18 NICE reports mentioned a negotiated discount, 16 out of 17 CADTH reports recommended price negotiations and 11 out of 15 PBAC/MSAC reports mentioned (proposed) special pricing arrangements.

2.3.2 The Netherlands

None of the 16 ZIN reports contained explicit information about manufacturers' costs or profit margins of the evaluated drugs. Four reports (daratumumab; palbociclib; ribociclib; atezolizumab) contained a statement regarding a perceived lack of transparency regarding price setting by the manufacturer, which was part of the concluding recommendations. It was not stated whether or how transparency was sought by the committee. One other report (venetoclax & rituximab) indicated that price negotiations were recommended, also because an increase in volume was expected and development costs could be recouped more quickly.

Four reports contained a positive recommendation for reimbursement without requiring further price negotiations. Two reports contained a negative advice, without recommending further price negotiations. The other ten reports recommended to start price negotiations (four reports) or indicated that price reductions would be a condition for reimbursement (six reports). The recommendations to start price negotiations were typically substantiated with arguments regarding uncertainty concerning

effectiveness, an unfavourable incremental cost-effectiveness ratio (ICER) and/or a high budget impact.

2.3.3 England

None of the 18 NICE reports contained information on manufacturers' costs or profit margins of the evaluated drugs. Moreover, none of the 18 reports contained considerations relating manufacturers' costs to prices, as put forward by NICE or the manufacturer.

One appraisal resulted in a negative recommendation (osimertinib), while the other appraisals resulted in recommending the product (nine for routine use, eight for use within the context of the cancer drug fund). The positive recommendations were made under the condition that confidential commercial agreements were followed, which explicitly included providing a negotiated discount in 16 reports.

2.3.4 Canada

None of the 17 CADTH reports contained information on manufacturers' costs or profit margins of the evaluated drugs. In an appendix of the ethics and implementation report of tisagenlecleucel, CADTH cited two references which both pointed to a lack of transparency in pricing of CAR T-cell therapies [47, 48]. In one of these, De Lima and colleagues [47] stated that there is a '*need for greater transparency about pricing given public investment into R&D*'. These two references [47, 48] were also present in the ethics review of axicabtagene ciloleucel, which additionally cited a recommendation of the Association of European Cancer Leagues to better explain the rationale behind the prices of CAR T-cell therapies [49]. No other considerations relating manufacturers' costs to price were found within the CADTH reports.

Of the reports, one resulted in not recommending the product (cetuximab),

all others resulted in a conditional positive advice given to the provinces. The relevant conditions in all cases involved making price-arrangements with the manufacturer and improving cost-effectiveness.

2.3.5 Australia

None of the 15 PBAC/MSAC reports, in which some data may be censored during redaction, contained information on manufacturers' costs or profit margins of the evaluated drugs, or any considerations which related manufacturers' costs to price.

Five reports recommended reimbursement in combination with some form of (price) arrangements. Ten reports either deferred making a decision (two reports) or did not recommend the reimbursement (eight reports) of the appraised drug based on the first submission. Rejections were accompanied by an invitation to resubmit the application, which is common practice in Australia, as can be illustrated by a sentence used in all included PBAC reports (e.g., osimertinib): "*A PBAC decision not to recommend listing or not to recommend changing a listing does not represent a final PBAC view about the merits of the medicine. A company can resubmit to the PBAC or seek independent review of the PBAC decision.*" It should be noted that drugs rejected for Pharmaceutical Benefits Scheme (PBS) listing may be available through self-funding or private insurance. Resubmissions may involve an adjusted price, and in that sense may be viewed as being part of a negotiation process with the HTA organisation. All reports contained a reference to (proposed) special pricing arrangements (11 reports), or to a required price reduction.

2.3.6 Orphan drugs case studies

From the websites of the four relevant HTA-organisations, 21 reports were identified concerning the (re)appraisal of eculizumab for PNH, eculizumab for aHus or lumacaftor/ivacaftor for the treatment of cystic fibrosis. One

document reported cost related arguments provided by the manufacturer to substantiate the relatively high price of the product (eculizumab aHus, NICE). This document, as well as five others (eculizumab PNH, ZIN, 2016; eculizumab PNH, ZIN, 2017; eculizumab aHus, ZIN, 2016; lumacaftor/ivacaftor, ZIN, May 2016; lumacaftor/ivacaftor, ZIN, December 2016), discussed manufacturers' costs as relevant information for the reimbursement decision (Table 2.2) (see Table S2 in the supporting information for references).

Table 2.2 Included HTA-reports related to orphan drugs case studies

Intervention (indication)	**HTA organisation**	**Publication date**	**Consideration of manufacturers' cost in relation to drug price?**
eculizumab (PNH)	PBAC	July 2008	No
	PBAC	March 2009	No
	CADTH	February 2010	No
	PBAC	July 2010	No
	ZIN	May 2016	Yes
	ZIN	June 2017	Yes
eculizumab (aHus)	PBAC	March 2013	No
	CADTH	July 2013	No
	PBAC	March 2014	No
	PBAC	August 2014	No
	NICE	January 2015	Yes
	CADTH	May 2015	No
	ZIN	November 2016	Yes
lumacaftor/ivacaftor (CF)	PBAC	March 2016	No
	ZIN	May 2016	Yes
	NICE	July 2016	No
	CADTH	October 2016	No
	PBAC	November 2016	No
	ZIN	December 2016	Yes
	PBAC	July 2017	No
	PBAC	July 2018	No

HTA = health technology assessment, PBAC = Pharmaceutical Benefit scheme and medical services Advisory Committee, CADTH = Canadian Agency for Drugs and Technologies in Health, ZIN = Zorginstituut Nederland, NICE = National Institute for Health and Care Excellence, PNH = paroxysmal nocturnal haemoglobinuria, aHus = atypical haemolytic uraemic syndrome, CF = cystic fibrosis

2.3.7 Eculizumab for treatment of PNH

The orphan drug eculizumab is used to treat patients with paroxysmal nocturnal haemoglobinuria (PNH). PNH is a life-threatening genetic disease which results in severe medical complications such as anaemia and thrombosis. In May 2016 ZIN published an HTA-report on eculizumab for PNH. This report concluded that eculizumab, at annual costs of € 360,000 per patient, was not cost-effective and ZIN recommended to suspend reimbursement and start price negotiations. Much weight was attached to the notion that the manufacturer was not transparent about the price structure. The appraisal committee indicated to take the position that rejection of reimbursement would be in order '*if the manufacturer does not take the trouble to submit an acceptable cost-effectiveness model and asks an extremely high price that, according to the ACP, is not transparent as well as being immoral*…'.

The report also mentioned that for interventions with an unfavourable cost-effectiveness, arguments may exist to justify reimbursement. Such justifications could include costs related to the development of the drug, to market access, and to production. Furthermore, a general call for transparency was included, directed to the pharmaceutical industry. This should help in making accountable public decisions regarding reimbursement. In addition, the Minister of Health (MoH) was advised to take into consideration, in the context of the process of price negotiations, that other indications for which eculizumab would be prescribed were expected. In a reassessment in 2017, ZIN reported that the requested transparency was still lacking.

No NICE report on eculizumab for the indication of PNH was found.

CADTH published a common drug review on eculizumab for PNH in 2010. In this report the Canadian Drug Expert Committee (CDEC) recommended

that eculizumab should not be listed at the price listed in the submission. The yearly cost of eculizumab per patient were labelled as exceptionally high. The CDEC noted that eculizumab had not been shown to be cost-effective, with the remark that this criterion is weighed against other criteria in making reimbursement decisions. The CDEC added: '*It has been argued that the costs of drugs to treat rare diseases are often high because of the relatively small number of patients for whom the drug is indicated*.' This sentence could imply that actual costs might be used to justify high prices, even if these would result in estimates beyond commonly used cost-effectiveness thresholds.

PBAC published a report on eculizumab for PNH in 2008. The PBAC rejected the manufacturers' submission on the basis of unacceptably high and highly uncertain costs per avoided death. Reassessment followed in 2009 when eculizumab was rejected on the basis of an unacceptably high and highly uncertain ICER. In 2010, the drug was appraised in the context of the Life Saving Drugs Program (LSDP), which provides access to expensive and potentially lifesaving drugs for very rare life-threatening conditions. In this appraisal the PBAC decided to defer the submission for eculizumab to allow the sponsor time to obtain further data about the magnitude of the survival gain. Since January 2011 eculizumab is funded through the LSDP. None of the PBAC reports mentioned manufacturers' costs or profits.

2.3.8 Eculizumab for treatment of aHus

Eculizumab is also indicated for the treatment of patients with atypical haemolytic uraemic syndrome (aHUS). Like PNH, aHus is a life-threatening genetic orphan disease, resulting in anaemia, thrombocytopenia and kidney failure. In November 2016 ZIN published their HTA-report on eculizumab for the treatment of aHus. ZIN concluded that eculizumab, at annual costs of € 478,000 per patient, was not cost-effective and recommended to only

allow reimbursement under strict conditions, including price negotiations. In their deliberations, ZIN considered that the manufacturer, although requested to do so, did not offer transparency to explain or justify the high price. Moreover, the manufacturer did not provide a sound estimate of cost-effectiveness. Similar to their eculizumab PNH reports, ZIN acknowledged that insight provided by manufacturers into their costs could potentially justify reimbursing an intervention with an unfavourable ICER.

In January 2015 NICE published their HTA-report on eculizumab for the treatment of aHus. Based on its high price, even when compared to the prices of other highly specialised technologies, NICE asked the manufacturer to provide a price justification. The response by the manufacturer, which highlighted the need to recoup development costs in a small number of patients, was deemed insufficient. An additional inquiry was made by NICE, specifically aimed at exceptional clinical or safety requirements during clinical development, costs of post-marketing research plans, and any other information to justify the proposed price. The manufacturer provided a response in which they stated that R&D costs accounted for only a small part of the additional costs for highly specialised drugs. Other elements mentioned in this context included the need to set up multiple sites for patient recruitment into clinical trials, investments in education, higher risk of failure, and reinvestment for new indications. This response did not convince NICE, as these aspects were not considered to be exclusively valid for eculizumab. Moreover, the number of patients treated with the drug was found not to be lower than that of some other highly specialised drugs. In their response, the manufacturer expressed concern that the committee was acting outside its remit with NICE's inquiries pointing towards a more cost-based price substantiation. However, the committee stated that it is within their remit to '… *also take into account what could be considered a reasonable cost for the medicine in the context*

of recouping manufacturing, research and development costs from sales to a limited number of patients.'

CADTH recommended in 2013 that eculizumab should not be listed since its clinical benefit could not be adequately established. In 2015 it confirmed this recommendation while adding the need to consider opportunity costs and healthcare system sustainability given the associated "very high cost per patient". Both reports did not contain information on manufacturers' costs or profit margins of the evaluated drug or any considerations relating price to manufacturers' costs.

In 2013 PBAC rejected eculizumab for the treatment of aHus on the basis of uncertainty regarding clinical effectiveness and an unacceptably highICER . Later (March 2014; August 2014) reassessments were published which concluded that eculizumab could be accepted through special arrangements, including a scheme of outcome-based price rebates. No comments on manufacturers' costs or profit margins were found within these reports.

2.3.9 Lumacaftor/ivacaftor for the treatment of cystic fibrosis

Lumacaftor/ivacaftor is an orphan drug used to treat patients with the inherited disease cystic fibrosis. Cystic fibrosis causes severe effects on the lungs and the digestive system of patients. In May 2016 ZIN published an initial HTA-report concerning lumacaftor/ivacaftor which concluded that it was not cost-effective. ZIN stated that acceptance of an ICER above the common threshold could be possible in cases where price setting would be transparent, but that such transparency was lacking for lumacaftor/ivacaftor. They explicitly called for more transparency regarding the price setting of lumacaftor/ivacaftor. This was also deemed important since the ICER of the drug to a considerable degree was influenced by its

price (resulting in annual costs of € 169.386 per patient) and was well above the relevant threshold. In a reassessment published seven months later ZIN concluded that the requested transparency was still not provided. Eventually, lumacaftor/ivacaftor was accepted for reimbursement after (confidential) price negotiations.

NICE published an HTA-report in July 2016 which concluded that the estimated ICERs for lumacaftor/ivacaftor were considerably higher than what is normally considered a cost-effective use of NHS resources. Similar to ZIN, NICE did not recommend reimbursement of lumacaftor/ivacaftor. Transparency regarding price setting was not discussed within the report. In 2019, NHS England announced a commercial agreement with the manufacturer that would allow access to the drug.

In 2016 CADTH also did not recommend reimbursement of lumacaftor/ivacaftor in a common drug review. This recommendation was justified by a lack of proven effectiveness. No comments on manufacturers' costs or profit margins were made.

The PBAC decided not to recommend lumacaftor/ivacaftor for Pharmaceutical Benefits Scheme (PBS) listing based on the uncertain and unfavourable ICER and the uncertain long-term effectiveness. No reference was made to manufacturers' costs or profit margins (March 2016). These were also not considered in the context of three subsequent resubmissions (November 2016; 2017; 2018). After their 2018 meeting, PBAC recommended lumacaftor/ivacaftor to be listed on the PBS, making it available via a Managed Access Program (MAP).

2.4 Discussion

This study aimed to investigate whether manufacturers' costs in relation to

prices, i.e., profit margins, are explicitly considered by HTA organisations within their reimbursement reports. A total of 66 HTA-reports on cancer drugs were studied to investigate whether information on manufacturers' costs or profit margin of the evaluated drug were included, and to see whether general considerations were included which relate manufacturers' costs to proposed prices. In addition, three extreme cases in the area of highly expensive orphan drugs were studied. In total, 21 HTA-reports on these three pharmaceutical products were investigated. In general, information on manufacturers' costs and profit margins of the evaluated drugs was not presented in the reports. Only one report contained (non-quantitative) information on manufacturers' costs (eculizumab aHus, NICE, 2015). This information was provided by the manufacturer as part of a price justification. The justification however did not convince the appraisal committee. In 13 of the 87 reports, general considerations relating manufacturers' costs to prices were provided by the HTA organisation, mostly in the form of statements on the (undesirable) lack of transparency on price setting. Requests for more transparency were not honoured.

The results indicate that information on manufacturers' costs in relation to prices is typically lacking and typically does not seem to be requested. Reflections of HTA-organisations on the relationship between manufacturers' costs and prices are rare and, if present, typically very general. At the same time, the instrument of price negotiations was recommended and used, explicitly or implicitly through resubmissions, quite common. This appears to signal the more general idea that proposed prices (and hence implied profits) could be lowered by manufacturers.

This study was the first, to our knowledge, to investigate this topic. That also implies that we cannot relate our results to previous findings in the literature. However, our findings do appear to align with a recent review of

methodological guidelines of 24 HTA-organisations, including the four jurisdictions we investigated. Manufacturers' costs nor profit margins were reported as criteria used in priority setting [14].

2.4.1 Limitations

We acknowledge several limitations of our study. First, this study is based on publicly available, partly censored, reports, which may not report all deliberations of the involved committees during their meetings, especially in jurisdictions which limit their publication to a summary.

Second, our study was focused on cancer drugs as well as purposely selected cases of orphan drugs. This sample of reports may not be representative for reports in general. One might expect considerations related to manufacturers' costs in relation to price to be relatively frequently present in our sample, which would make our findings in terms of attention for this topic (albeit rare) an overestimation of the attention for this issue in general.

Third, although we found little explicit attention for manufacturers' costs in relation to prices and profitability, we cannot exclude the possibility that this attention was present more implicitly through the more common criteria of cost-effectiveness and budget impact. For example, advising price reductions or negotiations, based on an unfavourable cost-effectiveness will (ceteris paribus) lead to lower profit rates. Moreover, when a high budget impact was used to provide a cue for price negotiations, this may be framed as a consideration of affordability, but could also relate to the assumption that fixed manufacturers' cost are recouped after a certain overall revenue. A related example of this type of influence is NICE's policy to use a higher threshold for some orphan drugs than for non-orphan drugs [50]. This may relate to the limited number of patients available to recoup investment

costs. In that sense, manufacturers' costs in relation to prices may play a larger role than observed. Such implicit considerations may also be enforced by the fact that decision frameworks typically do not use profitability (in some form) as criterion. Our study focused only on explicit information and consideration of this issue, which therefore is an important limitation.

Fourth, we had few reports with final unconditional negative decisions in our sample. In such cases one might expect profitability to more frequently play a role in justifying such a decision. This could be an interesting avenue for future research.

When interpreting our findings across jurisdictions it should be acknowledged that differences in reporting may complicate comparisons. For example, PBAC only publishes summaries of their appraisals, while in other jurisdictions more extensive HTA-reports are published. Additionally, important differences exist in the process of appraisal in relation to price negotiations, which also hamper international comparisons in this context. For example, in England, price negotiations may take place during the assessment and appraisal process. Therefore, part of the NICE HTA-reports were able to already take into account the final negotiation results (e.g. a lowered price, Managed Entry Agreements). It could be argued that in these cases, a public consideration of manufacturers' costs in relation to the proposed or negotiated price may not be that relevant or even desirable (also for the HTA agency). Similarly, in Australia, negotiations may take the form of resubmissions with a reduced price offer, allowing 'negotiation outcomes' through this process to be part of the final appraisal, although not necessarily of the preceding reports. In Canada and the Netherlands such negotiations follow after an HTA-report has been published and serve as input for the negotiations. Hence, for instance in a number of Dutch HTA-reports, subsequent price negotiations are recommended and the required

price reduction sometimes is even quantified [51]. In such a context, the emphasis on manufacturers' costs in relation to the proposed price might be expected to be larger. Additionally, the emphasis on profitability during reimbursement decision making may depend on pricing policies in place within a specific jurisdiction. In England for example, the growth of the medicines budget is capped by 2% per annum through an agreement between the pharmaceutical industry and the government (the 'Voluntary scheme for branded medicines pricing and access'). As a so-called portfolio-wide profit control scheme, this agreement may limit the need for considerations on profitability at the individual product level. The use of external reference pricing, in place for pharmaceuticals in for example the Netherlands, may also limit this need.

2.4.2 Interpretation and implications of the results

The investigated HTA-reports typically did not contain any information on manufacturers' costs or on the profit margin of the assessed drug. As a consequence, recommendations, including those concerning the start of a price negotiation, were not substantiated with an explicit weighing of manufacturers' costs or profit margins. In other words, the reports lacked the information required to take an informed, cost-based, or partly cost-based, approach to the appraisal. In a limited number of reports we found indications that such a partly cost-based approach implicitly is used in considering the desirability of reimbursing a particular drug. While such considerations may play a larger role in practice through the adopted processes, the use of price negotiations or through the use of other criteria (e.g. cost-effectiveness, budget impact), at least in the studied HTA-reports these considerations were not systematically mentioned.

To emphasize the way in which these issues may still enter the debate and HTA-reports, we highlight the NICE dinutuximab beta appraisal consultation

document. This document contained public comments which featured manufacturers' costs and profit margins (NICE, 2018). In that document, a carer urged the manufacturer '*to be as transparent as possible in laying out the basis on which it has developed its pricing to show it is offering the drug as cheaply as possible while meeting its obligations to shareholders*'. An NHS professional stated that '… *developing a new drug for an orphan indication such as high risk NB is never going to be profitable for a pharmaceutical company, particularly a relatively small one like EusaPharma unless it is priced above what NICE would normally consider cost-effective*'. Another NHS professional commented '… *drug development is never cheap. The costs can be recouped with relatively narrow profit margin if a drug has a market of tens of thousands of patients. If companies are squeezed too hard, then they will be disincentivised from researching drugs for rare conditions* …'. NICE replied that it noted these comments. While it is not clear how NICE considered them, they do highlight that HTA-organisations likely are aware of such considerations, even if they are not explicitly elaborated on in their reports.

In a general sense therefore, perceived manufacturers' costs in relation to proposed prices may at least play an implicit role, for instance in the decision to recommend to start price negotiations. However, this was not explicitly stated in the reports, which in general substantiate the need for price negotiations by pointing to insufficient cost-effectiveness. Invitations to manufacturers to provide a cost-based pricing substantiation to justify initially asked prices were observed in some cases, especially when the proposed price resulted in an ICER exceeding the relevant threshold. This again highlights the relevance of profitability, but also may suggest that exceeding common ICER thresholds could be considered acceptable for products with a limited profit margin.

While in the current situation price negotiations are typically advised based on common HTA criteria such as cost-effectiveness, policy makers could explore the desirability of starting price negotiations based on an expected large profit margin, also when the ICER does not exceed the threshold. This was not observed in any of the HTA-reports in our study. Starting price negotiations for cost-effective products with high profit margins could meet some concerns relating to high prices. Obviously, such decisions would normally be informed by information that is currently lacking: exact cost information. Moreover, in deciding on starting a price negotiation other aspects which influence expected negotiation outcomes, e.g. market position, patents and available alternative products, would also be relevant.

New price models that could guide such price negotiations have been proposed in the literature, all with their own advantages and disadvantages. Some of these models explicitly distinguish between manufacturers' costs and profits in price setting, also in evaluating a new drug. For example, Berdud et al. describe a method to establish a "reasonable" price for orphan drugs based on the costs of conducting research and the size of the patient population [24]. Balderrama et al. propose a model in which a drug price can be labelled "justifiable" or "unjustifiable" by considering the costs of its development and manufacturing [23]. Uyl et al. [27] propose to estimate reasonable prices for new cancer drugs by estimating manufacturers' total average unit costs for a pharmaceutical to which a relevant and acceptable profit margin could then be added. In their approach, this profit margin could be based on the anticipated clinical benefit, leading to a more intermediate position between fully cost-based and value-based prices. As a final example, van den Berg et al. [52] present a cost-based approach to calculate a fair price, specifically for a repurposed orphan drug.

Within the context of high drug prices, the public contributions to drug

development are a topic of interest in the literature.[53] Public funding of drug R&D is shown to be substantial in specific cases e.g. [54], and it is claimed that one to two-thirds of total upfront R&D costs are funded by taxpayers or charitable donations.[55] When manufacturers' costs are related to drug price, this public funding should be acknowledged and transparency should be provided, also to prevent governments to "pay twice".

If it is considered desirable to broaden the HTA process to more systematically and explicitly consider profitability, actively requiring information on manufacturers' costs seems necessary. In that context it is interesting to mention a recent French law which requires pharmaceutical companies to disclose the amounts of public investments in R&D for specific new drugs entering their market [56]. This information is then allowed to be used by the responsible government body (Comité économique des produits de santé, CEPS) during price negotiations with pharmaceutical companies. Note that legal amendments to enforce more extensive disclosures about manufacturing costs, for example on costs of active ingredients and profits, were not approved by the French parliament, emphasising the difficulty in obtaining and using such information in practice.

2.5 Conclusion

Despite the attention given to manufacturers' costs in relation to price within the literature and in public debates, and although they appear to have been considered relevant in some decisions, profitability levels do not seem to receive systematic explicit attention in HTA-reports for expensive drugs.

2.6 Supporting information

S1 Table. References to the HTA-reports on cancer drugs included in the analysis (n=66)

Active ingredient /generic name	HTA organisation	Publication year	Title, URL
abemaciclib	CADTH	2019	"Abemaciclib for advanced or metastatic Breast Cancer (PC0161-000)", https://www.cadth.ca/abemaciclib-advanced-or-metastatic-breast-cancer-details
abemaciclib	NICE	2019	"Abemaciclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta563
abemaciclib	NICE	2019	"Abemaciclib with fulvestrant for treating hormone receptor-positive, HER2-negative advanced breast cancer after endocrine therapy - Technology appraisal guidance", www.nice.org.uk/guidance/ta579
abemaciclib	PBAC	2019	"Public Summary Document – March 2019 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2019-03/files/abemaciclib-psd-march-2019.pdf
abemaciclib	ZIN	2019	"Pakketadvies abemaciclib (Verzenios®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/03/07/pakketadvies-sluisgeneesmiddel-abemaciclib-verzenios
atezolizumab	CADTH	2018	"Tecentriq for Non-Small Cell Lung Cancer (PC0115-000)", https://www.cadth.ca/tecentriq-non-small-cell-lung-cancer-details
atezolizumab	NICE	2018	"Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy - Technology appraisal guidance", www.nice.org.uk/guidance/ta520
atezolizumab	PBAC	2017	"Public Summary Document – November 2017 PBAC Meeting ", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-11/files/atezolizumab-psd-november-2017.pdf
atezolizumab	ZIN	2018	"Pakketadvies atezolizumab (Tecentriq®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2018/02/07/pakketadvies-atezolizumab-tecentriq-bij-gemetastaseerde-niet-kleincellige-longkanker
axicabtagene ciloleucel	CADTH	2019	"Axicabtagene Ciloleucel for Adults With Relapsed or Refractory Large B-cell Lymphoma", https://www.cadth.ca/axicabtagene-ciloleucel-adults-relapsed-or-refractory-large-b-cell-lymphoma
axicabtagene ciloleucel	NICE	2019	"Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies - Technology appraisal guidance", www.nice.org.uk/guidance/ta559
axicabtagene ciloleucel	MSAC	2020	"Public Summary Document Application No. 1587 – Axicabtagene ciloleucel (CAR-T therapy) for the treatment of refractory or relapsed CD19-positive lymphoma", http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1587-public
axicabtagene ciloleucel	ZIN	2019	"Pakketadvies axicabtagene ciloleucel (Yescarta®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/03/07/pakketadvies-sluisgeneesmiddel-axicabtagene-ciloleucel-yescarta
cetuximab	CADTH	2014	"Erbitux for Metastatic Colorectal Cancer (PC0031-000)", https://www.cadth.ca/erbitux-metastatic-colorectal-cancer-details

Active ingredient /generic name	HTA organisation	Publication year	Title, URL
cetuximab	NICE	2017	"Cetuximab and panitumumab for previously untreated metastatic colorectal cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta439
cetuximab	PBAC	2014	"Public Summary Document – November 2014 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2014-11/files/cetuximab-psd-11-2014.pdf
cetuximab	ZIN	2017	"Standpunt cetuximab (Erbitux®) bij gemetastaseerd coloncarcinoom (herbeoordeling)", https://www.zorginstituutnederland.nl/publicaties/standpunten/2017/08/10/cetuximab-erbitux-bij-gemetastaseerd-coloncarcinoom-herbeoordeling
dabrafenib & trametinib	CADTH	2019	"Tafinlar & Mekinist in combo Melanoma Adjuvant Therapy (PC0152-000)", https://www.cadth.ca/tafinlar-mekinist-combo-melanoma-adjuvant-therapy-details
dabrafenib & trametinib	NICE	2018	"Dabrafenib with trametinib for adjuvant treatment of resected BRAF V600 mutation-positive melanoma - Technology appraisal guidance", www.nice.org.uk/guidance/ta544
dabrafenib & trametinib	PBAC	2019	"Public Summary Document – March 2019 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2019-03/files/dabrafenib-and-trametinib-psd-march-2019.pdf
dabrafenib & trametinib	ZIN	2019	"Pakketadvies dabrafenib/trametinib (Tafinlar®/Mekinist®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/08/28/pakketadvies-sluisgeneesmiddelen-dabrafenib-in-combinatie-met-trametinib-tafinlar-en-mekinist-bij-de-adjuvante-behandeling-van-volwassen-patienten-met-stadium-iii-melanoom
daratumumab	CADTH	2017	"Darzalex for Multiple Myeloma (second-line or beyond) (PC0104-000)", https://www.cadth.ca/darzalex-multiple-myeloma-second-line-or-beyond-details
daratumumab	NICE	2019	"Daratumumab with bortezomib and dexamethasone for previously treated multiple myeloma - Technology appraisal guidance", www.nice.org.uk/guidance/ta573
daratumumab	PBAC	2017	"Public Summary Document – November 2017 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-11/files/daratumumab-psd-november-2017.pdf
daratumumab	ZIN	2017	"Pakketadvies daratumumab (Darzalex)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2017/11/gvs-advies-oordruppels-bij-gehoorgangontsteking-otitis-externa/pakketadvies-daratumumab-darzalex-bij-multipel-myeloom
dinutuximab	CADTH	2019	"Unituxin for Neuroblastoma (PC0154-000)", https://www.cadth.ca/unituxin-neuroblastoma-details
dinutuximab beta	NICE	2018	"Dinutuximab beta for treating high-risk neuroblastoma [ID910] – Single technology appraisal", https://www.nice.org.uk/guidance/ta538/documents/committee-papers-2
dinutuximab beta	ZIN	2018	"Standpunt dinutuximab bèta (Qarziba®) bij de behandeling van hoog-risico neuroblastoom bij patiënten van 12 maanden en ouder", https://www.zorginstituutnederland.nl/publicaties/standpunte

Active ingredient /generic name	HTA organisation	Publication year	Title, URL
			n/2018/12/12/standpunt-dinutuximab-beta-qarziba-bij-de-behandeling-van-hoog-risico-neuroblastoom-bij-patienten-van-12-maanden-en-ouder
durvalumab	CADTH	2019	"Imfinzi for Non-Small Cell Lung Cancer (PC0131-000)", https://www.cadth.ca/imfinzi-non-small-cell-lung-cancer-details
durvalumab	NICE	2019	"Durvalumab for treating locally advanced unresectable non-small-cell lung cancer after platinum-based chemoradiation - Technology appraisal guidance", www.nice.org.uk/guidance/ta578
durvalumab	PBAC	2018	"Public Summary Document – November 2018 PBAC meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2018-11/files/durvalumab-psd-november-2018.pdf
durvalumab	ZIN	2019	"Pakketadvies durvalumab (Imfinzi®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/04/01/pakketadvies-sluisgeneesmiddel-durvalumab-imfinzi-voor-volwassenen-met-lokaal-gevorderd-irresectabel-niet-kleincellig-longcarcinoom
ibrutinib	CADTH	2015	"Imbruvica for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma (previously treated) (PC0043-000)", https://www.cadth.ca/imbruvica-chronic-lymphocytic-leukemiasmall-lymphocytic-lymphoma-previously-treated-details
ibrutinib	CADTH	2016	"Imbruvica for Waldenström's Macroglobulinemia (PC0082-000)", https://www.cadth.ca/imbruvica-waldenstroms-macroglobulinemia-details
ibrutinib	NICE	2017	"Ibrutinib for previously treated chronic lymphocytic leukaemia and untreated chronic lymphocytic leukaemia with 17p deletion or TP53 mutation - Technology appraisal guidance", www.nice.org.uk/guidance/ta429
ibrutinib	PBAC	2017	"Public Summary Document – November 2017 PBAC Meeting ", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-11/files/ibrutinib-mcl-psd-november-2017.pdf
ibrutinib	ZIN	2017	"Pakketadvies ibrutinib (Imbruvica®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2017/06/08/pakketadvies-ibrutinib-imbruvica-bij-de-eerstelijnsbehandeling-chronische-lymfatische-leukemie-cll-bij-patienten-die-geen-del17p-of-tp53-mutatie-hebben
ipilimumab & nivolumab	CADTH	2018	"Opdivo in combo with Yervoy for Renal Cell Carcinoma (PC0132-000)", https://www.cadth.ca/opdivo-combo-yervoy-renal-cell-carcinoma-details
ipilimumab & nivolumab	NICE	2019	"Nivolumab with ipilimumab for untreated advanced renal cell carcinoma - Technology appraisal guidance", www.nice.org.uk/guidance/ta581
ipilimumab & nivolumab	PBAC	2018	"Public Summary Document – July 2018 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2018-07/files/nivolumab-and-ipilimumab-melanoma-psd-july-2018.pdf
ipilimumab & nivolumab	ZIN	2019	"Pakketadvies ipilimumab/nivolumab (Yervoy®/Opdivo®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/05/29/pakketadvies-sluisgeneesmiddelen-ipilimumab-yervoy-in-combinatie-met-nivolumab-opdivo

Active ingredient /generic name	HTA organisation	Publication year	Title, URL
osimertinib	CADTH	2019	"Tagrisso for Non-Small Cell Lung Cancer (first line) (PC0137-000)", https://www.cadth.ca/tagrisso-non-small-cell-lung-cancer-first-line-details
osimertinib	NICE	2020	"Osimertinib for untreated EGFR mutation-positive non-small-cell lung cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta621
osimertinib	PBAC	2019	"Product Summary Document – July 2019 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2019-07/files/osimertinib-psd-july-2019.pdf
osimertinib	ZIN	2018	"Pakketadvies osimertinib (Tagrisso®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2018/11/07/pakketadvies-sluisgeneesmiddel-osimertinib-tagrisso-bij-de-eerstelijnsbehandeling-van-patienten-met-gevorderde-of-gemetastaseerde-niet-kleincellige-longkanker-nsclc-met-activerende-egfr-mutaties
palbociclib	CADTH	2016	"Ibrance for Advanced Breast Cancer Resubmission (PC0093-000)", https://www.cadth.ca/ibrance-advanced-breast-cancer-resubmission-details
palbociclib	CADTH	2019	"Ibrance (with Faslodex) for Advanced or Metastatic Breast Cancer (PC0150-000)", https://www.cadth.ca/ibrance-faslodex-advanced-or-metastatic-breast-cancer-details
palbociclib	NICE	2017	"Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta495
palbociclib	NICE	2020	"Palbociclib with fulvestrant for treating hormone receptor-positive, HER2-negative, advanced breast cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta619
palbociclib	PBAC	2017	"Public Summary Document – March 2017 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-03/files/palbociclib-psd-march-2017.pdf
palbociclib	ZIN	2017	"Pakketadvies palbociclib (Ibrance®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2017/04/11/pakketadvies-palbociclib
ribociclib	CADTH	2018	"Kisqali for Metastatic Breast Cancer (PC0112-000)", https://www.cadth.ca/kisqali-metastatic-breast-cancer-details
ribociclib	NICE	2017	"Ribociclib with an aromatase inhibitor for previously untreated, hormone receptorpositive, HER2-negative, locally advanced or metastatic breast cancer - Technology appraisal guidance", www.nice.org.uk/guidance/ta496
ribociclib	PBAC	2017	"Public Summary Document - July 2017 PBAC meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-07/files/ribociclib-psd-july-2017.pdf
ribociclib	ZIN	2017	"Pakketadvies ribociclib (Kisqali®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2017/12/21/pakketadvies-ribociclib-kisqali-bij-gemetastaseerde-borstkanker
tisagenlecleucel ALL	NICE	2018	"Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years - Technology appraisal guidance", www.nice.org.uk/guidance/ta554
tisagenlecleucel ALL	MSAC	2019	"Public Summary Document - Application No. 1519 – Tisagenlecleucel (CTL019) for treatment of refractory CD19-

Active ingredient /generic name	HTA organisation	Publication year	Title, URL
			positive leukaemia and lymphoma", http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1519-public
tisagenlecleucel ALL	ZIN	2018	"Pakketadvies sluisgeneesmiddel tisagenlecleucel (Kymriah®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2018/12/18/pakketadvies-sluisgeneesmiddel-tisagenlecleucel-kymriah-voor-de-behandeling-van-b-cel-acute-lymfatische-leukemie-b-cel-all-bij-kinderen-en-jongvolwassenen-tot-25-jaar
tisagenlecleucel DLBCL / ALL	CADTH	2019	"Tisagenlecleucel (Kymriah) for Pediatric Acute Lymphoblastic Leukemia and Diffuse Large B-Cell Lymphoma", https://www.cadth.ca/tisagenlecleucel-kymriah-pediatric-acute-lymphoblastic-leukemia-and-diffuse-large-b-cell-lymphoma
tisagenlecleucel DLBCL	NICE	2019	"Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies - Technology appraisal guidance", www.nice.org.uk/guidance/ta567
tisagenlecleucel DLBCL	MSAC	2019	"Public Summary Document - Application No. 1519.1 – Tisagenlecleucel (CTL019) for treatment of relapsed or refractory diffuse large B-cell lymphoma (DLBCL)", http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1519.1-public
tisagenlecleucel DLBCL	ZIN	2019	"Pakketadvies tisagenlecleucel (Kymriah®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/03/07/pakketadvies-sluisgeneesmiddel-tisagenlecleucel-kymriah
venetoclax & retuximab	CADTH	2019	"Venclexta in combo Rituximab for Chronic Lymphocytic Leukemia (PC0162-000)", https://www.cadth.ca/venclexta-combo-rituximab-chronic-lymphocytic-leukemia-details
venetoclax & retuximab	NICE	2019	"Venetoclax with rituximab for previously treated chronic lymphocytic leukaemia - Technology appraisal guidance", www.nice.org.uk/guidance/ta561
venetoclax & retuximab	PBAC	2018	"Public Summary Document – November 2018 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2018-11/files/venetoclax-psd-november-2018.pdf
venetoclax & retuximab	ZIN	2019	"Pakketadvies venetoclax (Venclyxto®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2019/05/08/pakketadvies-venetoclax-venclyxto-in-combinatie-met-rituximab

ZIN = Zorginstituut Nederland, NICE = National Institute for Health and Care Excellence, CADTH = Canadian Agency for Drugs and Technologies in Health, PBAC = Pharmaceutical Benefit scheme and medical services Advisory, MSAC = Medical Services Advisory Committee

S2 Table. References to the included HTA-reports related to the cases on orphan pharmaceuticals (n=21)

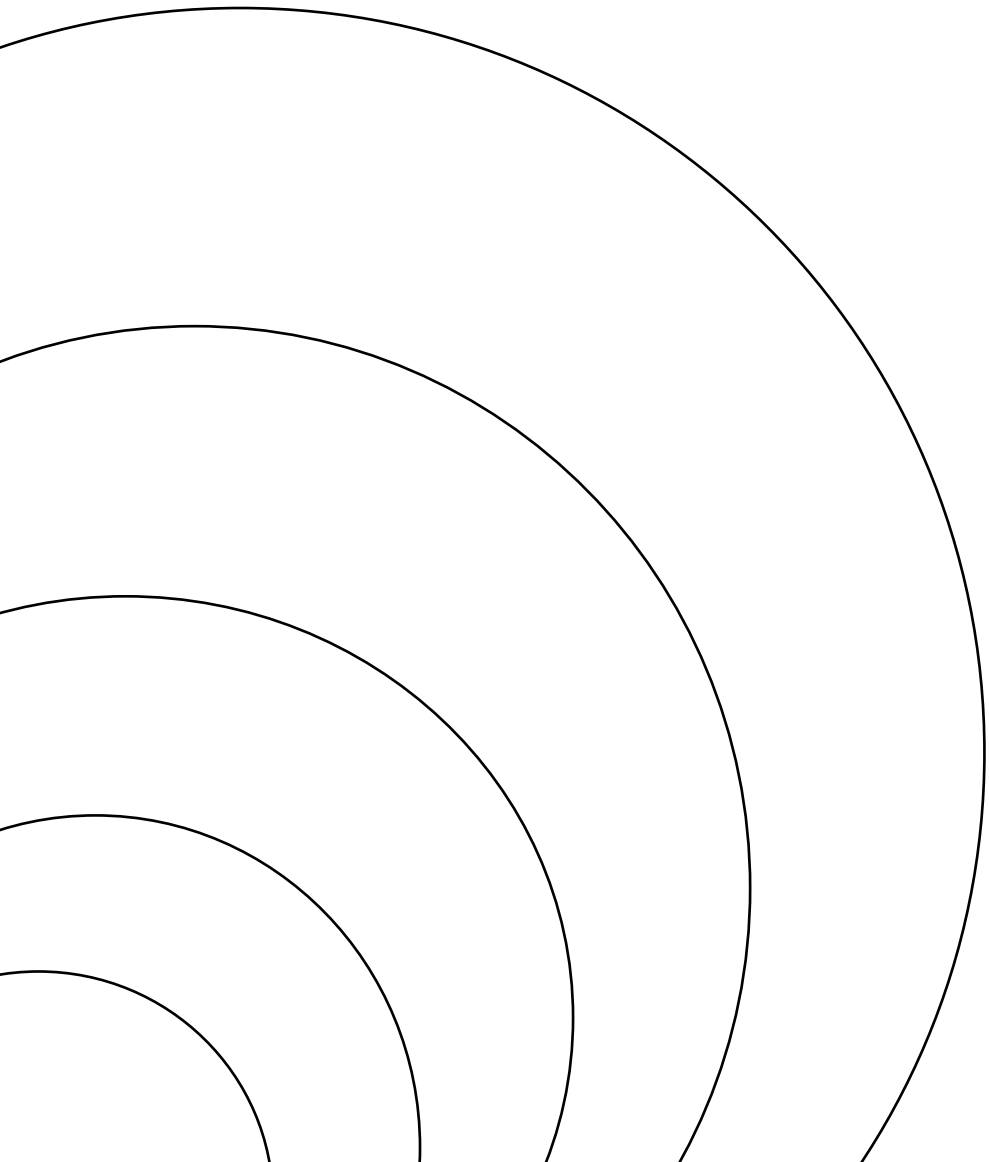
Active ingredient /generic name (indication)	**HTA organisation**	**Publication date**	**Title, URL**
eculizumab (PNH)	PBAC	July 2008	"Public Summary Document July 2008 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2008-07/Eculizumab_Final_PSD_Alexion_Pharmaceuticals_Inc.pdf
eculizumab (PNH)	PBAC	March 2009	"Public Summary Document March 2009 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2009-03/Eculizumab_Alexion_Pharmaceuticals_PSD_7-3_2009-03_Final.pdf
eculizumab (PNH)	CADTH	February 2010	"Common drug review - CEDAC Final Recommendation Eculizumab New Indication:: Paroxysmal Nocturnal Hemoglobinuria", https://www.cadth.ca/media/cdr/complete/cdr_complete_Soliris_February_18_2010.pdf
eculizumab (PNH)	PBAC	July 2010	"Public Summary Document July 2010 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2010-07/Eculizumab_SOLIRIS_Alexion.pdf
eculizumab (PNH)	ZIN	May 2016	"Pakketadvies eculizumab (Soliris®) bij behandeling van PNH-patiënten", https://www.zorginstituutnederland.nl/publicaties/adviezen/2016/05/13/pakketadvies-eculizumab-soliris-bij-behandeling-van-pnh-patienten
eculizumab (PNH)	ZIN	June 2017	"Pakketadvies eculizumab (Soliris®) bij behandeling van PNH-patiënten", https://www.zorginstituutnederland.nl/publicaties/adviezen/2017/06/09/pakketadvies-eculizumab-soliris-bij-pnh---herbeoordeling
eculizumab (aHus)	PBAC	March 2013	"Public Summary Document March 2013 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2013-03/eculizumab-psd-03-2013.pdf
eculizumab (aHus)	CADTH	July 2013	"Common drug review - CDEC Final Recommendation Eculizumab New Indication: Atypical Hemolytic Uremic Syndrome", https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_Soliris-aHUS_July-23-13.pdf
eculizumab (aHus)	PBAC	March 2014	"Public Summary Document March 2014 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2014-03/eculizumab-psd-03-2014.pdf
eculizumab (aHus)	PBAC	August 2014	"Public Summary Document July 2014 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2014-08/eculizumab-psd-07-2014.pdf
eculizumab (aHus)	NICE	January 2015	"Eculizumab for treating atypical haemolytic uraemic syndrome - Highly specialised technologies guidance", https://www.nice.org.uk/guidance/hst1
eculizumab (aHus)	CADTH	May 2015	"Common drug review - CDEC Record of Advice Eculizumab New Indication: Atypical Hemolytic Uremic

Active ingredient /generic name (indication)	HTA organisation	Publication date	Title, URL
			Syndrome", https://www.cadth.ca/sites/default/files/cdr/advice/cdr-advice-Soliris-aHUS-June-2-2015.pdf
eculizumab (aHus)	ZIN	November 2016	"Pakketadvies eculizumab (Soliris®) bij behandeling van aHUS-patiënten", https://www.zorginstituutnederland.nl/publicaties/adviezen/2016/11/21/pakketadvies-eculizumab-soliris-bij-behandeling-van-ahus-patienten
Lumacaftor/ ivacaftor (CF)	PBAC	March 2016	"Public Summary Document – March 2016 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2016-03/files/lumacaftor-ivacaftor-psd-march-2016.pdf
Lumacaftor/ ivacaftor (CF)	ZIN	May 2016	"GVS rapport 16/08 lumacaftor/ivacaftor (Orkambi®)", https://www.zorginstituutnederland.nl/publicaties/adviezen/2016/05/13/gvs-advies-lumacaftor-ivacaftor-orkambi-bij-cystische-fibrose-cf-bij-patienten-van-12-jaar-en-ouder-die-homozygoot-zijn-voor-de-f508del-mutatie-in-het-cftr-gen
Lumacaftor/ ivacaftor (CF)	NICE	July 2016	"Lumacaftor–ivacaftor for treating cystic fibrosis homozygous for the F508del mutation - Technology appraisal guidance", www.nice.org.uk/guidance/ta398
Lumacaftor/ ivacaftor (CF)	CADTH	October 2016	"Common drug review – CADTH Canadian Drug Expert Committee Final Recommendation Lumacaftor / Ivacaftor Indication: Cystic Fibrosis, F508del-CFTR mutation", https://www.cadth.ca/sites/default/files/cdr/complete/SR0471_complete_Orkambi-Oct-28-16.pdf
Lumacaftor/ ivacaftor (CF)	PBAC	November 2016	"Public Summary Document – November 2016 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2016-11/files/lumacaftor-ivacaftor-psd-november-2016.pdf
Lumacaftor/ ivacaftor (CF)	ZIN	December 2016	"Herbeoordeling lumacaftor/ivacaftor (Orkambi)" https://www.zorginstituutnederland.nl/publicaties/adviezen/2016/12/15/gvs-advies-lumacaftor-ivacaftor-orkambi-bij-cystische-fibrose-cf-bij-patienten-van-12-jaar-en-ouder-die-homozygoot-zijn-voor-de-f508del-mutatie-in-het-cftr-gen-herbeoordeling
Lumacaftor/ ivacaftor (CF)	PBAC	July 2017	"Public Summary Document – July 2017 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2017-07/files/lumacaftor-ivacaftor-psd-july-2017.pdf
Lumacaftor/ ivacaftor (CF)	PBAC	July 2018	"Public Summary Document – July 2018 PBAC Meeting", https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2018-07/files/lumacaftor-with-ivacaftor-age-12-over-psd-july-2018.docx.pdf

ZIN = Zorginstituut Nederland, NICE = National Institute for Health and Care Excellence, CADTH = Canadian Agency for Drugs and Technologies in Health, PBAC = Pharmaceutical Benefit scheme and medical services Advisory

Chapter 3
The role of information on profitability in hypothetical reimbursement decisions

Based on: Enzing JJ, Himmler S, Knies S, Brouwer WBF. Do Profit Margins of Pharmaceuticals Influence Reimbursement Decisions? A Discrete Choice Experiment Among Dutch Healthcare Decision Makers. Value Health. 2022 Feb;25(2):222-229. doi: 10.1016/j.jval.2021.08.007. Epub 2021 Sep 9. PMID: 35094795.



Abstract

The aim of this study was to investigate whether the profit margins of pharmaceuticals would influence the outcome of reimbursement decisions within the Dutch policy context.

We conducted a discrete choice experiment among 58 Dutch decision makers. In twenty choice sets, we asked respondents to indicate which of two pharmaceutical treatment options they would select for reimbursement. Options were described using five attributes (disease severity, incremental costs per QALY, health gain, budget impact and profit margin) with three levels each. Additionally, cognitive debriefing questions were presented, and for validation debriefing interviews were conducted. Choice data were analyzed using mixed logit models, also to calculate marginal effects and choice probabilities.

Results indicated that the specified levels of profit margins significantly influenced choices made. Decision makers were less likely to reimburse a product with a higher profit margin. The relative importance of profit margins was lower than that of the included traditional health technology assessment (HTA) criteria, but not negligible. When asked directly, 61% of respondents indicated that profit margin should play a role in reimbursement decision making, although concerns about feasibility and the connection to price negotiations were voiced.

Our results suggest that if available to decision makers the profit margin of pharmaceutical products would influence reimbursement decisions within the Dutch policy context. Higher profit margins would reduce the likelihood of reimbursement. Whether adding profit margin as an additional, explicit criterion to the HTA decision framework would be feasible and desirable is open to further exploration.

3.1 Introduction

Health care expenditures are increasing in many countries, leading to questions about financial sustainability of health care systems and optimal allocation of resources. Economic evaluations and, more broadly, health technology assessment (HTA) may help to control costs and inform allocation decisions within the health care sector.[57] In many countries, including the Netherlands, HTA is used to inform reimbursement and pricing decisions.

Traditionally, pharmaceutical products are relatively often subject to HTA before a decision is made on reimbursement.[58] HTA offers a systematic way of considering whether and under which circumstances pharmaceuticals offer value for money to the health system and society. Given the increase in the number of new pharmaceutical products, which sometimes may be perceived as relatively expensive,[10] a sound assessment of their costs and benefits may be considered necessary, especially given the pressure on overall health care budgets. In that context, one of the cost components in an HTA is the (initial, official, 'list', or requested) price of the pharmaceutical under evaluation. Together with information on, amongst other things, target population, clinical effectiveness, and broader cost consequences of using the pharmaceutical, this information on the price of a pharmaceutical is used to assess whether it may offer value for money.

More recently, in several jurisdictions questions have been raised about the sustainability and 'fairness' of the prices asked or set for pharmaceuticals (for an extensive definition of 'fair pricing' for pharmaceuticals we refer to Moon et al. (2020)).[10, 38] These questions appear to pertain to both the general question whether *given prices* some products can be perceived to still offer value for money, and, even if this is the case, whether the division

of surplus implicitly proposed through these prices can be considered 'fair'.[10, 59] While the first question is answered through common economic evaluations (often, by the way, equating prices with costs) and judging an ICER to a relevant threshold,[21] the second is more difficult to answer for several reasons. Prices will normally need to cover different elements, including components to cover the required R&D costs, the (marginal or average) production costs, distribution costs as well as a profit margin.[10] While the former elements may normally follow from the development, manufacturing and distribution process, the latter under some circumstances may be determined by the manufacturer (within the boundaries of existing thresholds). By setting the profit margin higher, more of the surplus generated by the pharmaceutical is appropriated by the manufacturer, at the expense of the payer. Commonly, however, these payers, often government authorities, have limited information on the exact cost components of the pharmaceutical under evaluation in relation to its price. Hence, the division of surplus, or the 'fairness' of the profit margin or price is not directly observed.

Nonetheless, given the increase in expensive new pharmaceuticals, as well as reports on the relatively high levels of profitability among pharmaceutical companies,[60, 61] prices and profit margins are receiving attention. The increased reliance on price negotiations in several jurisdictions, presumably aimed at reducing profit margins, and changing the division of surplus, is relevant to mention here as well. Moreover, new pricing models have been proposed [10, 23, 24] and calls for more transparency on cost components has been advocated, [59] to justify prices or to allow cost-based price models.[27] More transparency would allow more insight into profit margins of particular products, which could be relevant for and used within HTA.

At present, profit margins of health technologies are not part of the explicit

criteria considered during reimbursement decision making.[42] Obviously, this may reflect the fact that this information is typically not available. Nonetheless, it is interesting to understand how information on profit margins could influence reimbursement and pricing decisions, for at least two reasons. First of all, including information on profit margins more systematically in HTA is only relevant if doing so actually is expected to affect final decisions. If it does, profit margins could perhaps be considered as an explicit criterion in HTA. Currently, to our knowledge, direct evidence on this issue is lacking. Second, given increased attention for profit margins, 'fair pricing' of pharmaceuticals and price negotiations, and despite the fact that profit margin is not an explicit criterion at present, perceptions of or incidental information on profitability of specific interventions might play a role in reimbursement decisions. Indeed, in the context of some previous reimbursement processes, manufacturer costs in relation to prices appear to have been considered relevant to the reimbursement decision.[62, 63] This could suggest that in some instances, (perceptions of) profit margins may already play a role in (some) decisions.

Currently it is unknown whether information on profit margins, if available to decision makers, would influence reimbursement decisions when being used alongside more traditional HTA criteria, like clinical effectiveness, severity of illness, cost-effectiveness, and budget impact. Therefore, the aim of this study was to assess whether and to what extent actual healthcare decision makers would take information on profit margins into account in hypothetical reimbursement decisions, when presented alongside common information on pharmaceutical products. In order to investigate this further, in line with previous studies,[64, 65] we conducted a discrete-choice experiment (DCE) in a sample of Dutch policy makers. While we also acknowledge the importance of more normative consideration of the desirability and optimal way of presenting information on profit margins

in HTA informing reimbursement decisions, this provides first important insights into the potential role of profit margins in such decisions.

3.2 Methods

The setup of our study followed Koopmanschap et al. (2010),[66] who used a DCE to elicit preferences regarding the applied health priority setting criteria among Dutch healthcare professionals. Koopmanschap et al. asked respondents to select one of two different unlabeled, curative treatment options for reimbursement, using twenty-seven choice sets. Best-practice guidelines were applied for developing and analysing our DCE.[67-69] Ethical approval for this study was obtained from the Research Ethics Review Committee of the Erasmus School of Health Policy & Management.

3.2.1 Attribute and level development

We reused the four most influential attributes from Koopmanschap et al. (2010),[66] i.e. disease severity, incremental costs per QALY, individual health gain and budget impact. Their current relevance was confirmed in a recent description of the Dutch HTA process[58] and by reviewing recent Dutch National Health Care Institute (Zorginstituut Nederland; ZIN) reimbursement reports. Furthermore, we included the expected level of profitability (profit margin in %) of the pharmaceutical as an additional attribute. While other criteria may also be relevant in this decision context,[21] we limited the number of attributes to five to keep the choice tasks cognitively feasible and in light of our main research aim.

Three levels were set for each attribute, which were sought to represent a realistic and distinctive range of the respective characteristics in current Dutch reimbursement decision making practice. This was validated by reviewing recent ZIN reimbursement reports. Levels for disease severity and individual health gain in QALYs were directly extracted from

Koopmanschap et al. (2010).[66] Level values for incremental costs per QALY (ICER) and budget impact were adjusted upwards, to correct for changes since the time of that study. The levels for profit margin were specified to range from 5% to 50%, avoiding values which may be perceived as unrealistically low or excessively high, but still providing a distinct range. As values of profit margins on product level from current Dutch (or any other) reimbursement practice are unavailable, the mid-level (20%) was set in relation to the average profit margin on industry level. For the the U.S. context, a mean net income margin of 16.2% has been reported [60] while in Dutch context a net profit margin of the pharmaceutical industry of 17.5% has been suggested.[61]

Table 3.1 Attributes and levels

Concept	**Attribute**	**Levels**
Health gain	Number of gained QALYs per patient	0.5, 2, 4 (QALYs)
Cost-effectiveness	Incremental cost per QALY (ICER)	20,000, 60,000, 120,000 (€ per QALY)
Budget impact	Additional national medical costs per year	10, 50, 100 (million €)
Severity	Disease severity before treatment	Low, moderate, high
Profit margin	Expected level of profit margin of product (in % of price)	5%, 20%, 50%

Note. QALY, quality-adjusted life-year; ICER, incremental cost-effectiveness ratio.

3.2.2 Experimental design

The created attributes and levels (Table 3.1) were used in a set of 20 unlabelled, pairwise comparisons of hypothetical treatment options. An opt-out option was not included as our main interest concerned eliciting relative

preferences and to maximise the amount of information obtained per choice task. To create the choice tasks, we applied a Bayesian D-efficient design, implemented using Ngene Software (Version 1.2.1). The design was optimized for a standard multinomial logit model, based on a utility function including main-effects (the levels themselves) and two-way interactions between cost-effectiveness and budget impact, which was significant in the DCE performed by Koopmanschap and colleagues,[66] as well as the level of profitability and all other attributes. Lowest budget impact (€ 10 million) was prohibited to appear together with highest QALY gain (4) and highest ICER (€ 120,000/QALY) to prevent unrealistic choice sets considering the context information given to respondents (see below). Information on priors necessary for the Bayesian optimization were initially based on expert judgment, supported by the results from Koopmanschap et al. (2010),[66] and updated after the first 22 respondents completed their survey. The mean D-error of the updated design reported by Ngene was 0.033 (SD 0.010).

3.2.3 Survey design and information provided to respondents

Participants were informed that the aim of the survey was to provide insight into the influence of different characteristics of pharmaceuticals on hypothetical reimbursement decisions, without placing particular emphasis on profit margins. Thereafter respondents provided informed consent and were acquainted with the choice task format using a simple pairwise comparison of everyday products. Respondents were then asked to imagine being a health care decision maker, having to decide about reimbursing one of two pharmaceutical products in the Dutch health basic benefit package. Each of the five attributes and their levels were explained step by step, (re-)familiarising respondents with the concepts of budget impact, disease severity, QALYs and ICERs, partly using graphical support. It was explained to respondents that the level of profit margin specified is the price minus

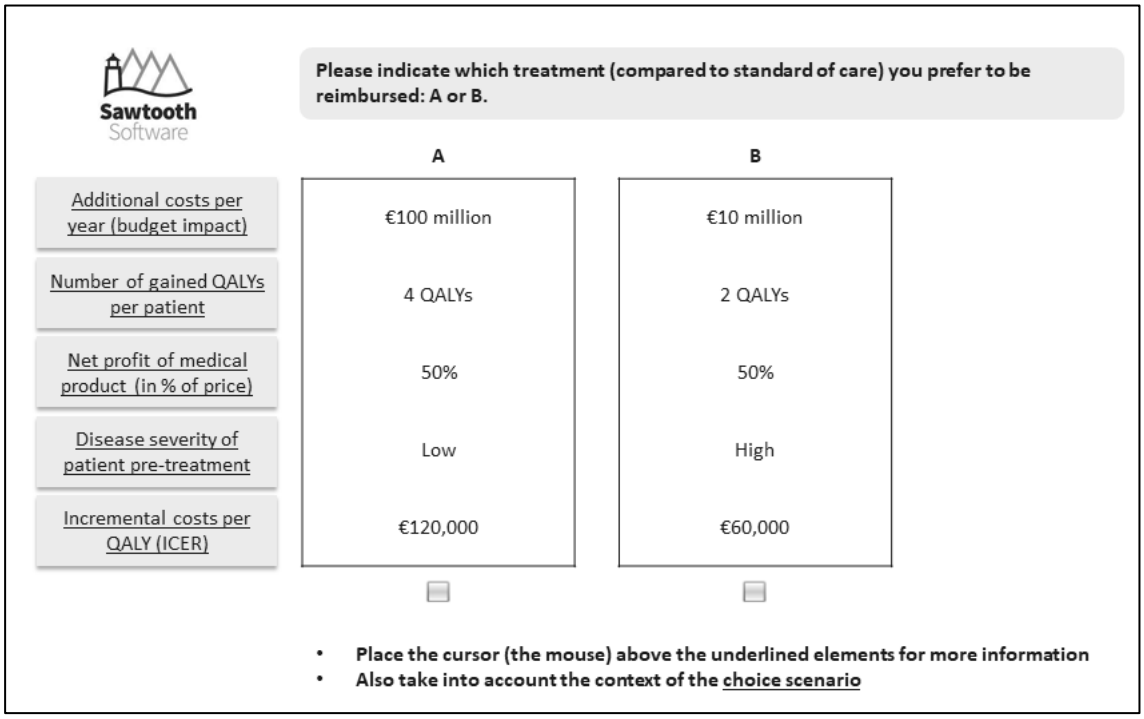
production costs, which specifically include the research and development costs for this product development cycle. It was mentioned that a certain degree of profitability would be required for manufacturers to not hamper innovation. Moreover, it was indicated that profits are not necessarily used for, for example, distributing profits to shareholders but could also be spent on developing new drugs.

To reduce bias by omitting potentially relevant attributes, we also specified the following scenario context, which was informed by the setup and results from Koopmanschap et al. (2010).[66] The respondents were asked to assume that the options related to a pharmaceutical product, which is not already reimbursed for other indications, the treatment recipients were men and women aged 50 to 75, with an average socioeconomic distribution, the product would be an addition to existing therapies in the disease area (and therefore to the basic benefit package), and that the composition of the health gains (duration and quality of life) as well as the number of treated patients (specified to be at least 1,000 per year) were equal across alternatives. This context information was added to avoid specific considerations (e.g. relating to orphan disease status, socioeconomic inequalities, or age profiles) from influencing the choices.[70, 71]

After a warm-up choice task with only two attributes (instead of five), the 20 choice tasks were separated into two blocks of 10 choice tasks each, intermitted by several background questions to reduce response fatigue. The order of attributes and choice tasks were randomised across respondents to prevent ordering bias. Attribute and scenario descriptions were accessible to respondents during all choice tasks, as shown in Figure 3.1, which contains an example choice task. The survey ended with some cognitive debriefing questions, and an open text question, in which respondents could indicate whether they felt profit margins should play a

role in reimbursement decisions. The survey was programmed using Sawtooth software version 9.8.1 (Sequim, WA).

Figure 3.1 Example choice task. Translated from Dutch.



3.2.4 Data collection

The target population of our survey were individuals employed at ZIN or members of ZIN related committees (the Insured Package Advisory Committee and the Scientific Advisory Board). Among other tasks, ZIN is responsible for the assessment of health technologies in the Netherlands. A further inclusion criterion was that individuals needed to be familiar with HTA. The number of eligible individuals who were selected (jointly by JJE and SK) based on these criteria and were subsequently invited via email to participate in the online survey was 92. Data collection took place in October and November 2020. The first 22 full responses were used to update the Bayesian priors for generating a more efficient experimental design. Data collection ended upon obtaining responses from roughly 50% of our respondent pool. Of the 67 instances the survey was started, 51 resulted in

complete responses and 7 in partial responses (i.e. 58 in total). The available choice task responses of the latter were included in the statistical analysis to increase statistical power, leading to a total sample size of 58 decision makers (although the analyses were also run on complete responses only). This sample size exceeded the calculated sample size (n=50) required to identify the main effect (Appendix A 1 in Supplemental Material found at https://doi.org/10.1016/j.jval.2021.08.007 describes the sample size calculation).

3.2.5 Statistical analysis

To account for heterogeneity in preferences towards the HTA criteria included in our experiment, we used mixed logit models to analyse the choice data. Attribute levels were dummy coded, with the (expected) most positive levels as reference categories. The mixed logit models were calculated using 1,000 Halton draws. All main effects were set to be random following a normal distribution, as heterogeneity was found for all attributes. A diagonal covariance matrix was specified, implying independence between the random coefficients, due to the low number of observations and therefore lack of statistical power. Based on model fit and testing for linearity, ICER (in €1,000) and level of profitability (in %) were included as linear attributes. We furthermore tested the inclusion of the two-way interactions we optimized our design for. As a result, an interaction between the level of profitability (as linear term) and the €100 milion budget impact level was included in the model. Finding no heterogeneity in preferences towards this interaction, it was set to be a fixed parameter. Standard errors were clustered at the respondent level. Mixed logit models were calculated using the mixlogit command in Stata 16 (StataCorp, College Station, TX). Based on the final model, marginal effects and choice probabilities were calculated using the *mixlpred* command.

3.2.6 Analysis of open text responses, and debriefing interviews

An inductive content analysis was conducted for the survey responses to the open text question on whether profitability should play a role in reimbursement decisions according to respondents. Two authors (JJE and SK) first independently attached one or more (non-predetermined) topic labels to each answer. Second, based on these labels, answers were categorised into clusters. Third, the two authors jointly selected the five most relevant clusters based on their frequency, homogeneity within clusters, and their distinctiveness compared to other clusters.

To validate the results of the choice experiment, short, semi-structured debriefing interviews were conducted with four DCE participants within a month after participation. Participants were selected to cover different expertise (cost-effectiveness assessment, effectiveness assessment and appraisal), all agreed to an e-mailed invitation. These short video-assisted personal interviews covered individuals' experiences with the survey, their views on the use of profit margins as an additional HTA criterion in practice, and whether including just one additional criteria may have biased the attention towards profitability, and a discussion of the preliminary results of the DCE.

3.3 Results

The characteristics of the survey participants providing their demographic and background information (53 of 58) are shown in Table 3.2. The sample is likely younger and has a higher share of females compared to the total pool of eligible respondents as invited. We observed a low share of non-completes (24%) and 76% of respondents completed the survey within a plausible range of 10 to 32 minutes (median 20 minutes). 78% of respondents agreed that the choice tasks were clear and that the number of choice tasks was manageable. Around half of respondents (partially)

agreed that they based their choices predominantly on just one or two characteristics/attributes.

Table 3.2 Characteristics of respondents providing individual information (53 of 58)

Characteristic	**N =53[1], n (%)**
Age range	
Younger 35	16 (30.2%)
35 to 44	16 (30.2%)
45 to 54	9 (17.0%)
55 to 64	10 (18.9%)
65 and older	2 (3.8%)
Gender	
Male	16 (30.2%)
Female	37 (69.8%)
Self-identified as	
Policy maker/advisor	38 (71.7%)
HTA expert	10 (18.9%)
Other (e.g. medical expert)	5 (9.4%)
Mean completion time in minutes	32.6 (median: 19.8)

Note. [1]For 5 respondents providing choice task data, no information was available.

3.3.1 Preference estimates

Our main results are based on the full sample of 58 respondents, including seven respondents who completed the survey partially. Note that a sensitivity analysis showed that excluding these seven respondents did not lead to noteworthy differences in coefficient estimates. Results from the preferred mixed logit model are shown in Table 3.3 (Appendix Table A 1 in Supplemental Material found at https://doi.org/10.1016/j.jval.2021.08.007 presents estimates from different model specifications, stepwise progressing from the simplest main effects model to the preferred specification). All attribute coefficients were statistically significant, implying that each of the included HTA criteria influenced choices in the experiment. The standard deviations (SD) of all random parameters were

significantly different from zero, indicating preference heterogeneity in all dimensions. Finding only negative coefficients is in line with a priori expectations, as the attribute levels we expected to be the most preferred, were specified as the reference categories.

The largest coefficients for attribute levels in absolute terms were found for low disease severity and a health gain of 0.5 QALYs (-4.535 and -5.215). These translate to marginal effects, i.e., changes in choice probability, compared to their respective reference categories of -32.1% in both cases. The coefficient (-0.054) and marginal effect (-0.3%) of the linear "ICER in €1,000" term, correspond to differences in choice probabilities between the two lowest ICER levels (€20,000 and €60,000) and the lowest and largest ICER levels (€20,000 and €120,000) of -15.7% and -36.3%, respectively. This indicates that cost-effectiveness was the most important HTA criterion in the experiment. The yearly budget impact of a pharmaceutical was less influential, with a marginal effect of the largest level of -20.4%. The coefficient (-0.023) and marginal effect (-0.2%) of a one percent change in profit margin translate to differences in choice probabilities of 3%, moving from 5% to 20% profit margin, and 9% moving from 5% to 50% profit margin. The level of profit margin, therefore, has a lower but nonnegligible impact on choices in the experiment. The marginal rate of substitution between the linear ICER and profitability coefficients (-0.023/-0.054 = 0.43) exemplifies this: a 20% increase in level of profit margin is equally weighted as an increase in the ICER of €8,600. The interaction term between the highest level of budget impact and profit margin (BI100 × Profit) was negative and significant, indicating that given a high budget impact, higher levels of profit margin are evaluated more negatively.

Appendix Figure A 1 in Supplemental Material found at https://doi.org/10.1016/j.jval.2021.08.007 plots preference patterns

according to the "seniority" of respondents, showing that older respondents put more weight on cost-effectiveness (ICER) and the profit margin.

Table 3.3 Preference results of the mixed logit model

	Preference estimates				Marginal effect
Attributes and levels	**Coefficient**	**95% CI**	**SD**	**95% CI of SD**	
Yearly budget impact					
10m €	Reference				Reference
50m €	-1.536***	[-2.060,-1.011]	0.807**	[1.371,0.244]	-12.0%
100m €	-2.584***	[-3.424,-1.743]	1.367***	[0.770,1.963]	-20.4%
Disease severity					
High	Reference				Reference
Moderate	-1.778***	[-2.394,-1.162]	1.302***	[0.587,2.017]	-13.4%
Low	-4.535***	[-6.032,-3.038]	1.682***	[0.745,2.620]	-32.1%
Cost-effectiveness					
Δ ICER in €1,000[1]	-0.054***	[-0.072,-0.035]	0.033***	[0.018,0.047]	-0.3%
ICER €20,000	Reference				Reference
ICER €60,000					-15.7%
ICER €120,000					-36.3%
Health gain in QALYs					
4 QALYs	Reference				Reference
2 QALYs	-2.118***	[-2.895,-1.341]	1.316*	[0.195,2.436]	-15.3%
0.5 QALYs	-5.215***	[-6.454,-3.976]	1.744***	[1.181,2.307]	-32.1%
Level of profitability					
Δ profit 1 %[1]	-0.023**	[-0.038,-0.009]	0.039***	[0.052,0.026]	-0.2%
Profit 5%	Reference				Reference
Profit 20%					-3.0%
Profit 50%					-9.0%
BI100 × Profit	-0.021*	[-0.039,-0.002]			

Log likelihood	-428.9
AIC	891.8
Observations	2,116
Respondents	58

Note. * p < 0.1, ** p < 0.05, *** p < 0.01; AIC, Akaike information criterion; [1]Coded as continuous variable in the model. BI100 × Profit is an interaction term between the highest budget impact and the linear profitability parameter.

To make results more tangible, we calculated the average predicted probabilities of accepting reimbursement for certain hypothetical pharmaceutical scenarios based on the mixed logit coefficients. The first three scenarios relate to products with the worst, middle and best levels for all attributes, respectively (Table 3.4). The corresponding probabilities of reimbursement were 2.1%, 65.3% and 94.4%. The last two scenarios deviated from the middle levels only in their profit margins. The predicted probability of reimbursement given a profit margin of 50% was 61.0% (67.3% for 5% profit margin).

Table 3.4 Choice probabilities of selected reimbursement scenarios

			Scenarios		
Attribute	**Worst**	**Midpoints**	**Best**	**Midpoints + high profit**	**Midpoints + low profit**
Yearly Budget impact	100m €	50m €	10m €	50m €	50m €
Disease severity	Low	Moderate	High	Moderate	Moderate
Cost-effectiveness (ICER)	€120,000	€60,000	€20,000	€60,000	€60,000
Health gain in QALYs	0.5 QALYs	2 QALYs	4 QALYs	2 QALYs	2 QALYs
Profit margin in %	50%	20%	5%	50%	5%
Probability of reimbursement	2.1%	65.3%	94.4%	61.0%	67.3%

Note. QALY, quality-adjusted life-year

3.3.2 Insights from open text responses and debriefing interviews

A total of 61% of respondents indicated that profit margins should have a role in reimbursement decision making, while 39% indicated they should not. In summarising the open text responses, containing their substantiations, five topic clusters could be formulated, based on seven and nine initial clusters formed by two authors independently. First, feasibility concerns regarding obtaining valid information on the profit margin of a pharmaceutical product. Second, inclusion of profit margins already during the reimbursement process may interfere with subsequent price negotiations. Third, cost-effectiveness should be leading in decision making, irrespective of profit margins. Fourth, (high) profit margins may be justified as a reward for innovation. Fifth, the use of profitability as a criterion can help to prevent paying too much as society.

The four individual debriefing interviews did not raise doubts on survey validity. Respondents expressed to have understood the choice tasks. One respondent doubted reproducibility of answers caused by her indifference in some decisions. When reflecting on their view on profit margins in the context of HTA, three respondents expressed to put some weight on profit margin, although less so than to the traditional criteria. Three respondents could recall a reimbursement dossier in which profit margin had been an issue in practice; these consistently involved a repurposed pharmaceutical with an increased price. One respondent suggested that the structure of the survey (choosing between two options) might have increased the attention given to profit margins. Finally, the preliminary results presented in the interviews seemed plausible to participants.

3.4 Discussion

The aim of this study was to investigate whether presenting information on profit margins of pharmaceutical products would influence the outcomes of

reimbursement advice or decisions in the Dutch policy context. In order to investigate this, we conducted a discrete choice experiment among 58 Dutch healthcare decision makers. Our results indicated that profit margins, at least at the levels specified in the DCE, significantly influenced the choices made in our experiment. In particular, decision makers were less likely to reimburse a product with higher profit margins. The importance of profit margins in comparison to other included HTA criteria (health gain, severity, ICER, budget impact) was relatively low, but certainly not negligible. For instance, an increase in the profit margin of 20% was equally influential as an increase in the ICER of €8,600. Arguably, this constitutes a relevant difference within reimbursement decisions. Furthermore, the majority of health care decision makers agreed that profit margins should play a role, although not further defined, in reimbursement decisions. Subgroup analyses indicated that older respondents put more emphasis on profit margins.

The open text responses indicated that main concerns in relation to including information on profit margins in the HTA process concerned the feasibility of measuring and obtaining the relevant information on profit margins. The issue of how such inclusion would relate to potential price negotiations (which may take place after the HTA in the Dutch situation) was also mentioned. Moreover, some respondents emphasized that highly effective innovations may justify also high profit margins.

3.4.1 Strengths and limitations

Our study was based on a unique respondent group, which consisted of persons who work on a day-to-day basis in reimbursement decision making. A further strength of our study is that we undertook several steps to assess the validity of our experiment and its estimates, with generally positive results. This relates to response rate, drop-outs, completion time and

cognitive debriefing questions, but also to the conducted individual debriefing interviews. In terms of the external validity, the estimated relative preferences seemed plausible to debriefing interviewees. Moreover, the incremental effects of the attributes compare well to previous estimates from the Netherlands[66], except for a higher relevance of health gain in this study. Additionally, the results of discrete choice experiments among decision makers in other jurisdictions are consistent with our results in the high relative importance they generally attributed to cost-effectiveness, clinical benefits and disease severity.[72-74]

Nevertheless, some limitations should also be noted. First, we need to acknowledge that we used a stated preference study, based on hypothetical choices, which were different from actual reimbursement decisions (e.g. in practice one does not decide between two alternatives). Hence, our results may not be directly representative of or transferrable to actual reimbursement decisions. Second, we used health care decision makers from the Dutch context, which may hamper the generalisability to other jurisdictions, in which cultural and political context may also play a role. Third, our modest sample size comes with limitations, e.g. not allowing for detailed subgroup analysis. Fourth, a more specific limitation of our experiment is that preferences may have been influenced by status quo bias. Respondents may have put more weight on HTA criteria they were more familiar with and which are more prominently used in current Dutch HTA practice. If this was the case, this may have resulted in an underestimation of the importance of profit margins. In contrast, by only including one 'new and additional' HTA criterion in the experiment, we may have increased the attention paid to profit margin by respondents. Furthermore, the framing of the choice scenarios and of the presented attributes, although aimed to be neutral and balanced, may have influenced the results of the experiment. This not only relates to the concept and

purpose of profit margins, which may be viewed differently by different respondents, but also to the selected levels of profit margin used in our experiment. More generally, the imposed scenario context information was selected to provide an average approximation of the importance of HTA criteria. This also means that in other contexts profitability could have been more (or less) influential (e.g., in the case of first-in-class drugs, or repurposed drugs). A last issue in relation to our study design we want to highlight is that we could only provide respondents with limited context information compared to a real-life decision making context. Additional information on market context (e.g. potential competitors) or price negotiations could have influenced respondents view on and weighting of profit margins. In general, prices of pharmaceuticals may be influenced by negotiations, market structure and other context variables, which could influence the weight placed on them in an HTA. Lacking information on such broader issues, respondents most likely formed an own opinion about this broader context when assessing the choice scenarios.

A general limitation that needs noting is that the type of information we provided in the DCE regarding profitability currently is not generally available. Systematically obtaining information on or estimating profit margins of particular products is not straightforward and would require overcoming many hurdles. E.g. non-trivial uncertainties surrounding profit margins may be unavoidable since profits typically depends on market developments unknown at time of decision making. As long as systematic information on profit margins is lacking, establishing profitability as a criterion may lead to an inconsistent consideration of profit margins in reimbursement decisions, which can have downsides. At the same time, given the feedback of respondents, this may already be the case now.
As an additional subgroup analysis (Appendix Table A 2 in Supplemental Material found at https://doi.org/10.1016/j.jval.2021.08.007 presents the

stratification output), we examined whether preferences differed between respondents who agreed or disagreed on a role for profit margin in the reimbursement decision by interacting the main effects with this indicator. The profitability coefficient is roughly twice as large in the former group (-0.038) compared to the main analysis (-0.023). This underlines the existence of heterogeneity in preferences related to the role of profit margin, also within an HTA organisation.

3.4.2 Implications and future research

The results of our study imply that if information on profit margins would be available within the assessment of a health technology, in general health care decision makers would take it into account in their decisions. Even though the traditional HTA criteria may receive more weight, the influence of profit margins was shown to be non-negligible in our study. This gives raise to several questions and avenues for future research.

A first question would relate to the desirability of having profitability as an additional criterion. Normative work in this area, in relation to new price models and 'fair pricing' of pharmaceutical has been performed, but also highlight divergent views on acceptable divisions of surplus (and thus on what is 'fair'). Adding this criterion at least requires concensus regarding its general relevance amongst the responsible decision makers. In the Netherlands, the overall assessment framework, including the basic benefit package criteria, is eventually determined by politicians.[58] Therefore, future research could investigate politicians', or maybe more importantly their constituencies', normative views on the potential role of the level of profitability in the health technology reimbursement decision context.

Related to this point, it also needs to be determined whether using information on profit margins in the context of an HTA would be the appropriate route to take to ensure an optimal division of surplus.

Alternative ways of addressing this issue are also conceivable, e.g. by conducting price negotiations (potentially guided by a 'fair pricing' framework) or through some form of price regulation.[10, 25, 27].

A second question relates to the feasibility of suffiently operationalising this criterion in practice. Can we obtain or estimate the required information within the full reimbursement process? Formulating a structured process to define and measure or estimate profit margins for the purpose of HTA would be a first step. Future research could also consider the complex relationship between reimbursement decision making, price negotiations and market context, as well as the link between profitability and incentives for innovation.

A final option for future research would be to investigate whether in specific cases, profit margins already play a role in actual decisions. While it currently is not an official criterion, our study not only indicates that profit margins would influence outcomes of decisions if they were known, but also that in some cases respondents felt that profit margins already played a role.

3.5 Conclusions

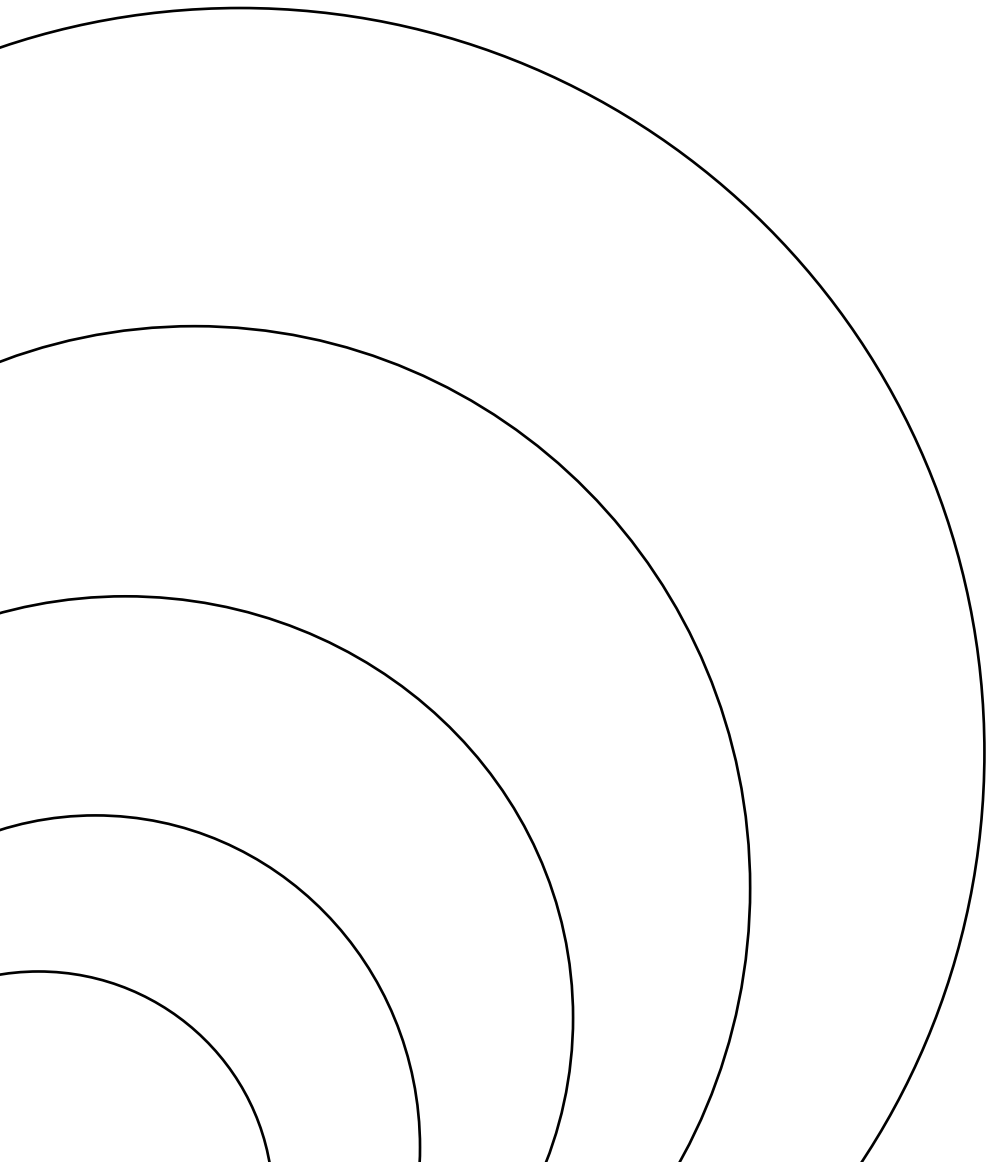
If available to health care decision makers during an HTA process aimed to inform reimbursement decisions, profit margins of pharmaceutical products could be influential, with higher profit margins lowering the likelihood of reimbursement. This highlights the importance of 'fair pricing' also in relation to reimbursement decisions. Whether adding "profitability" as an additional explicit criterion to the HTA decision framework is considered feasible or desirable needs further exploration.

Part 2

Challenges for HTA organisations when broadening the use of HTA towards healthcare interventions other than outpatient pharmaceuticals

Chapter 4
Challenges for reimbursement decision-making in the Netherlands

Based on: Enzing JJ, Knies S, Boer B, Brouwer WBF. Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? Health Econ Policy Law. 2021 Oct;16(4):440-456. doi: 10.1017/S1744133120000237. Epub 2020 Aug 6. PMID: 32758331; PMCID: PMC8460451.



Abstract

Currently, reimbursement decisions based on health technology assessments (HTA) in the Netherlands mostly concern outpatient pharmaceuticals. The Dutch government aspires to broaden the systematic application of full HTA towards other types of health care in order to optimize the content of the basic benefit package. This chapter identifies important challenges for broadening the scope of full HTA to other types of health care. Based on a description of the Dutch reimbursement decision-making process, five important characteristics of outpatient pharmaceuticals were identified, which are all relevant to the successful application of HTA: (i) closed reimbursement system, (ii) absence of alternative policy measures, (iii) existence of marketing authorisation, (iv) identifiable and accountable counterparty, and (v) product characteristics. For a selection of other types of health care, which may be subject to HTA more frequently in the future, deviations from these characteristics of outpatient pharmaceuticals are discussed. The implications of such deviations for performing HTA and the decision-making process are highlighted. It is concluded that broadening the application of HTA will require policy makers to meet both important policy-related and methodological challenges. These challenges differ per health care domain, which may inform policy makers which expansions of the current use of HTA are most feasible.

4.1 Introduction

Similar to many other countries, healthcare costs constitute a significant part of total public spending in the Netherlands, and their share in public spending has been growing [75]. This growth is a concern for the Dutch government which aims to maintain affordability of care, together with quality of care and accessibility of care [76]. Attaining these three public goals is an inherently difficult task, with which many countries struggle. Policy instruments to limit the observed cost increases, while contributing to maintaining an efficient and equitable healthcare system are required therefore, also because there are limits to the degree of risk and income solidarity in societies. Health technology assessment (HTA) can be seen as one such instrument. HTA is an established discipline [77], aimed at informing decision makers about relevant aspects of (new) health technologies, including pharmaceuticals, medical devices, surgical procedures and other healthcare interventions [78]. It is intended to provide a systematic assessment and appraisal of multiple aspects of health technologies relevant to a funding or reimbursement decision. Considered aspects can include for instance effectiveness, cost-effectiveness, legal, social and ethical aspects [79]. In practice, the emphasis in HTA research is often on providing evidence regarding cost-effectiveness of new interventions relative to a relevant comparator, but it can be broader. Through allowing the explicit consideration of all relevant aspects in the decision making process, HTA enables transparent decision making and allocations of scarce resources in line with the overall health system goals. Based on the information provided through HTA research, decision makers may decide to fund or reimburse those technologies within the publicly funded healthcare system, that meet relevant criteria, including those regarding efficiency and equity [18]. Likewise, they could exclude technologies from funding that do not meet the criteria, for instance when not being (sufficiently) effective or cost-effective.

In the Dutch context, a reimbursement decision making process has been gradually developed in which the evidence obtained in HTA research plays an important role. This decision making framework and process embeds the assessment of four main criteria: necessity, effectiveness, cost-effectiveness and feasibility [15]. This process has been operationalised and is currently most systematically applied for the evaluation of new outpatient pharmaceuticals (pharmaceuticals provided by community pharmacists or dispensing general practitioners [80]). Not all new outpatient pharmaceuticals are subject to a full evaluation, but in 2018 for instance twenty-six assessments of such pharmaceuticals were completed [81].

While HTA and the reimbursement decision making process based upon it by now appear well accepted in the field of pharmaceuticals (even though the final decisions may not always receive support), this is not the case for other health technologies. In fact, only a limited number of other healthcare technologies have been subject to an HTA process in the Netherlands. Outside the scope of outpatient pharmaceuticals, HTA appears to be most used in the context of (expensive) inpatient pharmaceuticals (pharmaceuticals exclusively provided by hospital pharmacists [80]). The limited number of evaluations and reimbursement decisions in other areas of healthcare than pharmaceutical treatments may be considered remarkable, especially since there is no a priori reason why policy makers would only be interested in evaluating pharmaceuticals. Moreover, other interventions than outpatient pharmaceuticals, form the larger part of public healthcare spending in the Netherlands [76]. The apparent bias towards applying HTA in the context of reimbursement decisions for outpatient pharmaceuticals, may therefore be difficult to justify. This can be seen as an international phenomenon [29, 30] although initiatives have been taken which may have reduced this bias in specific jurisdictions, e.g. the initiation of the NICE Medical Technologies Evaluation Programme (MTEP) in 2009 [82]

and the Canadian Health Technology Expert Review Panel in 2011 [57]. Indeed, also carefully selecting other technologies than pharmaceuticals for collective financed reimbursement seems an appropriate goal.

The Dutch government has therefore expressed the ambition of broadening the systematic use of full HTA beyond the current scope [11]. This should result in better use of HTA as a policy instrument, a more comprehensive evaluation of technologies in the (publicly financed) health system, and a fairer use of the decision making process across different health technologies. Such broadening requires expanding the systematic application of HTA towards inpatient pharmaceuticals and non-pharmaceuticals (e.g. medical devices, curative interventions and non-pharmaceutical mental healthcare). This intended expansion of the use of HTA in the Netherlands is likely to be challenging. Although on a general level performing an HTA may have clear similarities when applied in the context of different health technologies [29], in practice specific health technologies may require tailored HTA methods and decision making processes. This tailoring is importantly related to the characteristics of the technologies and the relevant healthcare settings considered, as for instance the findings of the European MedTecHTA project showed for medical devices [32].

This article aims to identify important challenges of broadening the application of HTA research and the decision making process based upon it, specifically for the Dutch context. Currently, an overview of HTA challenges covering the general expansion of the application of HTA to other health technologies, is not available in the literature. This chapter aims to present such a general, coherent overview of these challenges, and to subsequently explore possible solutions (also from other jurisdictions) to overcome them in the Dutch context. This article will provide Dutch policy makers who are

responsible for broadening the application of HTA the possibility to prioritise between different health technologies and to anticipate on some of the issues that need to be addressed in the coming years. In discussing these challenges, we take the decision processes and criteria developed for outpatient pharmaceuticals in the Dutch context as the comparator for other health technologies. Our results can also serve as an input to a research agenda aimed at the development of policy solutions and methodologies that would facilitate the broader application of HTA in research and policy. Furthermore, depending on their similarity to the Dutch context, findings may be of relevance for other countries considering the broadening of the systematic application of HTA.

In order to address these issues, we first introduce the Dutch reimbursement system and HTA process (section 2). Then important characteristics of outpatient pharmaceuticals in relation to the application of HTA in the Netherlands are highlighted (section 3). Section 4 discusses these characteristics and the resulting challenges of five types of health technologies which may be subject to Dutch HTA research in the future.

4.2 Reimbursement decisions and HTA in the Dutch context

Like most Western societies, the Netherlands has a healthcare system based on income and risk solidarity, through a number of insurance schemes. Here the focus is on the Health Insurance Act (*Zorgverzekeringswet*), which covers a broad range of curative interventions. This plan is provided by competing private health insurance companies, regulated under public law (van de Ven and Schut, 2009). All these health insurers are obliged to cover the same basic benefit package (BBP) [83], and all Dutch citizens are obliged to take out insurance from one of the insurers. The BBP covers a broad range of healthcare services; including general practitioners' care, hospital care, mental healthcare,

pharmaceutical care and medical devices [84].

The content of the BBP is decided on by the Minister of Health (MoH). Most of the content is described in legal descriptions of reimbursed healthcare, defining the healthcare domains concerned (e.g. as "care normally provided by medical specialists"). This allows for an "open system", which follows the developments in the relevant fields without interference from the MoH. One overarching requirement for the health technologies included in the BBP through the open system is that they have to be part of the "established medical science and medical practice". Otherwise, they need to be regarded in the relevant field as responsible and adequate care and services. This overarching requirement is referred to as the requirement of "effectiveness" [20]. In some cases, specific descriptions are provided regarding inclusions but also exclusions of specific health technologies in the open system. The latter are referred to as "negative lists" [85]. These, as examples, exclude liposuction of the abdomen [86] and fertility-related care for women over forty-two years old [87] from reimbursement.

Much of the practical content of the BBP thus is determined at the level of care providers and health insurers, without direct involvement of the MoH. Only in exceptional circumstances, the Dutch National Health Care Institute (Zorginstituut Nederland; ZIN), an independent governing body and advisor to the MoH regarding the BBP, determines whether specific health technologies meet the requirement of effectiveness. This is often done to inform care providers and health insurers in cases where they do not agree on inclusion [85]. In 2018, ten of such 'clarifications' were published [81].

The MoH can intervene in the "open" system by making changes to the legal framework itself, e.g. by excluding selected interventions from reimbursement by placing them on a negative list. These reimbursement

decisions are typically made ad hoc. For inpatient pharmaceuticals, being part of the open system, the MoH has created the option to suspend reimbursement (see Box 4.1) while reaching a decision on the suitability of the pharmaceutical to be covered under the BBP. The MoH is not legally limited to specific decision criteria or bound to a specific decision process when deciding on an intervention in the open system. However, in practice, the MoH often makes use of the HTA based reimbursement decision making process that is described next.

Box 4.1 The lock

The lock: suspending reimbursement of expensive inpatient pharmaceuticals

The MoH has the option to postpone reimbursement of new inpatient pharmaceuticals with disproportionately high costs per treatment or a high budget impact, by placing these interventions in a so-called "lock" [87]. Without such a policy intervention, these pharmaceuticals would be covered in the open system without further assessment. With the policy intervention, while being in the lock, these pharmaceuticals are not reimbursed. This situation can change in case of a positive reimbursement decision, which mostly would occur after successful price negotiations. The lock was implemented in 2015 in reaction to new, expensive inpatient pharmaceuticals which put increased pressure on hospital budgets.

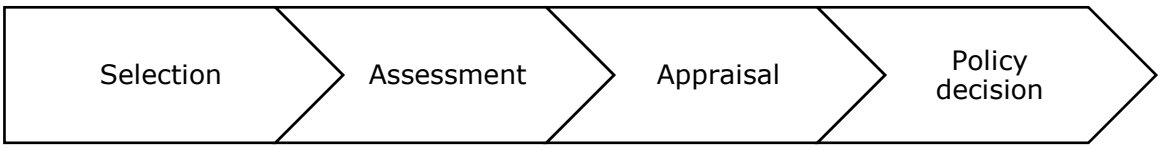
Outpatient pharmaceuticals, in contrast to all other health technologies regulated under the Health Insurance Act, are covered in a "closed system", which is called the Drug Reimbursement System (Geneesmiddelen Vergoedingssysteem, GVS). Their coverage within the BBP is arranged through a 'positive list'. Only when an outpatient pharmaceutical is on this list [86] it is reimbursed. To get on the list, the manufacturer needs to request admission. Only after a positive decision of the MoH, the

pharmaceutical is placed on the list and, hence, reimbursed [80]. In this context, the MoH can use an HTA based reimbursement decision making process.

The Dutch HTA framework

The Dutch HTA based reimbursement decision making process basically consists of four phases: the selection phase, the assessment phase, the appraisal phase and the policy decision phase (Figure 4.1). Together they form a full framework for the systematic application of HTA, especially used in the context of outpatient pharmaceuticals.

Figure 4.1 Phases in the reimbursement decision making process



The first phase is the selection phase which concerns selecting the interventions which become subject of the subsequent decision making phases, since it is not feasible nor desirable to evaluate all new technologies. In the closed system, this selection happens systematically. New products enter this phase when the manufacturer submits an application for admission to the GVS to the MoH [80]. This can be done after receiving marketing authorisation from the European Medicines Agency (EMA). In the open system, selection is less systematic, for instance induced by disputes between health insurers and care providers, direct questions from the MoH, or based on risk assessments of ZIN. Twenty-six technology assessments initiated by the GVS and seventeen technology assessments initiated otherwise (including ten 'clarifications') were published in 2018 [81].

The assessment phase starts with ensuring the selected interventions

belong to the healthcare domain and to be potentially covered under Health Insurance Act, which can be relevant for instance when dealing with lifestyle interventions [88]. If the intervention is deemed not to belong to the healthcare domain, the MoH is advised to exclude it from the BBP. Otherwise, the intervention is typically assessed on four criteria: effectiveness, cost-effectiveness, necessity and feasibility (for a detailed description see Box 4.2). The assessment is mainly based on scientific literature, assuming that the literature is indicative of the real world outcomes of the assessed intervention. For outpatient pharmaceuticals, two explicit rules exist which limit the extent of the assessment. First, when the estimated budget impact, three years after admission to the BBP, is less than €10 million per year, cost-effectiveness does not need to be assessed. That is, the criterion of cost-effectiveness is not used in cases with a 'low' budget impact. Five assessments in which this exemption rule was applied were published in 2018 [81]. Second, typically, if pharmaceuticals do not claim superior therapeutic value than already listed products, cost-effectiveness also does not need to be considered. If equivalent therapeutic value is established, these products are clustered with similar products (in terms of indication criteria, mode of administration and targeted patients) on "list 1A" of the GVS. One price limit applies to all clustered products, hence lowering the need for cost-effectiveness given that the products have a similar therapeutic value. Nine products were added to "list 1A" in 2018 [81]. In other cases, a full assessment is required, including cost-effectiveness. When interventions from the open system are made subject to HTA, the same process is normally followed, although the criterion of cost-effectiveness in practice often is evaluated more limitedly in these cases, which may be related to lack of information.

Stakeholders, like care providers, patient associations and health insurers, are consulted at an early stage of the assessment phase to give their input.

The stakeholders can also be consulted by ZIN during the assessment to obtain relevant information. ZIN moreover collects available scientific evidence. In case of outpatient pharmaceuticals or inpatient pharmaceuticals in the lock, this is done to complement the evidence submitted by the manufacturer. Frequently, available scientific evidence is generated by studies funded by the respective manufacturers [89]. After a dossier is built, the scientific advisory board (Wetenschappelijke Advies Raad; WAR) of ZIN can be consulted, in closed meetings, to assure the scientific quality of the assessment. This board consists of independent academics, clinicians and pharmacists, all appointed by ZIN. Draft versions of assessment reports are sent to stakeholders for comments.

The third phase, the appraisal phase, largely consists of the deliberations by the societal advisory board (Advies Commissie Pakket; ACP) of ZIN. The ACP consists of eight independent experts appointed by the MoH. Their fields of expertise range from clinical practice and patient representation to ethics and health economics. The ACP performs a deliberative societal weighing of the assessed criteria, combined with other aspects considered relevant for the decision, also in a societal context, like the availability of alternatives, orphan status of disease, patient vulnerability and palliative versus curative interventions. The ACP meetings are public and open to participation by external stakeholders. The minutes of these meetings are public as well. Note that not all interventions that go through the assessment phase also necessarily go through the appraisal phase. In fact, for most outpatient pharmaceuticals this latter phase is omitted, while highly expensive inpatient drugs are more often subject to elaborate appraisal. Four appraisals were conducted in 2018, three of which concerned expensive inpatient drugs, placed in the lock [81].

During the final phase, the executive board of ZIN, based on the information

obtained in the assessment and appraisal phase, formulate their advice to the MoH on inclusion in the BBP. The MoH subsequently decides, often in line with this advice. Besides direct inclusion or exclusion, price negotiations with manufacturers, coverage with evidence development (or conditional reimbursement), and arrangements with care providers aimed at the improvement of care delivery can be part of that decision. Such instruments are currently especially used in the context of expensive, inpatient drugs. The final outcome of a decision making process is published in the Law Gazette (Staatscourant).

Box 4.2 Assessment criteria and decision framework

Assessment criteria and decision making process

The assessment criteria used in the reimbursement decision making process are: necessity, effectiveness, cost-effectiveness and feasibility. These criteria have a long tradition in the Netherlands in discussions on choices in healthcare [16]. Each criterion addresses a specific question:

Necessity: Do the disease and the intervention needed justify a claim on solidarity?
Effectiveness: Does the intervention benefit the patient?
Cost-effectiveness: Are the costs of the intervention reasonable in relation to the effects of the intervention?
Feasibility: Is it feasible to include the intervention in the BBP?

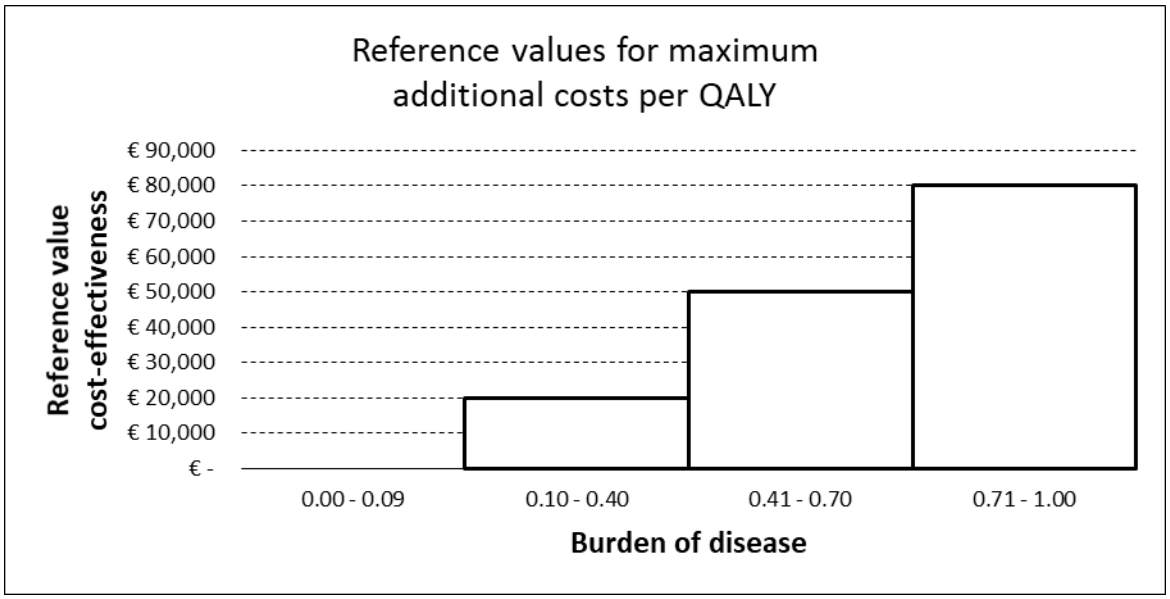
In the assessment phase these questions are answered in a structured and standardized way.
For the assessment of "necessity", the medical necessity to treat the disease is determined. This is captured in terms of so-called "burden of illness", which is calculated as the proportion of otherwise lived health that is lost due to the disease, i.e. *proportional shortfall* [90]. In addition,

it is investigated whether it is necessary to publicly insure the intervention. Inexpensive interventions may for instance be excluded from coverage. How to assess both aspects of 'necessity' has been explained by ZIN in manuals [85].

The criterion "effectiveness" is addressed using the principles of evidence based medicine. This determines whether the intervention conforms with "established medical science and medical practice", according to published standards [91].

Whether the intervention meets the criterion "cost-effectiveness" is investigated by gathering information on incremental costs and incremental health gains, in terms of quality adjusted life years (QALYs), of the intervention compared to a relevant comparator (like best current care) and calculating an Incremental Cost Effectiveness Ratio (ICER). This ICER is subsequently judged against a reference value to determine whether the intervention is cost-effective. Four different reference values are used in that context, depending on the burden of disease established under the criterion 'necessity': €0, €20,000, €50,000, and €80,000 per QALY. The highest reference value is used for interventions falling in the highest burden of disease category as shown in Figure 4.2 [11]. ZIN has issued methodological guidelines for calculating cost-effectiveness, which are publicly available [92].

Figure 4.2 Reference values costs per QALY



The criterion "feasibility" is assessed by mapping out pragmatic issues that can hamper or promote the successful coverage and implementation of an intervention in practice. It for instance explores the (im)possibility to provide the intervention in practice as well as the financial sustainability of covering the intervention in the BBP, using a supporting checklist [85].

Not every criterion has the same role or weight in the decision making process. "Effectiveness" is normally seen as a knock-out criterion - if the intervention fails to demonstrate effectiveness, the other three criteria need not be investigated further. The MoH will then be advised not to reimburse the intervention. Otherwise, for the final decision, all criteria are jointly evaluated, although each separately could lead to a negative advice.

4.3 Important characteristics of the Dutch outpatient pharmaceutical sector

The fact that the focus of HTA applications has been on pharmaceuticals,

seems to be related to both policy decisions and inherent characteristics of the market for and product of outpatient pharmaceuticals. Below we highlight five important characteristics, as assessed by the authors using previous descriptions of the Dutch reimbursement system, which in general are associated with the Dutch outpatient pharmaceutical sector. Note that this list serves as an illustration of relevant defining characteristics of (the Dutch market for) pharmaceuticals that are important in the context of the need and possibility for performing HTA, and is not intended to be exhaustive. In addition, we address pharmaceuticals here in a very general way, abstracting from the large underlying diversity in this area (e.g. in size of patient group, single or combination therapies, or in more or less predictable effectiveness).

4.3.1 Closed system for reimbursement

The structure of the GVS is an important feature. It obliges both stakeholders as well as policy makers to use HTA in the decision making process in the Netherlands. As highlighted above, the GVS forms a closed reimbursement system using a positive list. As a result, before and during the process of decision making a new pharmaceutical is not yet reimbursed. Only after completion of the HTA process and a positive reimbursement decision, the intervention may be reimbursed and becomes available to clinicians and patients in practice. This dependency provides a strong incentive for stakeholders to contribute to a timely start and completion of the HTA process by fulfilling their roles. These roles include applying for admission, providing evidence on effectiveness and cost-effectiveness, and participating in meetings aimed at determining the scope of the HTA. In addition, the closed reimbursement system obliges policy makers to apply HTA to all submitted interventions. It moreover does not require Dutch policy makers to actively search for new interventions that could be made subject to HTA; these interventions actively present themselves through

application for reimbursement by their manufacturers. Note that (the possibility of) having a closed system may also relate to the other mentioned characteristics of outpatient pharmaceuticals.

4.3.2 Absence of alternative policy measures

Related to the previous point is the absence of alternative policy measures to guide and control healthcare expenditures. Given this absence, the perceived need to apply HTA is likely to increase. Once outpatient pharmaceuticals are admitted to the BBP, no specific budgeting policies, or other policies enforcing economic considerations, are in place in the Netherlands. As a result, the consideration of the economic aspects of an intervention is formally limited to the HTA based reimbursement decision, giving this process a unique role. As a consequence, from a policy maker's perspective, this process is positioned as an important, non-optional component of the institutional constellation when economic aspects are to be considered.

4.3.3 Marketing authorisation

The requirement and process of marketing authorisation is an important feature of the market for pharmaceuticals. It provides a base for HTA assessments as it produces evidence on the safety and efficacy of these health interventions [93], as required by the European Medicines Agency (EMA). Although the evidence needed for marketing authorisation does not suffice to demonstrate effectiveness or cost-effectiveness, the conducted studies and their results provide a first fundament for studies on these assessment criteria in the Dutch context. In addition to this first evidence generation, the marketing authorisation process produces standardised documentation on the pharmaceutical and its intended use, which facilitates the HTA. Moreover, a clear marketing authorisation procedure highlights the new products coming on the market, which facilitates horizon scanning

and prioritising HTA research. Related, the requirements for marketing authorisation work as a hurdle, preventing interventions for which generating evidence proved unfeasible or with unfavourable characteristics from proceeding to the "fourth hurdle" of deciding on reimbursement.

4.3.4 Identifiable and accountable counterparty

The presence of a manufacturer capable of producing the required evidence is a fourth important issue typically associated with (outpatient) pharmaceuticals. Many manufacturers have the resources to initiate and finance the studies needed to obtain evidence on effectiveness and cost-effectiveness of their products. They can be and are held responsible by policy makers to produce this evidence if applying for reimbursement in the Dutch context. These characteristics are related to market features and the proprietary nature of pharmaceuticals. A manufacturer of a new pharmaceutical is typically holder of a patent which provides the exclusive right to manufacture and market this new intervention during several years. Consequently, the expected financial revenues of reimbursement of the intervention during those years will benefit a single, identifiable entity. This entity can thus be obliged by policy makers from the start of HTA to produce required evidence for a coverage decision. This makes clear who needs to produce the evidence, which normally is an entity who is in principle capable of actually producing it. The (financial) risk of evidence gathering is thus placed outside ZIN in the Dutch context, and even outside the public domain.

4.3.5 Product characteristics of pharmaceuticals

Pharmaceutical interventions are often standardised products with clearly defined use and functioning, which are aimed at improving patients' length and health related quality of life. Their effectiveness is mainly determined by the active substance, or substances, they contain. Consequently, when

these products are correctly dosed, their effectiveness is relatively independent of the person administering them [31] or the organisational context in which they are provided. This 'confined nature' allows making valid and general statements regarding their effectiveness and cost-effectiveness based on clinical studies (even when conducted outside of the Netherlands), including (double) blinded randomised controlled trails (potentially even placebo controlled), and using traditional HTA methodology, including outcome measures like the QALY.

4.4 Challenges when broadening the application of HTA

The above presented characteristics may not only partly explain the focus on outpatient pharmaceuticals in performing HTA and the use of HTA in decision making in the Netherlands, but they implicitly also point to important challenges when aiming to broaden the application of HTA to other health technologies. In this section, we reflect on five other types of health technologies in relation to the highlighted characteristics. These are: inpatient pharmaceuticals, medical devices, curative interventions (including surgical procedures), non-pharmaceutical curative mental healthcare interventions (including psychotherapy) and non-curative and social care (including care for the elderly). These types of health technologies were selected as an illustration, covering a broad range of health technologies, differing from outpatient pharmaceuticals in different ways. We will only generally address these health technologies, simplifying their characterisation and ignoring the large variations within each type of health technology, for the current purpose. Besides the identification of challenges specific to these types of health technologies, we highlight potential ways forward, again with a focus on the Dutch situation.

4.4.1 Closed system for reimbursement

In contrast to outpatient pharmaceuticals, the five other types of health

technologies used here as illustration of the challenges expanding the scope of HTA in the Dutch situation, are part of the open system for reimbursement in the Netherlands. Therefore, the requirement for manufacturers to apply for admission and to submit evidence on effectiveness and cost-effectiveness is not present for these types of health technologies. Consequently, policy makers are not 'automatically' provided with an inventory of subjects for assessment, let alone with HTA dossiers for these technologies. As a result, they will need to screen and select interventions for assessment, which requires additional effort, clear processes, and rules by which to do so. Horizon scanning methods [94] may contribute to meeting this aim by the identification of specific technologies that need to be subject to assessment. Since 2015 horizon scanning has been performed in the Netherlands, however, this has been limited to new pharmaceuticals [95]. In Canada [96] and in the United Kingdom [97] horizon scanning has already been implemented for a broader range of health technologies (including medical devices and surgical procedures). Experiences in these jurisdictions may provide extremely valuable information to Dutch policy makers aiming to broaden the scope of their horizon scanning and, ultimately, application of HTA. The identification and selection of already reimbursed and used interventions for further investigation may require specific methodologies and processes, also because withdrawing reimbursement of already provided interventions may be a difficult and sensitive topic [13]. In 2014, the 'Appropriate care' programme [98] was introduced by ZIN, which had as one of its aims to identify low-value care (i.e. especially ineffective or low effective care) provided in Dutch healthcare practice. Central to this programme is a systematic screening of the full Dutch BBP, in close cooperation with stakeholders, not limited to any type of health technology. This programme resulted in several studies and publications which pointed at potential areas for improvement in terms of the effectiveness of currently covered and

provided care (e.g. [99]). Although not directly intended for this purpose, this programme may also help to identify already reimbursed health technologies that should be subjected to a full HTA process. Additionally, this programme may be extended to include horizon scanning for new health technologies and, as such, may offer a platform to extend the systematic use of HTA across different health care sectors and also for existing care interventions.

In addition, having an open system often involves the challenge to obtain cooperation from stakeholders in making changes in the coverage of particular health technologies. An open system does not financially incentivise stakeholders to enrol in an HTA process, also because the intervention is already reimbursed during its assessment. Hence, the only change relative to that status quo resulting from an HTA would be negative (withdrawing reimbursement). This could lead to attempts to avoid, postpone or delay the HTA process and to provide less clear evidence (if withdrawal on that basis is less likely). Expanding the scope of "the lock" (box 4.1), currently limited to inpatient pharmaceuticals, may be one route forward for selected interventions. This expansion may require formulating explicit criteria to select interventions to be made subject to "the lock", as well as a clear indication of how the process of evidence gathering and decision making will take place for these interventions. In that context, dependency on stakeholders, especially regarding their provision of evidence, may be reduced by public funding of studies on effectiveness and cost-effectiveness, as will be discussed below.

4.4.2 Absence of alternative policy measures

Budget restrictions exist for each of the other types of health technologies, in the Dutch setting. Hospitals, for example, are funded by health insurers with whom they agree on budget ceilings and fixed budgets [100]. These

restrictions not only lead to a cap on total expenditures, but also require local budget holders (e.g. hospitals) to make choices about the use of these interventions, based on (for them) relevant criteria. As a consequence, on a national level the perceived need to apply HTA may be less pronounced compared to the perceived need to apply HTA to outpatient pharmaceuticals where no such cap on expenditures exists. At the national level, this can result in a lower perceived need to systematically engage in HTA for these interventions.

A downside of these budget restrictions is that differences between care providers can occur if budget holders make different choices about the use of certain interventions. This can lead to unwarranted treatment variation across care providers and patients (ZIP code healthcare). Moreover, the choices made at lower levels in the healthcare system may not align with the principles and goals set at the central level. Put differently, the resulting use of resources may not be the most necessary, effective and cost-effective. Expanding the use of HTA may help in overcoming such differences and suboptimal decisions, but requires central decision making bodies to perform an increased number of assessments as well as ensuring that other actors act in line with their decisions, which may require much effort and better instruments to ensure adherence to centrally made decisions or guidelines.

4.4.3 Marketing authorisation

For most non-pharmaceutical interventions no market authorisation procedure is in place. This means that information regarding effectiveness and safety, as well as regarding the exact intended use of the intervention is not available at the onset of an HTA process. Medical devices form an exception [101]. For medical devices, a system of marketing authorisation is in place, but this is quite different from that for pharmaceuticals. Marketing

authorisation requirements for devices depend on their risk class and range from providing a self-declaration (for devices with low risk) to providing clinical evidence showing that the device works as planned and is safe (for devices with high risk and without an equivalent device present in the market) [102, 103]. Evidence on clinical effectiveness is not required for any of the risk classes. Therefore, also for medical devices evidence on effectiveness is not available at the start of a potential HTA process. This leaves Dutch policy makers for most non-pharmaceuticals without a (systematically enforced) evidence base to start from. Information may be obtained from other sources (e.g. scientific literature generated to inform clinical guidelines), but this may be lacking, differ in form and strength, and needs to be actively collected and processed. When these sources prove insufficient, public funding of evidence generation may be required in order to enable a full HTA.

Systematic overviews of health technologies other than pharmaceuticals entering the market are likely to be absent as well and given the lack of information about their safety, costs and effects, any type of risk-based prioritisation of which technologies to evaluate in an HTA process may prove difficult to apply. Additionally, some of the interventions, including authorised devices, may not be suitable for common types of evaluation (like RCTs). Developing methodologies and processes to scan for new or 'risky' interventions, also to prioritise these for HTA may contribute to formulating solutions to these challenges. As mentioned above (see 'Closed system for reimbursement'), broader horizon scanning for new health technologies has been implemented in some other jurisdictions, which is important to take into consideration in implementing this in the Netherlands.

4.4.4 Identifiable and accountable counterparty

Like for inpatient and outpatient pharmaceuticals, medical devices normally have a manufacturer who may own the exclusive right to manufacture and market a particular device. This would make them an identifiable and accountable counterparty in the HTA process similar to pharmaceutical companies. However, numerous manufacturers in this sector are small and medium enterprises (SMEs) [104], although recent extension of marketing authorisation requirements (Regulation (EU) 2017/745, 2017) may change this situation over time. SMEs may lack the resources (financial and knowledge) to produce the evidence required in a common HTA process. This poses an important challenge, as it requires setting rules for which entities can and which cannot be held responsible for evidence gathering, and which (funding) mechanisms to apply when a manufacturer cannot produce the evidence required to have a meaningful HTA process.

For other types of health technologies, a single manufacturer may not even exist, which emphasises the importance of the issue of who is responsible for evidence gathering and starting the HTA process. Non-pharmaceutical interventions may not be patentable and as a result, no single entity may own the exclusive right to market the intervention. As a consequence, policy makers may not have a clear counterparty to obtain evidence from or to make (price) arrangements with. In the absence of such a counterparty both a (selectively) closed system of reimbursement and market authorisation requirements cannot ensure adequate evidence generation. Creating evidence in the absence of an accountable counterparty then logically would become a public task. This has been shown to be feasible in the Netherlands in previous research programmes [13, 105], and is also the case in the current 'Potentially Promising Care' programme [106]. This programme has an annual budget of €69 million available to provide temporary public funding for research into potentially promising

interventions, which are currently not reimbursed from the Dutch BBP. Health care providers are invited to submit grant applications to this programme, not limited to specific types of health technologies. The costs of care provision can be funded, as well as the costs of research activities. However, the funding of research activities is limited to 20 percent of the total grant. Hence, the 'Potentially Promising Care' programme may be seen as an example of how to overcome the issue of not having an accountable counterparty.

4.4.5 Product characteristics of pharmaceuticals

Except for inpatient pharmaceuticals, each of the other types of health technologies in general differs from outpatient pharmaceuticals in terms of important product characteristics. When reflecting on medical devices, at least three important differences from (outpatient) pharmaceuticals can be distinguished. First, their outcomes may be more context dependent: personal characteristics of the care provider and the organisational context can influence how a device is used and hence the associated costs and effects [31]. Second, medical devices may evolve in daily practice [107]. As a result, a device may develop during evidence collection, or the studied device might not be equal to its current version. In such contexts, research findings have a lower external validity and policy decisions may be based on outdated information. Third, learning effects in their use add to the complexity [108]. The (cost-)effectiveness of an intervention may improve over time due to such individual or organisational learning effects, raising questions about for instance optimal timing of data collection. These three differences challenge the common practice of making one single decision shortly after market access. They also emphasise the importance of the 'maturity' of devices, which is relevant in determining when and how to evaluate the device. Alternative adaptive HTA processes [109] may contribute to meet these challenges. Curative interventions, mental healthcare

interventions, and non-curative care may share these differences and challenges. Additionally, they are often more 'intangible' as 'products' and demarcating them and their use may prove difficult (e.g. [110]), hampering the use of HTA and arriving at clear policy conclusions. Close cooperation with practising care providers may be necessary for their evaluation, in order to standardise the investigated intervention as much as possible. However, other challenges related to such 'intangible' health technologies exist as well, as described by Ergina *et al.* [111] for surgical interventions.

Among other challenges [112], the diversity of the intended outcomes of non-pharmaceutical curative mental healthcare interventions adds to the methodological challenges of performing HTA in this context. More than for many other curative interventions, mental healthcare interventions may be aimed at improving outcomes beyond health-related quality of life of the individual patient. Intended outcomes of such interventions may include wellbeing, autonomy, reduced criminality or drug abuse, and social participation (e.g. [113]). Such goals stress the need for adequate outcome measures, which may not be readily available. Moreover, using different outcome measures for these interventions may improve relevant outcome measurement while at the same time complicate comparisons across diseases and therefore decision making. Future research could further strengthen both methods as well as the policy making process based on adequate outcome measures.

Finally, non-curative interventions, including care for the elderly and palliative care, may often be *primarily* aimed at improving well-being (rather than health) of care users. This focus on well-being brings specific methodological challenges. Although outcome measures aiming to capture wellbeing in different ways and comprising various life domains have been developed [114, 115], further investigation into their performance remains needed [116]. Moreover, a recent scoping review [117] concluded that

considerable disagreement exists on the question which outcomes and which outcome measures are appropriate to use in the evaluation of social care. To date, agreement on the appropriate measures to be used in these contexts appears to be lacking, which complicates interpreting and comparing the results of conducted studies in the decision-making phase. Solutions to these challenges may range from developing new instruments for outcome measurement to alternative adaptive HTA processes [109]. Making progress on actually developing and testing potential solutions for the various challenges may be stimulated by publicly funded research programmes targeted at these issues, such as the Dutch "HTA methodology" programme [118] that was in place in the Netherlands in the past. It will also benefit from close cooperation between HTA methodology experts and policy makers.

4.5 Health technologies and their challenges

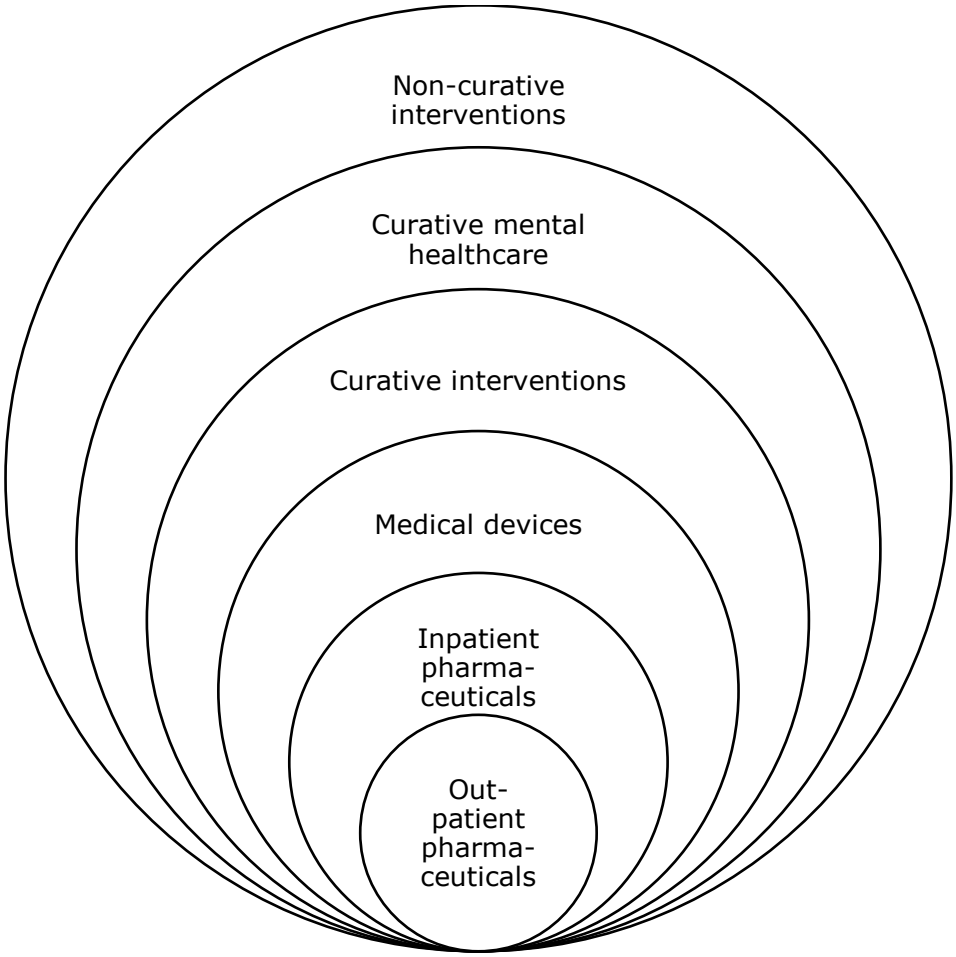
Using the discussed challenges, the five selected types of health technologies can be arranged according to their 'relative distance' in terms of number and degree of differences compared to outpatient pharmaceuticals. Most comparable to outpatient pharmaceuticals are inpatient pharmaceuticals. Differences with outpatient pharmaceuticals in the Dutch context are the open reimbursement system and the presence of alternative policy measures to control costs. As a result, new inpatient pharmaceuticals are not actively presented to policy makers by manufacturers in order to apply for reimbursement. Moreover, these manufacturers have no incentive to actively engage in an HTA process, since the default in the open system is reimbursement. Nonetheless, information on new inpatient pharmaceuticals can be easily obtained from EMA. Expanding the use of the recently introduced "lock" may change this situation, but does require criteria for selecting the interventions entering the lock. It also requires efforts to perform an increased number of

assessments for those interventions entering the lock. After inpatient pharmaceuticals, medical devices may be seen as closest to outpatient pharmaceuticals according to the here discussed characteristics and Dutch context. A limited marketing authorisation procedure for devices is in place and there may be a counterparty capable of supplying the evidence required for an HTA process in some cases, though certainly not all. At present, medical devices are part of the open system in the Netherlands, leading to similar problems as for inpatient pharmaceuticals. Moreover, medical devices may be less standardized, which can hamper the validity of research and the decision-making process. Hence, alternative adaptive HTA processes may need to be developed and tested [109, 119]. Furthermore, although guidance on the assessment of devices has become available (e.g. [120]; [92]), the availability of solutions to highlighted HTA challenges for the assessment of medical devices is limited [121, 122]. Additionally, the absence of an accountable counterparty and of marketing authorisation may result in the absence of evidence for assessment, in some cases. This challenge may be met by setting rules for which entities can be held responsible for evidence gathering, combined with creating funding mechanisms for evidence generation in other circumstances. Furthermore, although fewer European member states (systematically) assess medical devices than pharmaceuticals [30], an opportunity for European collaboration on (improving the methods for) the assessment of medical devices exist [123]. Such collaboration may lower resources needed to perform these assessments.

The other three types of health technologies share these differences and challenges, although arguably to an even stronger degree. In addition, their product characteristics pose additional challenges. The 'intangible' nature of some curative, mental healthcare, and non-curative interventions, may impede demarcating specific interventions and arriving at clear policy conclusions. Furthermore, curative mental healthcare and non-curative

interventions both may differ in their intended outcomes as compared to outpatient pharmaceuticals. This may require other outcome measures than QALYs. In the absence of agreement on such outcome measures, interpreting and comparing results may be complicated, as well as decision making based on outcomes of evaluations. Further development of methods, procedures and decision making processes may be required. Arguably, this will be most challenging for non-curative interventions, which may be seen to have the largest distance to outpatient pharmaceuticals. The resulting order is illustrated in figure 4.3, in which a larger distance from outpatient pharmaceuticals signals additional or more pronounced challenges for HTA in the Dutch context.

Figure 4.3 Illustration of five types of health technologies and their relative distance from outpatient pharmaceuticals.



4.6 Discussion

It has been advocated to broaden the use of HTA in the context of delineating the Dutch BBP. Currently, HTA is especially used systematically when deciding on reimbursement of outpatient pharmaceutical products. This practice may relate to certain characteristics of the outpatient pharmaceuticals in the Dutch context, which make performing HTA there more feasible or desirable. After a description of the Dutch decision making process regarding reimbursement within the BBP, we highlighted five important characteristics of outpatient pharmaceuticals in the Dutch context, which facilitate and stimulate the use of HTA there. Given the aim of expanding the use of HTA, we discussed other types of health technologies in relation to these five characteristics and in the Dutch context. These differences create challenges in applying HTA, which can relate to all phases of the decision making process. They range from the challenge to identify interventions for assessment to the challenge of making meaningful decisions about actual products. Some suggestions for solutions for the highlighted challenges were mentioned, some of which are already partly in place in the Netherlands. Overall, the picture emerges that broadening the systematic application of HTA in the Netherlands requires creating a suitable regulatory and policy framework as well as developing specific methodologies to be able to perform HTA in particular circumstances. Expanding the application of HTA therefore may be a worthwhile goal, but not an easy objective.

To the authors' knowledge this is the first article identifying and discussing important challenges in HTA application for different health technologies from a policy makers' perspective for the Netherlands. The highlighted differences and related challenges should be interpreted in the context of describing and discussing the different health technologies in very general terms, ignoring much of the variation, including in outpatient

pharmaceuticals. Moreover, we discussed the relevant differences from a Dutch perspective, in relation to the overall aim of the chapter to identify important challenges of broadening the application of HTA in the Netherlands. Some of these differences may also be relevant to other countries (e.g. marketing authorisation, accountable counterparty, product characteristics), others may be specific to the Dutch jurisdiction (e.g. open system, absence of alternative policy measures). Nonetheless, the presented challenges will, at least partly, exist in other jurisdictions than the Netherlands as well. For instance, Drummond *et al.* [29] already provided explanations for the focus on pharmaceuticals in international HTA-based reimbursement decision making. Two of the here distinguished characteristics are clearly aligned with their observations: (i) pharmaceuticals are subject of a rigorous licencing procedure (which is in line with our characteristic "market authorisation"), and (ii) pharmaceuticals need to be approved for reimbursement (in line with "closed system for reimbursement"). These similarities emphasise the importance and relevance of the here distinguished characteristics, also in an international context. Drummond *et al.* [29] also mention the sharp increases in pharmaceutical prices, the easily identifiable purchasing chain pharmaceuticals have as discrete products (related to "Identifiable and accountable counterparty"), and an assignment limited to pharmaceuticals for some HTA-programmes, as reasons for the international focus on pharmaceuticals. These reasons provide additional insight in the international dominance of pharmaceuticals in HTA. Future research may focus on identifying and understanding the existing challenges (e.g. using systematic reviews or in depth interviews with various stakeholders). Moreover, international HTA initiatives (e.g. NICE MTEP, CADTH horizon scanning) may provide valuable and relevant international experiences and solutions, also relevant for overcoming (part of) the described challenges in the Netherlands.

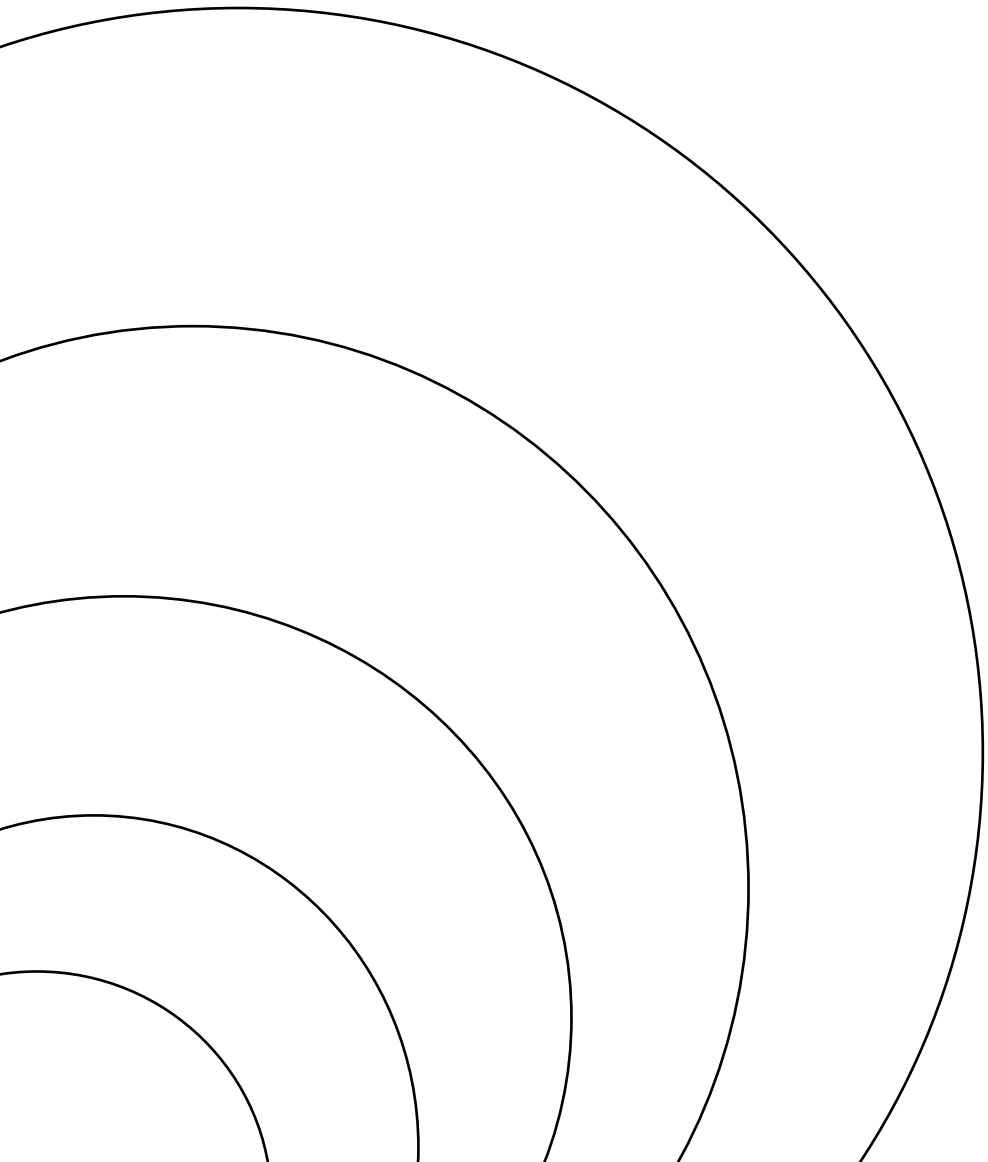
4.7 Conclusion

In light of the discussed differences, and the heterogeneity of health technologies in terms of (intensity of) deviations from the characteristics of outpatient pharmaceuticals, broadening the scope of HTA may be challenging – and more so in some areas than in others. Consequently, it is important for (Dutch) policy makers aspiring to broaden the application of HTA, to do so gradually and aware of the various challenges they are likely to face. A logical route forward may be to start the expansion in those areas in which the number and difficulty of the existing challenges may be least. Interventions relatively similar to outpatient pharmaceuticals, like inpatient pharmaceuticals and medical devices may be logical first steps in coming to a broader use of HTA in defining the Dutch BBP. Meanwhile, necessary preparatory steps can be taken that would facilitate a further expansion of the use of HTA in more challenging health technologies, both in terms of policy context as well as in methodological development. Such a route forward in the broader application of HTA is encouraged. While a bumpy road may lay ahead, a conscious planning may ease the travel, and the destination certainly is worthwhile.

Part 3
Case studies regarding methodological challenges when broadening the use of HTA towards healthcare interventions other than outpatient pharmaceuticals

Chapter 5
Challenges for economic evaluations of medical devices

Based on: Enzing JJ, Vijgen S, Knies S, Boer B, Brouwer WBF. Do economic evaluations of TAVI deal with learning effects, innovation, and context dependency? A review. Health Policy and Technology. 2021;10(1):111-9. Doi: 10.1016/j.hlpt.2020.09.006. ISSN: 2211-8837.



Abstract

Most collectively funded healthcare systems set limits to their benefit package. Doing so requires judgements which may involve economic evaluations. Performing such evaluations brings methodological challenges, which may be more pronounced in non-pharmaceutical interventions. For example, for medical devices, the validity of assessment results may be limited by learning effects, incremental innovation of the devices and the context-dependency of their outcomes.

To review the extent to which "learning effects", "incremental innovation" (related to outcomes) and "context-dependency" are included and/or discussed in peer reviewed economic evaluations on medical devices using Transcatheter Aortic Valve Implementation (TAVI) as an example.

A systematic review was performed including full economic evaluations of TAVI for operable patients with aortic stenosis identified using the Pubmed database. Study characteristics, study results and text fragments concerning the aforementioned aspects were extracted. The quality of the studies was assessed using a quality checklist (CHEC-extended).

Within 207 screened records, 15 studies were identified. Two studies referred to all three aspects, four studies referred to none. "Learning effects" were discussed in five studies, one of which described a method to cope with this challenge. “Incremental innovation” was described in seven studies. Limitations in generalizability of results related to context of care provision were discussed in seven studies.

The challenges related to economic evaluations of TAVI and their influence on the validity of reported results, are typically only partly discussed and rarely dealt within peer reviewed studies. It is important for better informed

policy decisions that this improves.

5.1 Introduction

Collectively funded healthcare systems in Western countries set limits to their benefit package. Setting these limits requires designated authorities to make judgements on whether specific healthcare interventions merit a claim on collective means. These policy judgements may be based on the assessment and appraisal of multiple aspects of health technologies, for instance on effectiveness, legal, social and ethical aspects [79]. Cost-effectiveness may be among these considered aspects. This aspect can be assessed using an economic evaluation. A growing number of these economic evaluations are conducted: i.e. until 2009 almost 2,500 cost-utility analyses (a specific form of economic evaluation) were published in English [57], in 2017 this number had grown to more than 7,000 [124]. Guidelines on how to perform economic evaluations in healthcare are available for many jurisdictions (e.g. [125]). However, despite the growing number of published evaluations, and the existence of guidelines, performing economic evaluations is still not without methodological challenges. As a result, estimates of interventions' incremental cost-effectiveness ratios (ICERs) may be inaccurate and thus policy makers may be misinformed. While some of the methodological challenges in performing economic evaluations are relevant to all types of healthcare, others are more pronounced in specific types of interventions. For medical devices three of such specific challenges have been repeatedly identified as important: learning effects, incremental innovation and context-dependency of outcomes [31, 32, 107, 126, 127]. Although more specific challenges may exist, these three thus seem particularly relevant in the context of medical devices. The concept of learning effects, or learning curves, refers to the situation in which the (cost-)effectiveness of an intervention is related to the experience (and resulting competence) of care providers with using a particular procedure or device. Learning effects can be relevant when accumulating experience and knowledge of care

providers, e.g. during a period of proctoring, lead to an increase in the average effectiveness and/or a decrease in the average costs. Incremental innovation refers to incremental changes through time of the medical device itself (e.g. alterations of its technical specifications) or its provision/use (e.g. alterations in the surrounding clinical pathway), which may cause changes in the efficacy and/or costs of the intervention as well. Finally, context-dependence of outcomes refers to a dependency of (cost-)effectiveness on the (organisational) context of care provision (e.g. organisational size or academic versus non-academic hospitals). All three aspects may thus influence the observed cost-effectiveness, leading to questions of whether this observed cost-effectiveness is generalisable in time, context and place, and therefore most relevant in informing a policy decision (which of course also depends on the policy problem that needs to be addressed). Flexible modelling and appropriate data collection may be among possible solutions to cope with these challenges [128]. Alternatively, researchers may provide a discussion of (the relevance of) these challenges to, at least, inform policy makers on limitations of their study, or present specific sensitivity analyses. Otherwise, when these aspects are (potentially) relevant yet ignored when conducting and reporting an economic evaluation, the reported results may misinform policy makers, who may not be aware of these specific challenges and their impact on the results. This raises the question to which extent "learning effects", "incremental innovation" (related to outcomes), and "context-dependence of outcomes" are accounted for in peer reviewed, full economic evaluations of medical devices. This review aims to answer this question, discuss some policy consequences of not dealing with these challenges, and through that to raise awareness about these challenges and their handling in applied economic evaluations, and ultimately improve the quality of economic evaluations of medical devices and decisions based upon these.

In this review Transcatheter Aortic Valve Implementation (TAVI) is used as

a case study. TAVI is a recently developed, minimal invasive technology initially aimed at inoperable patients with symptomatic aortic valve stenosis. In this context, TAVI was shown to be cost-effective [129]. Currently, the indication of TAVI has broadened towards patients with aortic valve stenosis (AS) who are also eligible for surgical aortic valve replacement (SAVR) [129]. This review focuses on economic evaluations of TAVI with SAVR as comparator. For TAVI, as a complicated, recent and developing technique, each of the three challenges mentioned above is potentially relevant when performing an economic evaluation. Recent economic evaluations for TAVI in this context are available, making this intervention a suitable case for this study. In addition, the aforementioned, recently broadened indication of TAVI may have influenced its costs and outcomes, making TAVI, especially compared to SAVR, a currently relevant topic for policy makers.

As part of the MedtecHTA project Tarricone et al. [130] previously reviewed published economic evaluations (published until December 2014) in order to investigate how they handled four distinctive features of medical devices, including "learning effects" and "incremental innovation". Based on two case studies, TAVI (for all indications) and implantable cardioverter defibrillators (ICD), it was concluded that general awareness of specific features of medical devices is low in the context of health technology assessments (HTA). Meanwhile, the results of the MedtecHTA project have been published and have informed methodological guidance for the assessment of medical devices issued by EUnetHTA [120]. The current review therefore updates the study by Tarricone et al. in the specific context of economic evaluations of TAVI with SAVR as comparator, enabling to assess whether the awareness about / inclusion of learning effects and incremental innovation has increased in published economic evaluations since 2015.

5.2 Methods

5.2.1 Search strategy and inclusion criteria

The systematic review was conducted according to PRISMA guidelines [131]. On November 12th 2018 PubMed was searched to identify publications which fulfilled the inclusion criteria: these publications should contain information on costs and benefits, aortic valve stenosis, transcatheter valve implantation, and surgery. No time restriction was applied. Subsequently, two reviewers (JE & SV) independently reviewed the results, excluding publications which did not report full economic evaluations of TAVI versus SAVR for patients with AS, based on the titles and abstracts of the identified publications. As a result, cost studies, editorials and letters to the editor were excluded. In case of differences between the reviewers, agreement was found through discussion between the two reviewers. Using the full articles of the remaining publications, the two reviewers independently determined whether articles could be regarded as full economic evaluations of TAVI versus SAVR for patients with AS. Again, differences were resolved through discussion between the two reviewers. Systematic reviews were excluded from this final selection, however, their references were cross-checked for relevant full economic evaluations. No search for grey literature was performed, also based on the assumption that policy makers would typically prefer to obtain evidence from peer reviewed studies in the decision making process.

5.2.2 Methodological quality assessment

To determine the quality of the included economic evaluations, the extended Consensus Health Economic Criteria list (CHEC-extended) [132, 133] was used. This tool was selected since it was developed to assess both trial based as model based full economic evaluations, both included in the review. The CHEC-extended list has twenty questions, with response options “yes” or “no”. The two reviewers separately scored the included

economic evaluations using this checklist. In case of differences in scoring, agreement was found through discussion between the two reviewers. For each economic evaluation, a quality ratio was composed by relating the number of positive answers to the number of applicable questions. Since the impact of the individual questions on quality may be incomparable, this ratio must be interpreted with care.

5.2.3 Data extraction

General study characteristics (e.g. perspective) were extracted using a data extraction form (JE, validated by SV). This data extraction form was designed by the authors and implemented in Microsoft Excel. Publications were read in full by the two reviewers and for each publication text-elements (and their section titles) were copied to the extraction form when they were regarded to concern:

- Learning effects: (potential) changes in the efficacy and/or costs of the intervention (TAVI) related to the cumulative experience of operators and/or centres;
- Incremental innovation related to outcomes: (potential) changes in the efficacy and/or costs of the intervention related to its (incremental) innovation through time;
- Context-dependency of outcomes: influence of personal characteristics of the care provider and/or the organisational context (e.g. organisational size and organisational structure) on the efficacy and/or costs of the intervention.

Based on the presence of text-elements on these methodological challenges, these challenges were regarded as undiscussed or discussed within a specific publication. Additionally, the reviewers determined whether any of the challenges mentioned in the text resulted in methodological choices to account for these challenges. If this was the case, this was noted

as an analytical solution in our review. Differences between judgements were resolved after discussion between the two reviewers.

5.2.4 Analyses

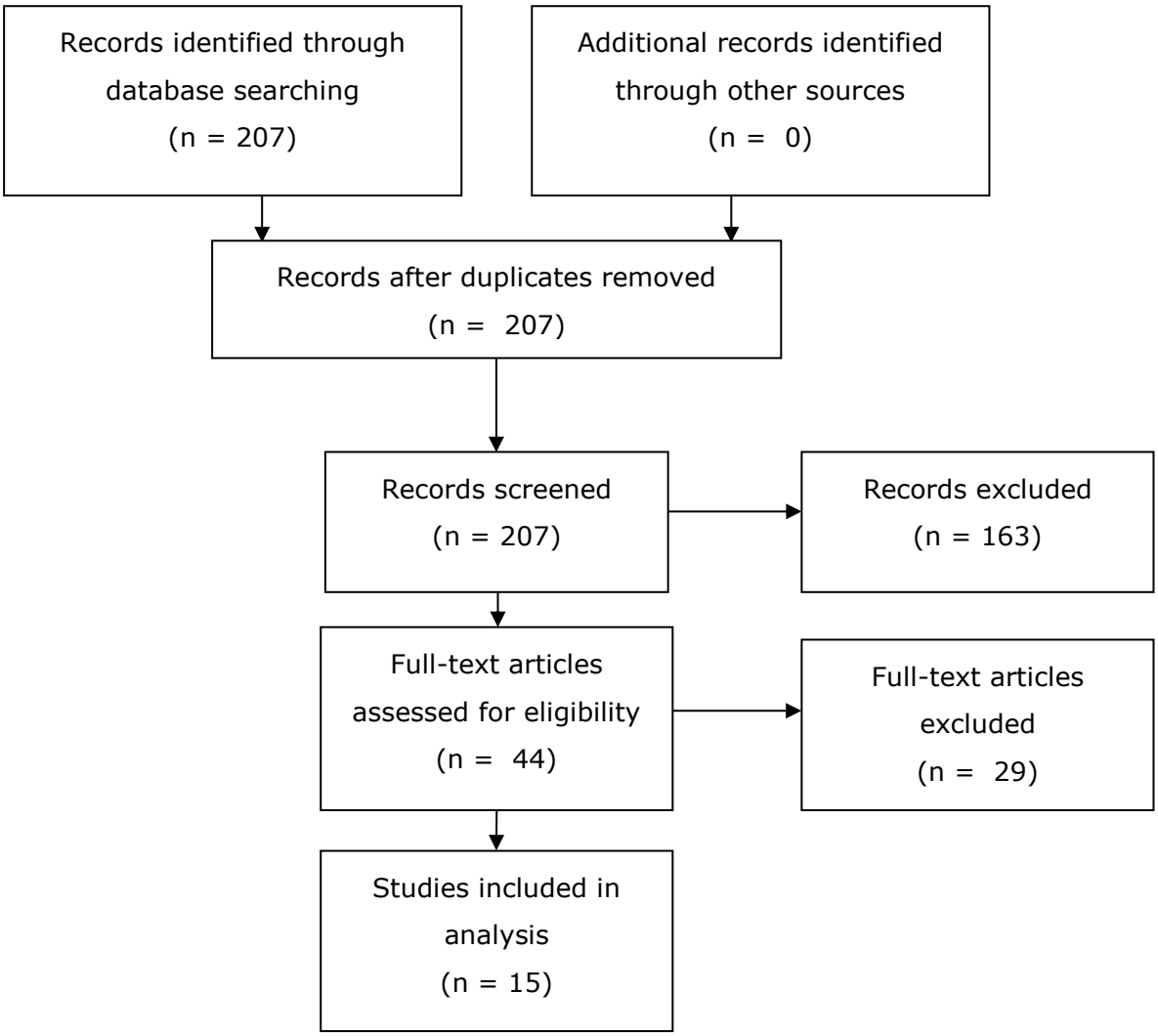
Publications before and after 2015 were compared in terms of the number of challenges discussed per study.

As additional information the results of the economic evaluations (e.g. ICERs) were extracted, also to explore whether TAVI outcomes improved over time, which could suggest learning effects and/or incremental innovation in subsequent studies.

5.3 Results

The literature search resulted in 207 studies, of which 15 studies were finally included (see figure 5.1). Studies were excluded for not being a full economic evaluation (e.g. cost studies) (n= 147), or subsequently for not concerning a comparison of TAVI to SAVR for operable AS patients. Ten systematic reviews were found and used to check their references to find additional peer reviewed full economic evaluations. This did not result in additional studies. Noteworthy, one included HTA-report [134] concerned an update of another included HTA-report [135] which is also described in a journal article [136]. This overlap was not considered problematic, so both were retained.

Figure 5.1 PRISMA Flow Diagram



5.3.1 Methodological quality assessment

The assessment of methodological quality of the included studies using the extended CHEC-list resulted in scores ranging from 12/20 (60 percent) to 17/19 (89 percent). Ten of the checklist items did not differentiate between studies, e.g. all clearly described their study population. No study discussed each validation type required by the checklist. Studies differed in terms of their scores regarding appropriateness of their costs measurement and valuation. Some equated costs with an assumed reimbursement tariff [137]. Differences were also observed in the explicit indication of potential conflict

of interest in the published papers. Ethical and distributional issues were rarely discussed.

5.3.2 Study characteristics

Characteristics of the fifteen included studies are provided in table 5.1. The studies were published from 2012 until 2018, most (12/15) were model based, and most used a payer's perspective (12/15). Most studies (10/15) were North American (Canada, USA) or European (four; United Kingdom, Belgium, Spain) and one was Japanese. Ten studies used a time horizon of ten years or more. Most studies (10/15) were based on the industry-sponsored, multi-centre, randomized controlled Placement of Aortic Transcatheter Valves (PARTNER) trial. Studies targeted two types of operable patients: those with high surgical risk (11/15) and those with intermediate risk (4/15). Most studies investigated a TAVI valve system of Edwards LifeSciences (11/15), others investigated a TAVI valve system of Medtronic, Inc. (3/15), and one study investigated systems of both manufacturers. One of the studies was limited to transapical (TA) implantation of TAVI, the other studies investigated (the less-invasive) transfemoral (TF) implantation or a combination of both routes.

5.3.3 Cost-effectiveness results

Cost-effectiveness outcomes as reported in the included studies are provided in table 5.2. All studies reported incremental effects measured in QALY's, while one-third (5/15) also reported incremental effects measured in life-years-gained. The reported incremental effect of TAVI in QALY's was mostly positive (9/15), and most studies reported additional costs (12/15). All studies presented a sensitivity analysis to quantify uncertainties. Five studies reported that TAVI was dominated by SAVR, two studies reported SAVR was dominated by TAVI, the other studies reported TAVI to have an incremental cost-effectiveness ratio (ICER) ranging from (expressed in

euros) €31,000 to €750,000 per QALY. The variation in the extracted cost-effectiveness results may partly be explained by observed differences in study characteristics, among which country, perspective and the cost and efficacy sources used were prominent ones. Furthermore, the reported incremental QALY's appeared to have an upward trend over time. All studies included some form of sensitivity analysis, among which one-way deterministic sensitivity analysis and probabilistic sensitivity analysis were the most common types of sensitivity analysis (in eleven and ten studies, respectively). None of the sensitivity analysis attempted to quantify the potential impact of learning effects or incremental innovation on the ICER. One of the analyses demonstrated the potential impact of context on cost-effectiveness by imputing country specific costs in two scenarios, changing the reported ICER from dominated in the base case to dominant in (some of) these scenarios [138].

5.3.4 Methodological challenges

Table 5.3 provides the results of the review per study. This table shows that each of the three methodological challenges was discussed in one or more of the studies. Two studies discussed all three challenges [139, 140], while four studies discussed none [134, 135, 141, 142]. The two studies which discussed all three were among the studies with the highest CHEC-list scores. Studies which discussed no or one challenge had a mean CHEC-score of 76%, studies which discussed two or three challenges had a mean CHEC-score of 85%, suggesting a potential relationship between number of discussed challenges and assessed methodological quality. Obviously, sample size prohibits formal testing or firm conclusions. Challenges were discussed in the "Discussion" (or "Comment") sections of the studies or, in one study, in the "Introduction" [143]. Two studies used 'analytical solutions' to deal with identified challenges. The first study was restrictive in the selection of registry data. It selected those registries that allowed inclusion of data after

an initial learning effect, hence avoiding data from situations in which proper training and experience was not yet realized. This was highlighted in the "Discussion" section [144] of the publication. The second study concerned additional analyses to deal with (high-level) context dependency, i.e. an international comparison of results, by using information (i.e. imputing unit costs) from other countries and health care systems to understand cost-effectiveness in these contexts, rather than the country of origin. This issue was described in the "Methods" section [138] of the study. The results highlighted that cost-effectiveness estimates were quite sensitive to these country specific unit cost parameters.

5.3.5 Learning effects

Five studies discussed learning effects, three of which labelling it as "learning curves" or "learning curve effects". For example, Reynolds [140] wrote that "*most PARTNER sites did not perform enough TA-TAVR procedures to move beyond the point of learning curve effects*". The remaining two studies described a positive relation between experience and outcomes. For example, "*In centres experienced in conducting TAVIs, procedural success may be around 90% or more and closely linked to experience, with greater learning resulting in better patient selection and outcomes*" [143]. One of the five studies described how learning effects were taken into account in its model-based analysis: "*Given the recent development of transapical TAVI, we did not include data from registries emphasizing results of a 'learning curve'. Only registries that separated recent procedures, once proper proctoring and training had been completed, were included in the data employed in the model*" [144].

5.3.6 Incremental innovation

In seven of the included studies potential developments of TAVI or its comparator were explicitly related to (future) outcomes, costs and/or the

ICER. For example, "*It is reasonable to expect that iterative improvements in TAVR technology in the short to intermediate term, coupled with increased clinical experience, will lead to reduced complication rates, more efficient care, reduced costs, and improved cost-effectiveness relative to SAVR, a much more mature therapy.*" [139]. No methodological solutions to cope with incremental innovation (e.g. specific sensitivity analysis) were found in the articles.

Besides these seven studies, three other studies contained a text fragment that implied that innovation of the intervention is continuing in daily practice, although without explicitly relating this phenomenon to (future) outcomes, costs and/or the ICER. As examples, Gada et al. labelled TAVI as "*a developing technique*" [144] and Orlando et al. stated that "*more sophisticated delivery systems have been developed.*" [143].

5.3.7 Context dependent outcomes

Seven studies discussed that their results may not be generalizable to other contexts (e.g. jurisdictions or treatment settings). One study [138] conducted a scenario analysis to demonstrate results for additional countries by imputing observed unit costs, as highlighted above. Four of the seven studies specifically discussed the context of care provision (e.g. the specific hospital). For example, "*We recognize that there is substantial institutional heterogeneity with respect to procedural location and resources, and this factor may potentially affect the ICER.*" [145].

To assess in a general fashion whether the awareness, measured as being discussed, of the three challenges increased since the end of 2015, we compared publications before and after 2015 in terms of the number of challenges discussed per study. Ten studies were published in 2015 or before. In these ten studies on average one challenge was discussed (see table 3). Five studies were published after 2015. These studies on average

discussed 1.8 challenges (i.e. nine in total). Notwithstanding the low numbers and rough indicator, this may suggest at least an increase in awareness of the challenges related to the economic evaluation of TAVI. Whether this increased awareness is representative for other medical devices, or e.g. results from the elapsed time since the introduction of TAVI (time effect) requires further research.

5.4 Discussion

This chapter reviewed the extent to which three methodological challenges of particular importance to medical devices, were discussed in peer reviewed full economic evaluations of TAVI, and whether analytical solutions were provided. It was observed that these challenges and their influence on the validity of reported results of economic evaluations, are typically only partly discussed and rarely quantitatively dealt with in the reviewed economic evaluations of TAVI. This seems inappropriate from a general HTA perspective. Within an HTA process, economic evaluations are part of the information which is systematically collected and synthesized during the assessment, to inform a subsequent appraisal phase. During the latter, the available evidence is critically appraised in terms of validity, significance and relevance, along with known uncertainties and all societal and ethical considerations deemed relevant. Information on methodological challenges, both resolved and unresolved, is needed to inform these deliberations. It seems this information is mostly lacking in reports on economic evaluations of TAVI.

For each of the three challenges, this observed absence of information may have specific consequences for policy makers. First, only one study explicitly corrected for the influence of accumulated experience on outcomes. However, the literature highlights that significant learning effects exist in TAVI care provision, both effectiveness as costs are influenced by experience [146-149]. Consequently, when care providers' experience levels

within trials used in economic evaluation differ from those in current or expected practice, the reported ICERs may not reflect actual clinical practice. For example, ICERs, which may aim to represent long-term cost-effectiveness of the use of an intervention, may be overestimated when short-term trial results are extrapolated without correcting for short-term inefficiencies such as learning effects [150]. Given that readily available techniques to deal with these issues are lacking, one may argue that it cannot be expected from applied economic evaluations that they deal with or correct for this issue in a quantitative fashion. However, the relevance of (the potential influence of on outcomes of) learning effects also mostly remained undiscussed, which could result in unawareness about these issues among policy makers, and lead to an overestimation of the validity of the reported ICER by them. As an illustration of the potential impact of this issue, a combination of strong confidence with an overestimated ICER may result in rejecting an intervention that might be cost-effective in the longer run. Moreover, discussing the potential discrepancy between short-term and long-term efficiency may also avoid disappointment with short-term results after implementation of the intervention [151]. Second, none of the studies provided or applied an explicit analytical solution to cope with incremental innovation which may influence outcomes, and numerous studies did not discuss this aspect. However, multiple innovations which influenced TAVI outcomes have occurred, for examples new generations of valves and new strategies for procedure optimization were introduced [152, 153]. Furthermore, new innovations, including those concerning alternative access routes, are expected. Such incremental innovations may be relevant for policy makers. As an illustration, one could consider the extremely divergent cost-effectiveness outcomes of the transfemoral (dominant) and the transapical access route (dominated) reported within a single study [140]. As a consequence of incremental innovation, reported ICERs may be especially relevant in the short-term. This aspect often is not mentioned

explicitly, and remained undiscussed in almost half of the studies. It is clear that one should try to avoid evaluations reporting on already obsolete technologies or application procedures to inform reimbursement decisions that do not pertain (only) to the studied interventions but also those currently in place. Policy makers therefore need to be aware of this, to avoid suboptimal reimbursement decisions. Awareness of incremental innovation may lead policy makers to apply a more adaptive approach to health technology assessment [109] which may help in dealing with this challenge. Third, except for one of the reviewed studies (which highlighted a scenario analysis for other countries) [138] none of the studies provided or used an analytical solution approach to cope with the dependence of outcomes on the context of care provision, and most did not discuss this dependency. Nonetheless, context dependency of outcomes is of relevance for TAVI; e.g. hospitals differ in their mix of access routes, in devices used, and in operation settings [154], which are elements affecting the 'local ICER'. For example, Ribera et al. presented ICERs for both major valve manufacturers (Edwards Lifesciences and Medtronic) separately, suggesting differences between these ICERs [138]. It can be argued that the challenge of context dependency has been mitigated to a certain extent by using parameters from the PARTNER trial as these are based on multiple centres. However, although PARTNER was a multicentre RCT, it was limited to valves of Edwards Lifesciences. Moreover, reported average ICERs may still not be valid for all contexts and in centres with other characteristics than the included ones. This limitation mostly remained undiscussed, potentially leaving policy makers unaware of risks in generalising the results of the studies to the context of the relevant policy question at hand. However, the potential policy relevance of this issue may be illustrated by considering the different scenario's reported by Ribera et al. [138], ranging from dominated (a policy argument to reject reimbursement) to dominant (a policy argument to allow reimbursement).

Taken together, the distinctive features of medical devices result in methodological challenges which were typically not accounted for in economic evaluations of TAVI. As a result, dealing with these challenges is, mostly implicitly, passed on to policy makers. When policy makers are unaware of these challenges, they may overestimate the relevance of reported cost-effectiveness results for their decision context. This could result in non-optimal decisions regarding funding these technologies or to a lack of additional information gathering to come to more relevant and up-to-date estimates of cost-effectiveness.

It could be suggested that the three methodological challenges were omitted in the included economic evaluations because of a presumed small impact on the ICER. However, this would require a quantification of their impact which was not provided within these evaluations. Also, some of the examples above suggest that their influence can indeed be substantial.

Exploring the impact on ICERs of dealing with (any of) the three challenges is hampered by the fact that only one study reported handling learning effects [144]. This study reported an incremental effect (-0.04 QALY) slightly below the average (0.02 QALY) but falling well within the range of incremental effects (-0.61 QALY to 0.23 QALY) reported in the included studies. The incremental costs reported in this study also fell within the range of reported values.

It should be noted that it may be unreasonable to expect individual economic evaluations to find and use technical solutions for the fundamental and complicated challenges highlighted here, without clear guidance how to do so. Although ready to use technical solutions may not be available, in current international methodological guidance on the assessment of medical devices [e.g. [120]] and in national HTA guidelines

(e.g. England, France, the Netherlands, and Sweden) [122] the specific methodological challenges are extensively acknowledged. Consequently, it could be reasonably expected that studies would at least mention these challenges and particularly their potential impact on the results, also to inform policy makers who may use the results of studies. Reporting study details on the level of operators' experience, organisational context and interventions, would allow policy makers to judge their similarity with health provision in their own context.

While our results suggest some improvement over time, they also show that still not all current studies mention these challenges and their potential impact. To stimulate further improvement, policy makers could enforce submitters of economic evaluations to specify how they handled specific methodological challenges of the intervention concerned. Moreover, future research could contribute to the further development of methodology dealing with these challenges, and the development of best practices to illustrate how to do so in economic evaluations.

As mentioned, the consideration of learning effects and incremental innovation in economic evaluations of TAVI has been subject of previous research. Tarricone and colleagues [130] showed in their review among other results that a minority of HTA-reports and journal articles on TAVI considered "learning curves" (42 percent of included publications) and "incremental innovation" (37 percent of included publications). Our results were in line with their results, showing moderate improvement over time in terms of the consideration of these challenges. Based on their results, combined with comparable results for economic evaluations of implantable cardioverter defibrillators (ICD), they concluded that the general awareness of specific features of medical devices is low in the context of economic evaluation. Our review confirms their conclusion, despite the developments in this field since their study, including the publication of specific guidance.

Hence, more effort is needed to increase the awareness about these challenges, their explicit mentioning in economic evaluations, and the availability of methodological techniques to deal with these issues.

As an additional observation, acknowledging the low number of included studies, an upward trend appears to be observed over time, in terms of the reported incremental QALY's gained. This potential trend may suggest a relative improvement of TAVIs effectiveness over time. However, it needs to be noted that the fifteen studies included in our review differed in terms of the risk class of their target population as well as the applied time horizon (ranging from 1 year to lifetime). Such differences warrant caution in the interpretation of these effectiveness results.

5.4.1 Limitations of this review

A number of limitations of this review deserve mentioning. First, this review only dealt with one particular medical device: TAVI. Hence, generalisations to other medical devices cannot be made, especially since medical devices consist of a large and heterogeneous collection of technologies [155]. For example, while TAVI is an artificial body part implanted by a medical procedure, other devices may concern assistive devices directly used by patients. In the latter category, in contrast to TAVI, a learning curve on patient side may be expected. For diagnostic technologies other methodological challenges may apply compared to therapeutic technologies like TAVI. Finally, for pragmatic reasons the search for this study was limited to one digital database (Pubmed) although several other digital databases (e.g. Embase, Web of Science) are available. However, given that the identified systematic reviews and the study of Tarricone et al. did not include peer reviewed studies that did not show up in our results, this suggests our search strategy was quite adequate in retrieving relevant studies.

5.5 Conclusion

The challenges related to economic evaluations of medical devices and their influence on the validity of reported cost-effectiveness results, are typically discussed incompletely and rarely dealt with in peer reviewed studies on TAVI. It is important for research and policy that this improves. Best practices should be developed to support the application of technical solutions, and policy makers should require submitters to at least reflect on specific methodological challenges of the intervention concerned.

Table 5.1 Study characteristics (1/2)

Author	**Country**	**Target population**	**Interventions**	**Comparator**	**Analytic approach**
Reynolds et al. (2012)	USA	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	Trial based
Neyt et al. (2012)	Belgium	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	Model based
Gada, Kapadia, Tuzcu, Svensson, & Marwick (2012)	USA	high surgical risk	TAVI (TF) Edwards SAPIEN	SAVR	Model based
Gada, Agarwal, & Marwick (2012)	USA	high surgical risk	TAVI (TA) Edwards SAPIEN	SAVR	Model based
Sehatzadeh et al. (2012)	Canada	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	Model based
Doble et al. (2013)	Canada	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	Model based
Fairbairn et al. (2013)	UK	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	Model based
Sehatzadeh et al. (2013)	Canada	high surgical risk	TAVI (TF and TA) Edwards SAPIEN	SAVR	*not stated*
Orlando et al. (2013)	UK	high surgical risk	TAVI Edwards SAPIEN	mixture of SAVR (90%) and medical management (10%)	Model based
Ribera et al. (2015)	Spain	intermediate surgical risk	TAVI (TF) Edwards SAPIEN Medtronic CoreValve	SAVR	Trial based
Reynolds et al. (2016)	USA	high surgical risk	TAVI Medtronic CoreValve	SAVR	Trial based
Health Quality Ontario (2016)	Canada	high surgical risk	TAVI Medtronic Corevalve	SAVR	Model based
Kodera et al. (2018)	Japan	intermediate surgical risk	TAVI (TF) Edwards Sapien XT	SAVR	Model based
Tam, Hughes, Fremes, et al. (2018)	Canada	intermediate surgical risk	TAVI (TF and TA) Edwards Sapien XT	SAVR	Model based
Tam, Hughes, Wijeysundera & Fremes (2018)	Canada	intermediate surgical risk	TAVI (TF and non-TF) Medtronic CoreValve Medtronic Evolut R	SAVR	Model based

TAVI = Transcatheter Aortic Valve Implantation, TF = transfemoral , TA = transapical, SAVR = Surgical Aortic Valve Replacement, n/a = not applicable

Table 5.1 Study characteristics (2/2)

Author	Time horizon	Efficacy source	Cost source	Perspective	Discount-ing
Reynolds et al. (2012)	1 year	PARTNER A	US hospital billing and resource based accounting	US healthcare system	n/a
Neyt et al. (2012)	1 year	PARTNER A Continued Acces study (non-published)	Belgian hospital billing data (n=183, treated with Edwards SAPIEN valve)	Belgian healthcare payer	n/a
Gada, Kapadia, Tuzcu, Svensson, & Marwick (2012)	lifetime	Published reports, registries, European PARTNER	Reimbursement data, primarily published reports (DRG, Medicare payments).	Healthcare provider	5%
Gada, Agarwal, & Marwick (2012)	lifetime	Registries	Reimbursement data, registries, DRGs. Medicare payments	Healthcare funding body	5%
Sehatzadeh et al. (2012)	20 years	PARTNER	Ontario Case Costing Initiative (OCCI) cost data	Canadian healthcare payer	5%
Doble et al. (2013)	20 years	PARTNER US	Ontario Case Costing Initiative	Canadian healthcare payer	5% (costs)
Fairbairn et al. (2013)	10 years	Utility data from a UK high-risk AS population PARTNER A	UK costs UK care pathway	UK National Health Service	3.5%
Sehatzadeh et al. (2013)	*not stated*	2-year follow-up of the PARTNER trial	Ontario Case Costing Initiative (OCCI) cost data	Canadian healthcare payer	5%
Orlando et al. (2013)	25 years / lifetime	PARTNER B	Reference prices and literature	UK National Health Service	3.5%
Ribera et al. (2015)	1 year	Collected within study	Collected within study, cost accounting, reimbursement tarrifs	Spanish health service	n/a
Reynolds et al. (2016)	lifetime	CoreValve U.S. High Risk Pivotal Trial	CoreValve U.S. High Risk Pivotal Trial (resource utilization, hospital billing data)	US Healthcare system	3%
Health Quality Ontario (2016)	5 years	U.S. CoreValve Pivotal Trial	Ontario Case Costing Initiative	Canadian healthcare payer	5%
Kodera et al. (2018)	10 years	PARTNER 2 cohory A Optimizes Catheter vAlvular	Previous studies, and estimations	Japanese public healthcare payers	2%

Author	Time horizon	Efficacy source	Cost source	Perspective	Discount-ing
		iNtervention (OCEAN) TAVI registry			
Tam, Hughes, Fremes, et al. (2018)	lifetime	PARTNER 2 cohory A Optimizes Catheter vAlvular iNtervention (OCEAN) TAVI registry	Canadian Institue of Health Information, Ontario Schedule of Benefits / literature review	Canadian healthcare payer	1.5%
Tam, Hughes, Wijeysundera & Fremes (2018)	lifetime	SURTAVI trial / CoreValve US High Risk Pivotal Trial (EQ-5D)	Canadian Institute of Health Information	Canadian healthcare payer	1.5%

TAVI = Transcatheter Aortic Valve Implantation, TF = transfemoral , TA = transapical, SAVR = Surgical Aortic Valve Replacement, n/a = not applicable

Table 5.2 Reported cost-effectiveness outcomes

Author	Incremental effect (QALY)	Incremental costs	ICER	Sensitivity analysis
Reynolds et al. (2012)	TAVI-TF: 0.068 TAVI-TA: -0.07 TAVI Overall: 0.027	TAVI-TF: $ -1.250 TAVI-TA: $ 9.906 TAVI Overall: $ 2.070	TAVI-TF dominates SAVR TAVI-TA is dominated by SAVR TAVI overall $76.877/QALY	Bootstrapping (and boundry testing valve pricescenarios)
Neyt et al. (2012)	0.03	€20,400	> €750,000/QALY	PSA and one-way sensitivity analysis
Gada, Kapadia, Tuzcu, Svensson, & Marwick (2012)	0.06	$3,164	$52,773/QALY	threshold analyses, one-way 2-way sensitivity analyses, PSA
Gada, Agarwal, & Marwick (2012)	-0.04	$ 100	TA-TAVI is dominated by SAVR	PSA and one-way sensitivity analysis
Sehatzadeh et al. (2012)	-0.102	CAN $11,153	TAVI is dominated by SAVR	Comprehensive sensitivity analyses
Doble et al. (2013)	-0.102	CAN $11,153	TAVI is dominated by SAVR	DSA, PSA and scenario analyses
Fairbairn et al. (2013)	0.06	£-1,350	TAVI dominates SAVR	Deterministic and PSA
Sehatzadeh et al. (2013)	-0.069	CAN $ –4,642	CAN $66,985/QALY	3-way deterministic sensitivity analyses

Author	Incremental effect (QALY)	Incremental costs	ICER	Sensitivity analysis
Orlando et al. (2013)	-0.6087	£7,963	TAVI not available option dominates the TAVI available option/ patients suitable for SAVR	PSA and DSA (one-way sensitivity analysis)
Ribera et al. (2015)	0.036 (Edwards) -0.011 (Medtronic)	€8,800 (Edwards) €9,729 (Medtronic)	€148,525 (Edwards) Dominated by SAVR (Medtronic)	Bootstrapping
Reynolds et al. (2016)	0.32	$17,849	$55,090/QALY	One-way sensitivity analysis, bootstrapping
Health Quality Ontario (2016)	0.181	CAN $9,412	CAN $51,988/QALY	one-way and probabilistic sensitivity analyses, as well as scenario analyses.
Kodera et al. (2018)	0.22	Y 1,723,516	Y 7,523,821/QALY	PSA, one-way sensitivity analysis/DSA, threshold analysis
Tam, Hughes, Fremes, et al. (2018)	0.23	CAN $10,548	CAN $46,083/QALY (TF-TAVI CAN $24,790/QALY)	one-way DSA, PSA
Tam, Hughes, Wijeysundera, & Fremes (2018)	0.15	CAN $11,305	CAN $76,736/QALY	one-way DSA, PSA

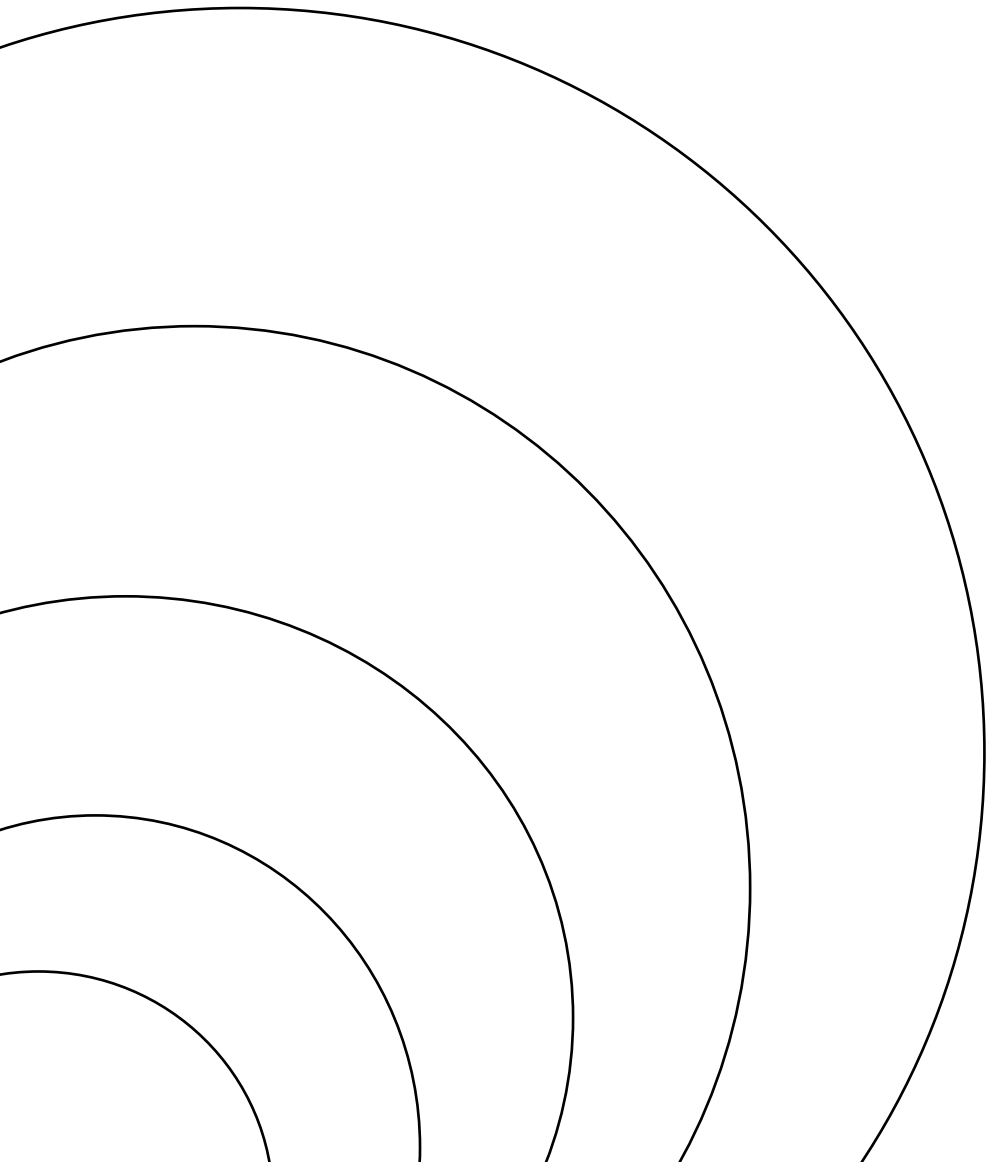
TAVI = Transcatheter Aortic Valve Implantation, TF = transfemoral , TA = transapical, SAVR = Surgical Aortic Valve Replacement, PSA = Probabilistic Sensitivity Analysis, DSA = Deterministic Sensitivity Analysis, ICER = Incremental Cost Effectiveness Ratio, QALY = Quality-adjusted Life Year

Table 5.3 Discussed challenges (1 = discussed, 0 = undiscussed)

Author	**Learning effects (discussed)**	**Incremental innovation (discussed)**	**Context dependence of results (discussed)**
Reynolds et al., 2012 [140]	1	1	1
Neyt et al., 2012 [141]	0	0	0
Gada, Kapadia, et al., 2012 [137]	0	1	0
Gada, Agarwal, & Marwick, 2012 [144]	1	0	0
Sehatzadeh et al., 2012 [135]	0	0	0
Doble et al., 2013 [156]	0	1	1
Fairbairn et al., 2013 [157]	0	0	1
Sehatzadeh et al., 2013 [134]	0	0	0
Orlando et al., 2013 [143]	1	0	0
Ribera et al., 2015 [138]	0	0	1
Reynolds et al., 2016 [139]	1	1	1
Health Quality Ontario, 2016 [158]	1	1	0
Kodera, Kiyosue, Ando, & Komuro, 2018 [159]	0	1	1
Tam, Hughes, Fremes, et al., 2018 [142]	0	0	0
Tam, Hughes, Wijeysundera, & Fremes, 2018 [145]	0	1	1
Total	**5**	**7**	**7**

Chapter 6
Challenges for outcome measurement of mental healthcare

Based on: Enzing JJ, van Krugten FCW, Sabat I, Neumann-Böhme S, Boer B, Knies S, Brouwer WBF; ECOS consortium. Psychometric evaluation of the Mental Health Quality of Life (MHQoL) instrument in seven European countries. Health Qual Life Outcomes. 2022 Sep 1;20(1):129. doi: 10.1186/s12955-022-02041-6. PMID: 36050766; PMCID: PMC9434504.



Abstract

To make efficient use of available resources, decision-makers in healthcare may assess the costs and (health) benefits of health interventions. For interventions aimed at improving mental health capturing the full health benefits is an important challenge. The Mental Health Quality of Life (MHQoL) instrument was recently developed to meet this challenge. Evaluating the pyschometric properties of this instrument in different contexts remains important.

A psychometric evaluation of the MHQoL was performed using existing international, cross-sectional data with 7,155 respondents from seven European countries (Denmark, France, Germany, Italy, Portugal, The Netherlands and the United Kingdom). Reliability was examined by calculating Cronbach's alpha, a measure of internal consistency of the seven MHQoL dimensions, and by examining the association of the MHQoL sum scores with the MHQoL-VAS scores. Construct validity was examined by calculating Spearman's rank correlation coefficients between the MHQoL sum scores and EQ-5D index scores, EQ-VAS scores, EQ-5D anxiety/depression dimension scores, ICECAP-A index scores and PHQ-4 sum scores.

The MHQoL was found to have good internal consistency for all seven countries. The MHQoL sum score and the MHQoL-VAS had a high correlation. Spearman's rank correlation coefficients were moderate to very high for all outcomes.

Our results, based on data gathered in seven European countries, suggest that the MHQoL shows favourable psychometrical characteristics. While further validation remains important, the MHQoL may be a useful

instrument in measuring mental health-related quality of life in the Western European context.

6.1 Introduction

Worldwide, more than one billion people are affected by mental health problems. The burden of disease related to mental health problems is very large. For instance, the disability-adjusted life-years (DALYs) lost due to such problems represent seven percent of the total global burden of disease. [160] Additionally, mental health problems are known to have a major economic impact. This can be illustrated by the worldwide costs of lost productivity related to depression and anxiety, the two most commons mental health problems, which have been estimated to amount to US$ 1 trillion annually (1.6% of worldwide GDP). [161] Such figures highlight the importance of effectively preventing and treating mental health problems.

However, healthcare systems have limited resources to improve the health of the populations they serve. Interventions aimed at preventing or treating mental health problems, even effective ones, in that context compete for scarce resources with interventions aimed at other diseases. Therefore, healthcare decision-makers are confronted with difficult allocation decisions. They may wish to prioritize those interventions that contribute most to the goals of the healthcare system, including using the available resources efficiently. In this context, economic evaluations are increasingly used, in which the benefits and costs of interventions (relative to some relevant comparator) are assessed in order to determine whether they are cost-effective. Such evaluations are most commonly and systematically used in relation to reimbursement decisions regarding pharmaceuticals. However, they are also increasingly applied to other types of interventions, but this broadening of the application of economic evaluations in other contexts comes with specific challenges.[58] Economic evaluations can and are also sometimes used to investigate the cost-effectiveness of mental health interventions, with the aim of informing decisions regarding their reimbursement. For mental health interventions an important challenge is

capturing all relevant benefits related to these interventions. In conventional economic evaluations, health benefits are typically measured using a generic health-related quality-of-life measure, such as the EQ-5D [162] or SF-6D [163]. These instruments are prescribed in many guidelines for economic evaluations of health interventions, including the UK[164] and Dutch guidelines[92]. However, it has been questioned whether these commonly used generic health-related quality of life instruments capture all relevant quality of life domains impacted by mental health problems and interventions aimed at improving mental health.[36] This is problematic since it could lead to inaccurate estimations of health benefits related to mental health interventions, hampering well-informed decisions. Such decisions require that the quality of life impact of mental health interventions are captured accurately and completely.

In an attempt to overcome this problem, the Mental Health Quality of Life instrument (MHQoL) was developed.[165] The MHQoL is a self-report, seven-item questionnaire that captures and values dimensions relevant to the quality of life of people with mental health problems, such as self-image, independence and hope. While the instrument was based on previous research highlighting these relevant quality of life dimensions, its psychometric properties, like feasibility, reliability and validity, also need to be demonstrated. This is even more important if healthcare decisions are informed by economic evaluations using the MHQoL as outcome measure. The psychometric properties of the MHQoL so far have only been examined in the Dutch context.[165] This evaluation was performed in a sample of 110 members of the Dutch general public as well as a sample of 479 Dutch mental healthcare users. The results of this study suggested that the MHQoL is a promising instrument, demonstrating favourable psychometrical properties. However, further psychometric evaluation in different populations and contexts remains warranted. This is also true for

psychometrical evaluation in an international context. Such evaluation may be considered relevant, not only to investigate the performance of translated versions of the MHQoL, but also given that cultural differences may impact how mental health problems may be experienced, evaluated and perceived.[166]

The objective of this study, therefore, is to evaluate the psychometric characteristics of the MHQoL using a large dataset obtained in seven European countries. Specifically, we will investigate the reliability of the MHQoL by examining its internal consistency as well as the construct validity of the MHQoL by investigating the association of the MHQoL sum scores with other validated outcome measures.

6.2 Methods

6.2.1 Data source

For this study, we used cross-sectional data obtained in the fourth survey wave of the European COvid Survey (ECOS) project, which is described in detail elsewhere [167]. Generally, this online survey examined support for COVID-19 containment policies, including vaccinations, worries about COVID-19, and trust in different information sources. The data in the fourth wave of this survey was obtained between 5 and 16 November 2020 . Respondents (n=7,115) were recruited from the general public in seven European countries (Denmark, France, Germany, Italy, Portugal, the Netherlands, and the United Kingdom) by the market research company Dynata using multisource online panels. To ensure that the sampling frame was representative of the population in each country, the company used various recruiting procedures for different subgroups of the population in each country. It used for example advertised/open recruitment, loyalty programs, affiliate networks and mobile apps [167]. Quotas based on age category, regional distribution and gender were implemented by the

authors using the Qualtrics research suite to ensure and control the representativeness based on the country specific census data. Dynata ensured representativeness with regard to educational categories based on their expertise in the differences in educational degrees for each country. The authors proceeded by excluding incompletes answers and speeders (faster than 1/3 of the median time in each country), both of which were replaced by Dynata to ensure the representativeness of the sample. The resulting sample of respondents from each country (with n~1,000) was representative of its adult population in terms of region, gender, age group and education level.

The questionnaire was available in the seven languages of the included countries. The MHQoL had existing official versions in Dutch, English and German, which were used in this survey. For the other four countries and languages, the MHQoL instrument was translated by native speakers with a background in health economics.

Respondents completed the MHQoL instrument along with the EuroQol (EQ-5D-5L and EQ-VAS), ICECAP-A (the ICEpop CAPability measure for Adults), and PHQ-4 (Patient Health Questionnaire for Depression and Anxiety) instruments. For these instruments, official translations available from the developers of these instruments were used. Respondents also were asked to answer questions about their demographic characteristics including gender, age, relationship status, and level of education, next to COVID-related questions.

6.2.2 Outcome measures

The MHQoL is an instrument intended to be used to describe and value respondents' current mental health-related QoL.[165] In the descriptive part, respondents are asked to describe their mental health state using seven specific dimensions of mental health-related quality of life and four

answering levels per dimension. The seven dimensions are: self-image, independence, mood, relationships, daily activities, physical health and hope. Levels for self-image for example range from 'I think very positively about myself' to 'I think very negatively about myself'. Preference-based tariffs, allowing scores on the different levels to be converted into utility scores anchored on 0 (dead) and 1 (full mental health-related QoL), are not yet available for the MHQoL. In the absence of tariffs, the MHQoL sum score is used as an alternative, which ranges from 0 (lowest level on all seven dimensions) to 21 (highest level on all seven dimensions), with higher scores indicating better mental health-related quality of life. Next to the descriptive part, the MHQoL instrument also has a direct valuation part in which respondents are asked to rate their psychological wellbeing using a horizontal visual analogue scale (MHQoL-VAS) ranging from 0 (representing 'worst imaginable psychological wellbeing') to 10 (representing 'best imaginable psychological wellbeing').

The five-level EQ-5D (EQ-5D-5L) is a generic instrument to measure and value respondents' current health-related QoL. [162] Within the questionnaire, respondents are asked to describe their health using five dimensions and five answering levels per dimension. These five dimensions are: Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. The five answering levels range from having no problems to having extreme problems. Using country-specific, preference-based tariffs, answer scores can be converted into utility scores, with 0 as the equivalent of the state 'dead' and 1 as the equivalent of the state 'perfect health'. In addition to scoring the five dimensions, respondents are asked to rate their current overall health using a vertical visual analogue scale (EuroQol Visual Analogue Scale; EQ-VAS) ranging from 0 ('worst imaginable health state') to 100 ('best imaginable health state').

The ICE-CAP-A (ICEpop CAPability measure for Adults) is an instrument to measure and value respondents' overall capability wellbeing and is grounded in Sen's capability approach. [168] [35] Within the questionnaire, respondents are asked to describe their capabilities in relation to five important life domains: stability, attachment, autonomy, achievement and enjoyment. Each domain is scored using four levels, ranging from full capability to no capability. Using preference-based tariffs, answer scores can be converted to standardised index scores, ranging from 0 (no capability) to 1 (full capability). Currently, tariffs for the United Kingdom [168] and the Netherlands [169] are available.

Finally, the PHQ-4 is a four-item self-complete screening instrument that measures respondents' likeliness of an anxiety disorder and/or depression.[170] The PHQ-4 is based on the Patient Health Questionnaire (PHQ), a more extensive instrument used by care providers to diagnose patients with mental health disorders. Dimensions are "Feeling nervous, anxious or on edge", "Not being able to stop or control worrying", "Feeling down, depressed or hopeless" and "Little interest or pleasure in doing things". The PHQ-4 has four levels which range from "Not at all" to "Nearly every day". A sum score can be calculated, which can be used to categorise the respondent's psychological distress level as none (0-2), mild (3-5), moderate (6-8), or severe (9-12).

6.2.3 Statistical Analysis

Using descriptive statistics, the basic characteristics of the respondents were summarized. Mean MHQoL sum score and mean MHQoL-VAS score were compared to scores known for the Dutch general population.[165] For subgroups based on country, gender, and age group, MHQoL sum scores were calculated, also to provide a reference for future studies. In doing so, we do acknowledge and stress the exceptional situation due to the COVID-

19 pandemic.

We additionally investigated the MHQoL dimension scores, comparing the youngest and oldest age groups. A lower and upper MHQoL quartile were distinguished using cut-off values (respectively <12 and >16) for the MHQoL sum score. We investigated the membership of these quartiles focusing on background characteristics of the respondents including income, mean, minimum and maximum values of the four other outcome measures. EQ-VAS mean scores per country were compared to population norm scores, also given the fact that the survey was conducted during the COVID-19 pandemic.

To examine the internal consistency of the MHQoL, Cronbach's alpha of the seven dimensions was calculated, both overall and by country. Cronbach's alpha is a coefficient that represents the extent to which items of a measure are correlated, indicating the extent to which these items measure the same construct, here mental health-related quality of life. Cronbach's alpha is expressed as a number between 0 (no correlation) and 1 (full correlation), and a score of > 0.7 is seen as indicating good internal consistency[171]. In addition, Spearman's rank correlation coefficients of the MHQoL sum score and the MHQoL-VAS were calculated, again overall and by country. This correlation was expected to be high and positive since both aim to measure mental health-related QoL. When interpreting results, a correlation coefficient is seen as trivial when < 0.1, as small when 0.1-0.3, as moderate when 0.3-0.5, as high when 0.5-0.7, as very high when 0.7-0.9, and as nearly perfect when > 0.9, following previous evaluation studies like Hoefman et al. [172] Additionally, the association between the MHQoL-VAS and the MHQoL dimension scores was investigated using a linear regression model (ordinary least squares regression; OLS). A moderate to high positive correlation and a moderate to high adjusted R^2 were expected since these

dimensions levels represent mental health states which determine overall mental health-related QoL. An adjusted R^2 of > 0.2 was expected, based on previous models which modelled EQ-VAS as the dependent variable and the levels of the EQ-5D dimensions as independent variables. [173] [174]

To examine convergent validity, Spearman's rank correlation coefficients between the MHQoL sum score and respectively: EQ-5D index score, EQ-VAS, EQ-5D anxiety/depression dimension score, ICECAP-A index score, and PHQ-4 sum score were calculated, overall and by country. The resulting coefficients inform whether these measures of theoretically interrelated constructs are correlated, which is an indication of construct validity. We expected the MHQoL sum score to have a moderate to strong positive correlation (0.3-0.7) with the EQ-5D index score, the EQ-VAS, and the ICECAP-A index score. This was based on the assumption that having a better mental health-related QoL is associated with both a better health-related QoL and a better wellbeing. It is acknowledged that these are complex associations, e.g. mental health has direct and indirect effects (e.g.by life style choices) on physical health and vice versa.[175]

Furthermore, we expected the MHQoL sum score to have a strong negative correlation (0.5-0.7) with the EQ-5D anxiety/depression dimension and with the PHQ-4 sum score, since a better mental health-related QoL is strongly related to the absence of mental health problems. Tariffs to compute EQ-5D-5L index scores were obtained from the EuroQol website [176]. Tariffs to compute ICECAP-A index scores were obtained from the website of the University of Birmingham.[168] United Kingdom (UK) tariffs were used for EQ-5D-5L and ICECAP-A for all countries. This was done given that tariffs were not available for all countries. Since the current UK tariffs for EQ-5D-5L have been disputed[177] the analyses were also performed using Dutch EQ-5D-5L tariffs[178] to check whether this would influence results.

Beforehand, we did not formulate expectations regarding mean MHQoL sum scores in different countries although they are expected to differ, e.g. based on country differences in depression stigma [179].

All analyses were performed using STATA version 16.1 (StataCorp, 2019. Stata Statistical Software: Release 16. College Station, TX).

6.3 Results

6.3.1 Descriptive statistics

Table 6.1 shows the general characteristics of our respondents. About half of the respondents was female (52%), a third of the respondents lived alone (29%), and a majority of respondents were of working age (18-64 years; 79%). The mean MHQoL sum score was 14.1 (SD ±3.8), which is below the MHQoL sum score reported for the Dutch general population (15.5; SD ±2.9)[165]. The mean MHQoL-VAS was 6.6 (SD ±2.2), which is below the mean MHQoL-VAS reported for the Dutch general population (7.5; SD ±1.5)[165]. For the dimension scores (scale 0 to 3; higher is better), means ranged from 1.8 (Hope) to 2.2 (Physical health).

Table 6.1 Respondent characteristics (n = 7,115)

	%	N
Gender		
Female	52.0	3,699
Age category		
18-24	9.9	707
25-34	15.8	1,125
35-44	18.9	1,343
45-54	18.7	1,332
55-64	16.1	1,145
65+	20.6	1,463
Relationship status		
Married / registered partnership	48.0	3,416
Living together (relationship)	14.4	1,027
Living alone (single)	24.0	1,710

	%	N
Living alone (in a relationship)	4.5	319
Widow / widower	3.0	212
Other	6.1	431
Country		
Germany	14.7	1,043
United Kingdom	14.1	1,006
Denmark	14.2	1,012
The Netherlands	14.3	1,020
France	14.3	1,017
Portugal	14.3	1,015
Italy	14.1	1,002
Outcome measures	**Mean (Min, Max)**	**Standard deviation**
MHQoL sum score[a]	14.1 (0,21)	±3.8
MHQoL-VAS[a]	6.6 (0,10)	±2.2
EQ-VAS	72.3 (0,100)	±23.3
EQ-5D-5L UK tarrif	0.85 (-0.285,1)	±0.2
ICECAP-A UK tarrif	0.78 (-0.001,1)	±0.2
PHQ-4 sum score[b]	3.4 (0,12)	±3.3
MHQoL dimension score[a]	**Mean (Min, Max)**	**Standard deviation**
Self-image	1.9 (0,3)	±0.7
Independence	2.0 (0,3)	±0.8
Mood	2.2 (0,3)	±0.8
Relationships	2.1 (0,3)	±0.8
Daily activities	1.9 (0,3)	±0.8
Physical health	2.2 (0,3)	±0.8
Hope	1.8 (0,3)	±0.8

[a] Higher is better mental health-related QoL. [B] Lower is better mental health.
Note: Min = minimum value, Max = maximum value, MHQoL = Mental health quality of life, EQ-VAS = EuroQol Visual Analogue Scale, EQ-5D-5L = EuroQol 5 dimensions 5 levels, ICECAP-A = ICEpop CAPability measure for Adults , PHQ = Patient Health Questionnaire-4, UK = United Kingdom.

The observed mean MHQoL sum score differed somewhat between subgroups, see Table 6.2. This Table also shows that the mean MHQoL sum score increased with age.

Table 6.2 Mean MHQoL scores by country, gender and age group

	Mean MHQoL sum score[a]	**Standard deviation**
Overall	14.1	±3.8
Country		
Germany	14.5	±3.6
United Kingdom	14.0	±4.4
Denmark	14.7	±3.7
the Netherlands	14.9	±3.7
France	13.3	±3.8
Portugal	14.1	±3.4
Italy	13.3	±3.8
Gender		
Male	14.6	±3.7
Female	13.7	±3.9
Age category (years)		
18-24	13.4	±3.9
25-34	13.9	±4.0
35-44	13.9	±4.0
45-54	13.8	±3.9
55-64	14.4	±3.6
65+	14.9	±3.5

[a] Minimum 0, Maximum 21. A higher MHQoL sum score means a better mental health-related QoL.

Mean MHQoL sum scores, by country and age group, are presented in Table 6.3. These scores ranged from 12.3 in young adults (18-24 years) in the UK, to 15.9 in Danish elderly (65+ years).

Table 6.3. Mean MHQoL sum score[a] by country and age group

Age group (years)	Germany	United Kingdom	Denmark	The Netherlands	France	Portugal	Italy
18-24	13.9	12.3	14.0	13.6	12.8	13.6	12.5
25-34	14.5	13.4	14.6	14.6	13.2	14.2	12.8
35-44	14.1	13.5	14.3	14.7	13.3	14.1	13.6
45-54	14.5	13.6	14.0	14.6	13.0	14.0	12.9
55-64	14.4	14.8	15.0	15.4	13.6	14.2	13.4
65+	15.0	15.4	15.9	15.7	13.8	14.8	13.9
Overall	*14.5*	*14.0*	*14.7*	*14.9*	*13.3*	*14.1*	*13.3*

[a] Minimum 0, Maximum 21. Higher is better mental health-related QoL.

Given the observed increase in MHQoL with age, we decomposed the MHQoL sum score into domain scores for the youngest (18-24 years) and oldest (65+ years) age groups, between which the difference was the largest (1.6 points) (Additional file 1 in Supplementary Information found at https://doi.org/10.1186/s12955-022-02041-6). The youngest group scored worse than the oldest group on all dimensions except for Physical health (+0.2), while especially their scores on the dimensions Mood (-0.6) and Independence (-0.5) were relatively low.

Characteristics and mean measure scores are presented in Table 6.4 contrasting the respondents with a high (>16) MHQoL score and those with a low (<12) MHQoL scores. Compared to those in the high group, respondents in the low group were younger, less often in a relationship, more often female, had lower wellbeing scores, and a lower health-related QoL.

Table 6.4 Descriptive statistics of the highest and lowest MHQoL quartiles

	High MHQoL score group[a] (SD) (Min, Max)	**Low MHQoL scores group[b] (SD) (Min, Max)**
Female	47%	61%
Age (mean)	50 (±16.5) (18,90)	44 (±15.6) (18,85)
% Household's total monthly income measured as being able to make ends meet (fairly) easily.	76%	33%
Relationship status		
Married / registered partnership	56%	36%
Living together (relationship)	15%	15%
Living alone (single)	19%	33%
Age category		
18-24	7%	14%
25-34	16%	18%
35-44	18%	20%
45-54	17%	21%
55-64	17%	13%
65+	26%	14%
Measures (mean)		
MHQoL sum score	18.46 (±1.4) (17,21)	8.58 (±2.6) (0,11)
EQ-VAS	83 (±19.6) (0,100)	57 (±24.7) (0,100)
EQ-5D UK tarrif	0.94 (±0.12) (-0.094,1)	0.70 (±0.22) (-0.285,1)
ICECAP-A UK tarrif	0.91 (±0.10) (0.156,1)	0.57 (±0.21) (-0.001,1)
PHQ-4 sum score	1.51 (None) (±2.5) (0,12)	6.07 (Moderate) (±03.2) (0,12)

[a] MHQoL sum score > 16 (n = 1,849), [b] MHQoL sum score < 12 (n = 1,564), SD = Standard deviation, Min = minimum value, Max = maximum value

The mean score on the EQ-VAS in our sample was 72.3 (SD ±23.3), which represents a moderate health-related QoL. For all six countries for which EQ-VAS population norm scores were available - these were not available for Portugal - the mean EQ-VAS-score was below this norm (which ranges from 76.8 in France to 83.7 in Denmark). [180] Overall, health-related QoL was relatively low compared to reference values.

In contrast to population norms [180], the mean EQ-5D scores in our sample did not decrease with age. For example, overall in the age group 18-24 the EQ-VAS score was 71.6 (SD ±25.1) while it was 73.0 (SD ±21.5), for those over 65. We also did not observe a decrease with age in mean scores for capability wellbeing (ICECAP-A scores).

6.3.2 Internal consistency

The seven dimensions of the MHQoL were found to have good internal consistency, with an overall Cronbach's alpha of 0.82. Comparable levels of internal consistency were observed for all individual countries (presented in Table 6.5), with alphas ranging from 0.78 for Portugal to 0.87 for the United Kingdom.

Table 6.5 Chronbach's alpha by country

Country	**Scale reliability coefficient**
Germany	0.81
United Kingdom	0.87
Denmark	0.82
the Netherlands	0.84
France	0.80
Portugal	0.78
Italy	0.83
Overall	*0.82*

6.3.3 MHQoL and MHQoL-VAS correlation

As expected, a significant positive correlation (0.65; $p < 0.00$) was found between MHQoL sum score and the MHQoL-VAS. For the individual countries, Spearman's rank correlation coefficient ranged from 0.53 (Italy) to 0.67 (Portugal and the United Kingdom).

The adjusted R^2 of the multivariate linear model (OLS; Table 6.6) with MHQoL-VAS as the dependent and the MHQoL dimension level scores as the independent variables was 0.43 for the total sample. For the individual countries, this ranged from 0.36 (Italy) to 0.51 (Portugal).

MHQoL-VAS scores were most strongly associated with Mood (all levels), Self-image ('very negative') and Future ('very gloomy about'). This might suggest that these will also receive the most weight in future MHQoL tariffs. Within each dimension lower levels were associated with lower MHQoL-VAS scores. The only exception were the scores associated with the third and fourth level for the domain Independence, which were almost identical. All but two coefficients were statistically significant.

Table 6.6 Multivariate regression analysis of MHQoL-VAS

	MHQoL-VAS	**Coefficient**	**95% Confidence interval**
Self-image (MHQoL1)	I think very positively about myself	Reference	
	I think positively about myself	-0.10	[-0.23, 0.03]
	I think negatively about myself	-0.68 *	[-0.85, -0.51]
	I think very negatively about myself	-1.28 *	[-1.57, -0.99]
Independence (MHQoL2)	I am very satisfied with my level of independence	Reference	
	I am satisfied with my level of independence	-0.42 *	[-0.53, -0.32]
	I am dissatisfied with my level of independence	-0.58 *	[-0.72, -0.43]
	I am very dissatisfied with my level of independence	-0.57 *	[-0.78, -0.36]
Mood (MHQoL3)	I do not feel anxious, gloomy, or depressed	Reference	
	I feel a little anxious, gloomy, or depressed	-0.95 *	[-1.05, -0.86]
	I feel anxious, gloomy, or depressed	-1.53 *	[-1.68, -1.38]
	I feel very anxious, gloomy, or depressed	-2.02 *	[-2.25, -1.78]
Relationships (MHQoL4)	I am very satisfied with my relationships	Reference	
	I am satisfied with my relationships	-0.19 *	[-0.29, -0.09]
	I am dissatisfied with my relationships	-0.48 *	[-0.62, -0.34]
	I am very dissatisfied with my relationships	-0.53 *	[-0.75, -0.30]
Daily activities (MHQoL5)	I am very satisfied with my daily activities	Reference	
	I am satisfied with my daily activities	-0.24 *	[-0.35, -0.12]
	I am dissatisfied with my daily activities	-0.44 *	[-0.59, -0.29]
	I am very dissatisfied with my daily activities	-0.86 *	[-1.09, -0.62]

Physical health (MHQoL6)	I have no physical health problems	Reference	
	I have some physical health problems	-0.19 *	[-0.28, -0.10]
	I have many physical health problems	-0.49 *	[-0.64, -0.35]
	I have a great many physical health problems	-0.85 *	[-1.08, -0.63]
Future (MHQoL7)	I am very optimistic about my future	Reference	
	I am optimistic about my future	-0.04	[-0.17, 0.10]
	I am gloomy about my future	-0.67 *	[-0.83, -0.51]
	I am very gloomy about my future ()	-0.95 *	[-1.17, -0.72]
Constant		8.70 *	[8.57, 8.82]
Adjusted R-squared		0.43	

*p < 0.001

6.3.4 Convergent validity

To examine convergent validity, an element of construct validity, Spearman's rank correlation coefficients were calculated for MHQoL sum scores with other relevant outcome measures. The resulting rank correlations are presented in Table 6.7. As expected, strong positive correlations were observed between MHQoL sum score and EQ-VAS, EQ-5D index scores and ICECAP-A index scores. The additional analysis using Dutch instead of UK EQ-5D tariffs resulted in comparable correlation coefficients (not presented). Furthermore, expected strong negative correlations were observed with the EQ-5D dimension anxiety and depression and with the PHQ4 score.

Table 6.7. Spearman's rank correlation coefficients

Examined rank correlation of MHQoL sum score	**Spearman's rho**
EQ-5D-5L dimension Anxiety and depression (Higher is worse)	
Overall	-0.63*
Germany	-0.60*
United Kingdom	-0.71*
Denmark	-0.61*
the Netherlands	-0.66*
France	-0.63*
Portugal	-0.59*
Italy	-0.60*
EQ-VAS (Higher is better QoL)	
Overall	0.49*
Germany	0.50*
United Kingdom	0.58*
Denmark	0.53*
the Netherlands	0.53*
France	0.45*
Portugal	0.43*
Italy	0.46*
EQ-5D-5L index score (UK tariffs) (Higher is better QoL)	
Overall	0.56*
Germany	0.57*
United Kingdom	0.62*
Denmark	0.56*
the Netherlands	0.62*
France	0.56*
Portugal	0.58*
Italy	0.54*
PHQ4 sum score (Higher is worse mental health)	
Overall	-0.60*
Germany	-0.54*
United Kingdom	-0.73*
Denmark	-0.45*
the Netherlands	-0.65*
France	-0.58*
Portugal	-0.57*
Italy	-0.60*
ICECAP-A index score (UK tariffs) (Higher is better wellbeing)	
Overall	0.68*
Germany	0.64*
United Kingdom	0.75*

Denmark	0.70*
the Netherlands	0.71*
France	0.64*
Portugal	0.67*
Italy	0.64*

*p < 0.001

6.4 Discussion

In this chapter, we presented a psychometric evaluation of the MHQoL questionnaire in seven European countries. This study, therefore, represents one of the first psychometric evaluations of the MHQoL and the first international one to our knowledge. Importantly, this builds on the evidence of the performance of the MHQoL as a reliable and valid measure of mental health-related QoL across Western European countries.

Specifically, we examined reliability by investigating internal consistency and construct validity by investigating convergent validity, using existing survey data obtained through the ECOS survey, which was conducted in seven European countries. Overall as well as for the separate counties, we found good internal consistency between the dimensions. Additionally, the MHQoL-VAS score, in general, was significantly associated with the MHQoL sum scores and domain scores in the expected way. Convergent validity was investigated by investigating the correlations between the MHQoL instrument with the EQ-5D, ICECAP-A as well as specific mental health instruments and showed favourable results. Overall our results suggest that in the investigated countries, the MHQoL appears to be a psychometrically sound measure of quality of life in people with mental health problems, which highlights it is a promising instrument to use and validate further.

The importance in doing so may be emphasised by the fact that sound and concise instruments, capable of adequately capturing important life domains impacted by mental health problems, are currently lacking.[33] Instruments like the MHQoL, therefore, can facilitate broadening the

application of HTA to the field of mental health in a way that the benefits of mental health interventions are indeed captured. This may ultimately lead to better (informed) decision-making and, therefore, a more efficient and equitable allocation of resources.

6.4.1 Limitations of the study

The main strengths of our study are its large sample size, the representativeness and international character of its sample, the inclusion of several relevant outcome measures, and the evaluation of multiple psychometric properties. Besides these strengths, several limitations need to be acknowledged. First, the psychometric evaluations performed were not without limitations. Given the nature of the ECOS survey, we did not have any information on clinical diagnosis of respondents. Consequently, we could not evaluate whether MHQoL scores were associated with the absence, presence or severity of mental health problems as determined by clinical professionals. Such an evaluation would have been a valuable and relevant addition to the results presented here, which relied on self-reported outcome measures. Furthermore, since the MHQoL was only included in one of the ECOS survey waves, we were not able to perform a test-retest validation. This would have allowed further evaluation of the reliability (consistency in measurement) and responsiveness (ability to capture changes in mental health) of the MHQoL. In addition, in the absence of MHQoL tariffs, we based our analyses on the sum scores and VAS scores of the MHQoL, while basing them on utility scores would have been interesting and insightful as well.

Second, the data were collected using an online survey, which may have caused some selection bias in respondents and may be associated with lesser engagement with the survey by respondents. Moreover, and very important to stress, the data were collected during the COVID-19-crisis.

This situation, both the threat or experience of falling ill as well as the strong measures imposed in most countries to mitigate the outbreak, may have affected mental health of the population. The fact that the survey focused on COVID-related aspects (including risks, worries and government intervention) may have increased the awareness of negative consequences and, therefore, negative feelings. The QoL and wellbeing scores observed in our study were, on average, indeed lower than previously observed. The age-profile for mental health, with younger people scoring lower than expected and lower than older people in all domains except for physical health, may also be related to this context, if the mental health of young adults on average would be affected more due to COVID (measures) than that of the old (e.g. by affecting normal activities and social life more in the young). As a competing explanation, the online form of the survey may have caused a selection bias in the group of older people towards those with a higher level of mental health. While this may not have affected the results in relation to our research question on the validity of the MHQoL, it is important to emphasize this when using the here presented MHQoL scores as a reference in future studies.

As a third limitation, the MHQoL instrument was translated from Dutch to four other languages by native speakers with a background in health economics. A more extensive translation procedure, as was done for English and German, using forward-backward techniques, would have been superior.

6.4.2 Future perspective

Given the findings presented here, the MHQoL appears to be a valid and reliable measure of mental health-related QoL in the Western European context. More research, confirming these findings, and expanding the investigation to other aspects of feasibility, reliability and validity, is

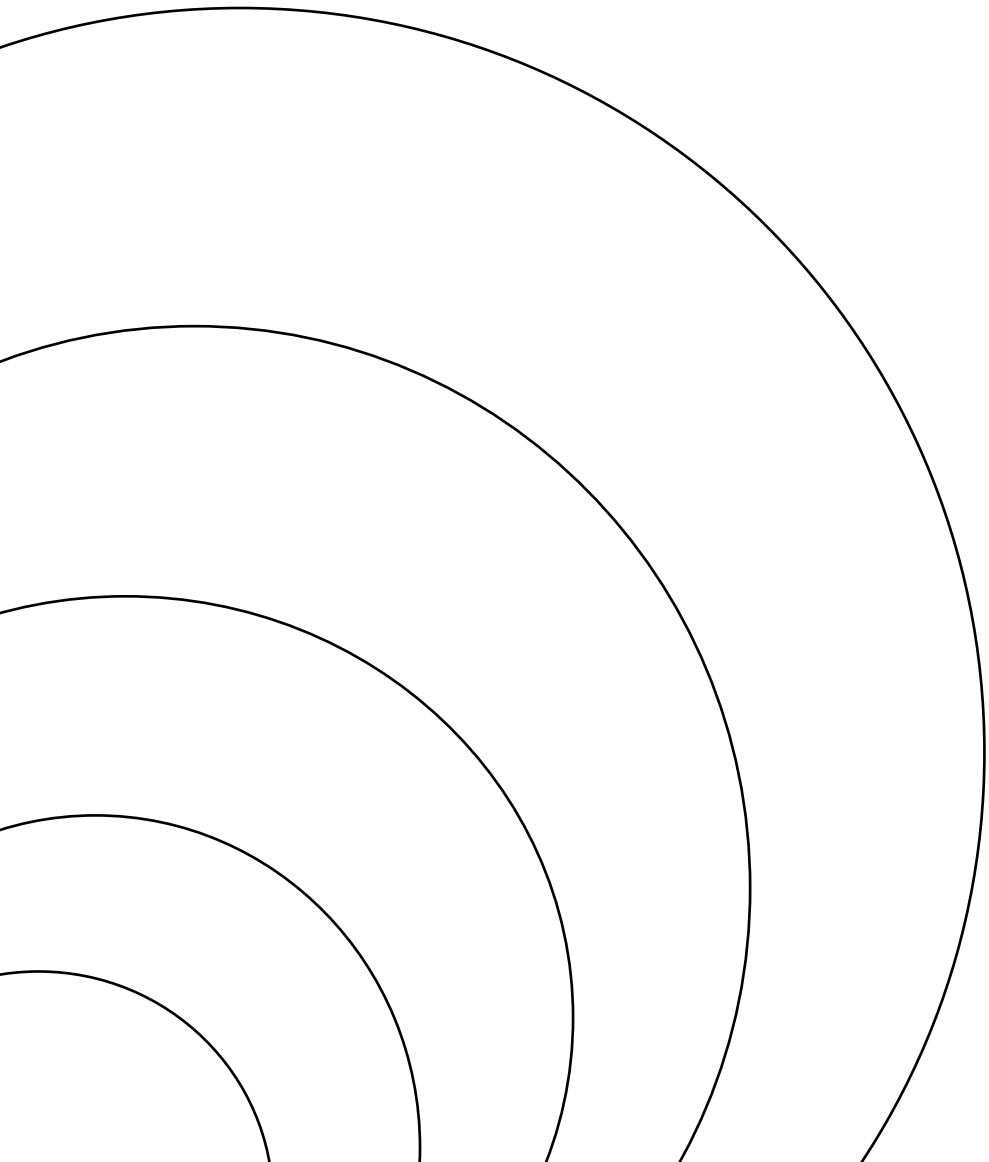
required in order to gain a fuller understanding of the psychometric properties of the MHQoL, also in different groups (including people with known mental health problems). If the current positive findings are confirmed, the MHQoL can be used in different settings to monitor or evaluate (changes in) mental health-related quality of life. This comprehensive measure may also be a valuable addition to disease-specific clinical outcomes in clinical decision making e.g. to inform patients on the expected impact of an intervention on their quality of life. It could also be applied in economic evaluations, informing decision-makers about the full costs and benefits of mental health interventions. It needs noting that in that context, using the MHQoL rather than generic HR-QoL measures, has the advantage of being more comprehensive in terms of relevant health domains covered but comes at the price of limiting comparability between outcomes. This may be less problematic for decisions within the mental health domain. Moreover, if adequate thresholds would be established for the MHQoL (like for generic HRQoL), decision-making would be more straightforward. However, given that in several Western European countries (with Germany as an exception) [181] the EQ-5D currently is prescribed to be used in economic evaluations (also within the mental health field), for these countries administering the MHQoL instrument *alongside* the EQ-5D is recommendable in the context of economic evaluations. This will also offer the opportunity to further investigate the relationship between the EQ-5D and MHQoL instruments.

The use of the MHQoL in economic evaluations would require having valid tariffs that transform a MHQoL state as described on the descriptive system into a utility score (typically anchored on the states 'dead' (0) and 'perfect health' (1). Such tariffs are expected to be developed soon (for the Netherlands).

6.5 Conclusion

Our results suggest the MHQoL is a psychometrically sound measure of mental health-related quality of life in the Western European context. While more research remains necessary, this makes the MHQoL instrument interesting to be used in (economic) evaluations of mental health interventions, as it more comprehensively captures their benefits.

Chapter 7
General discussion



7.1 Introduction

This thesis aims to contribute to the optimization of the decision-making framework for making decisions on including healthcare interventions into the basic benefit package (BBP) based on explicitly stated criteria, as well as of the underlying HTA methodology. [29] To this end, two research areas were covered. The first concerned investigating the broadening of the *scope* of HTA. This scope could be broadened by adding criteria to the traditional set of assessment criteria (i.e., in the Netherlands: necessity, effectiveness, cost-effectiveness, and feasibility). While other potentially relevant additional criteria could have been explored, here we focused on "profitability to the manufacturer". This criterion was considered relevant to investigate given the societal as well as scientific debate regarding fair pricing of technologies, especially pharmaceuticals. The second research area concerned investigating broadening the *use* of HTA. This relates to the ambition of HTA organisations, including the Dutch National Health Care institute, to perform more systematic assessments of healthcare interventions other than pharmaceuticals. Only by also systematically assessing other types of healthcare, it can be ensured that care included in the BBP satisfies the relevant criteria, and therefore sufficiently contributes to the goals of the healthcare sector. However, broadening the scope and use of HTA both come with their own challenges.

In this final chapter we will first address the research questions stated in the general introduction (chapter 1) based on our findings as presented in chapters 2 to 6. Furthermore, we will reflect on the strengths and limitations of the research presented in this thesis and highlight implications of our findings for policy and research.

7.2 Main findings

Part 1 (chapters 2 and 3) focused on broadening the scope of HTA by examining the role of profitability as a potential additional assessment

criterion. **Chapter 2** investigated whether manufacturers' costs in relation to the price of a pharmaceutical, or "profitability to the manufacturer", is currently explicitly considered by HTA organisations. This was investigated by studying reimbursement reports of expensive pharmaceuticals from four jurisdictions, namely those from the National Health Care Institute (Zorginstituut Nederland or ZIN; the Netherlands), the National Institute for Health and Care Excellence (NICE; England), the Canadian Agency for Drugs and Technologies in Health (CADTH; Canada), as well as the Pharmaceutical Benefit Advisory Committee (PBAC) and the Medical Services Advisory Committee (MSAC) (both Australia). We found that profitability was considered explicitly only in a limited number of reports, perhaps suggesting that it may not play a large role in reimbursement decisions at present. One of the 87 investigated HTA-reports contained a cost-based justification of the demanded price, as provided by the manufacturer. Thirteen reports contained general considerations of the HTA organisation relating the proposed prices to manufacturers' costs, including invitations to manufacturers to justify high prices by demonstrating high costs. We concluded that despite the attention given to manufacturers' costs in relation to price within the scientific literature and in public debates, it does not seem to receive systematic, explicit attention in HTA-reports for expensive pharmaceuticals.

Chapter 3 explored whether and to what extent actual healthcare policy makers take information on profit margins into account in hypothetical reimbursement decisions when presented alongside common HTA information on pharmaceuticals. In a discrete choice experiment (DCE) it was found that if presented to Dutch policy makers (n=58) profit margins of pharmaceuticals are influential, with higher profit margins lowering the likelihood of reimbursement. The importance of profit margins in comparison to other included assessment criteria (namely: effectiveness,

disease severity/necessity, cost-effectiveness, and budget impact) was relatively low, but not negligible. Interestingly, 61% of respondents indicated that profit margins should play a role in reimbursement decisions, whereas 39% indicated it should not. These findings suggest that, if available to Dutch policy makers, profit margins of pharmaceutical products in general would influence reimbursement decisions but also that these policy makers are not aligned regarding the normative question of whether profitability should play a role.

Part 2 (chapter 4) explored the challenges HTA organisations face when broadening the use of HTA towards healthcare interventions other than outpatient pharmaceuticals. To do so, **chapter 4** described important challenges that ZIN, as an HTA organisation, will have to overcome when broadening the use of HTA and the decision-making process based upon it towards inpatient pharmaceuticals, medical devices, curative interventions, non-pharmaceutical curative mental healthcare, and non-curative care. After a description of the Dutch decision-making process regarding reimbursement within the BBP, five important characteristics of outpatient pharmaceuticals were highlighted which facilitate and stimulate the use of HTA. These characteristics are (i) a closed reimbursement system, (ii) the absence of alternative policy measures, (iii) the existence of marketing authorization, (iv) an identifiable and accountable counterparty, and (v) the product characteristics. We discussed the challenges and some solutions for the selected types of healthcare interventions in relation to these characteristics. Overall, it was clear that the investigated types of healthcare interventions differed in various ways and degrees from outpatient pharmaceuticals in terms of these characteristics, creating important challenges for all phases of the reimbursement decision-making process. These challenges include the need for identification of interventions for assessment, the unavailability of evidence, and specific methodological

challenges in performing HTA. Importantly, for some types of healthcare interventions more challenges were expected than for others. Solutions could be suggested for some of the identified challenges, ranging from horizon scanning, implementing a "lock" for medical devices, public funding of evidence generation, to alternative adaptive HTA processes [109]. In conclusion, broadening the systematic use of HTA requires creating a suitable regulatory and policy framework as well as developing specific methodologies to be able to perform HTA research in particular circumstances.

Part 3 (chapters 5 and 6) consisted of two case studies on methodological challenges related to broadening the use of HTA specifically in the fields of medical devices and mental health. For medical devices, three of these challenges can be distinguished: their outcomes may be more context-dependent, medical devices may be subject to incremental innovation, and learning effects may occur in the early phases of their use. **Chapter 5** therefore examined to which extend these specific characteristics of medical devices were accounted for in economic evaluations of Transcatheter Aortic Valve Implantation (TAVI). It was observed that these challenges and their influence on the reported results of the economic evaluations, were typically only partially discussed and rarely quantitatively dealt with in the reviewed economic evaluations of TAVI. Only two out of the 15 included studies used some type of analytical solution to deal with the identified challenges. The first study selected registries that allowed inclusion of data after an initial learning effect, hence avoiding data from situations in which proper training and experience was not yet realized. The second study provided additional country specific analyses to deal with (high-level) context dependency. No methodological solutions to cope with incremental innovation (e.g., specific sensitivity analyses) were found in the included articles. The limited attention given to these specific characteristics of medical devices within

the included economic evaluations seems inappropriate from a general HTA perspective since it may lead to misinformed decisions.

Broadening the use of HTA may also require measuring outcomes beyond or instead of the traditional outcome of health-related quality of life of the individual patient. As an example, mental healthcare interventions may be aimed at improving well-being, autonomy, social participation, and at reducing criminality or drug abuse[113]. Such outcomes may not be adequately captured in common generic health-related quality of life measures such as the EQ-5D and therefore may require outcome measures that do capture all relevant outcomes and are tailored to the aims of the intervention under evaluation. Such measures may not be readily available for all different healthcare domains. [116] **Chapter 6** in that context explored whether the recently developed Mental Health Quality of Life (MHQoL) instrument, which aims to be adequate in capturing outcomes of mental health interventions, can be considered psychometrically sound. The MHQoL is able to describe (and later value) current mental health-related QoL of respondents.[165] The instrument invites respondents to describe their current mental health state using seven dimensions and four response levels per dimension. Preference-based tariffs should become available soon to allow converting the mental health descriptions into utility scores anchored on 0 (dead) and 1 (full mental health-related QoL). As part of this thesis a psychometric evaluation of the MHQoL was performed using existing international, cross-sectional data from 7,155 respondents from seven European countries (Denmark, France, Germany, Italy, The Netherlands, Portugal and the United Kingdom).[167] Reliability and construct validity were examined and the results suggested that the MHQoL is a psychometrically sound measure of mental health-related quality of life in people with mental health problems in the Western European context. This contributed to expanding the set of outcome measures which may

support the broadening of the use of HTA also to mental health interventions in a valid way.

7.3 Answers to the main research questions

Based on the findings of this thesis, we provide answers to the three main research questions stated in **chapter 1** and relate these answers to our aim of optimizing the current decision-making framework and underlying HTA methodology. The first question was whether broadening the scope of HTA with "profitability" as an additional assessment criterion would influence decisions regarding the inclusion of interventions in the BBP. We conclude that this influence is expected for three reasons. First, part of the reviewed reimbursement reports contained considerations which related prices to manufacturers' costs, including invitations to justify high prices by demonstrating high costs. Thus, HTA organisations to some degree already appeared to consider, be interested in, or open to information on profitability. Second, a majority of surveyed Dutch policy makers stated that profit margins *should* have a role in reimbursement decisions. Third, it was shown that higher profit margins indeed lowered the likelihood of reimbursement in hypothetical decisions. Thus (part of the) individual policy makers did include this information in their decisions, when the information was available to them. This conclusion, that "profitability" is expected to have an influence on reimbursement decisions, raises new and important questions about the desirability and feasibility of adding this criterion to the scope of HTA. These issues need to be considered carefully before it can be decided whether adding this criterion will improve the current decision-making framework.

The second research question was "which are challenges when broadening the use of HTA towards non-pharmaceutical interventions". Several challenges were mapped out in chapter 4. These challenges differ between

types of healthcare interventions and relate to all phases of the reimbursement decision-making process. They concern both policy related challenges, including the need for identification of subjects for assessment and the unavailability of evidence, and intervention type dependent methodological challenges. These challenges also highlight the need to improve the current decision-making framework and its underlying HTA methodology if the aim of HTA organisations and policy makers indeed is to use HTA more broadly and systematically in the future. However, the multitude and complexity of the identified challenges, and their variation between healthcare domains and interventions, also highlights the need for prioritization in moving from the current practice towards a broad use of HTA across all healthcare domains.

Our third main research question was whether specific methodological challenges can be resolved when broadening the use of HTA towards non-pharmaceutical interventions. We conclude that this may be the case for certain challenges, but more difficult for others. Indeed, we found that three important often mentioned methodological challenges of medical devices largely remained unaddressed (let alone resolved) in almost all reviewed economic evaluations of TAVI. Also, current international methodological guidance on the assessment of medical devices and national HTA guidelines provide little clarity on how to deal with these challenges - although they do acknowledge their relevance. More research is needed to further expand and refine HTA methodology to deal with such challenges. The psychometric evaluation of the MHQoL showed that a new instrument, potentially useful to measure outcomes in the mental health domain, has promising psychometric characteristics. This is an example of how existing HTA methodology may be improved and complemented in order the allow HTA in other domains. This may contribute to the optimization of the decision-making framework and underlying HTA methodology that is used for

decisions on including healthcare interventions into the BBP by allowing it to be used more broadly. Both examples also highlight the clear relationship between HTA methodology and the decision-making and policy process. If certain methodological aspects are not adequately dealt with, this puts more pressure on decision makers as the uncertainty around specific estimates will increase and their relevance may diminish. Moreover, solving certain methodological issues (e.g. being able to use the MHQoL in the context of mental health interventions to capture the relevant outcomes of such interventions) may also create new policy questions. For instance, which thresholds would be appropriate to use for gains in MHQoL. This highlights the need for policy and methodological development to ideally go hand in hand.

7.4 Strengths and limitations

The results reported in this thesis should be considered in the light of some noteworthy strengths and limitations. A first strength is that by especially focusing on the Dutch context, we hope that our findings are directly relevant for ZIN. Despite this focus, the findings presented in this thesis probably will be useful in other jurisdictions as well, as many healthcare systems are dealing with similar challenges. The relevance of our findings for other jurisdiction does require further investigation. A second strength is that the research presented in this thesis used and combined different, complementary research methods. First, a review of policy documents was combined with a discrete choice experiment among policy makers which allowed to contrast published policy considerations with preferences of policy makers. Second, an analytical review providing an overview of challenges related to broadening the use of HTA to other interventions than pharmaceuticals was supplemented with a systematic review and a psychometric evaluation to enrich the overview with two illustrative case studies.

Another strength is the close connection of the research team to the ZIN organisation, also by being part of the Research Network HTA NL (in Dutch: 'Academische Werkplaats Verzekerde Zorg'). This not only provided us with access to policy makers as respondents and quick access to relevant policy documents, but also meant that the research was closely connected to the development of ideas within ZIN. This alignment with the context of and development within ZIN may subsequently be beneficial in translating the research findings into further developments within ZIN. An obvious potential disadvantage of the connection to ZIN could be that the openness to unconventional ideas or critique on the status-quo would be more limited. This risk was mitigated by discussing these issues within the research team and challenging ideas both within the team as through presentations to broader audiences (both scientific and policy-oriented).

As an important limitation, it should be acknowledged that this thesis is based on selected case studies within a very large and diverse field. To investigate broadening the scope of HTA, we selected profitability as subject of research. However, profitability does not represent all options to broaden the scope of HTA with new criteria. For instance, elements like labor reducing technology or environmental impact of interventions could also have been investigated and are highly relevant. Also, the case studies used to investigate broadening the use of HTA are illustrative and concern relevant types of healthcare, but clearly do not represent the whole range of research areas and methodological challenges. If we would have chosen other case studies, different results would have emerged. We could for example have included paramedical care, palliative care or dental care as cases.

The choice of TAVI when looking at methodological challenges in economic evaluations of medical devices will have had a clear effect on our findings.

Results may have been different when, for example, glucose monitoring systems, implantable cardioverter-defibrillators or robotic-assisted surgical systems would have been the subject of the review. Nonetheless, while we in a direct sense only can conclude that there are still unresolved methodological challenges in the field of economic evaluations of TAVI, it seems unlikely that the situation would be completely different for other medical devices – although it would be interesting to see this confirmed in future studies. Moreover, validation of another outcome measure than the MHQoL (e.g. for palliative care or wellbeing of the elderly [182]), may also have led to different results. While, based on the research presented in this thesis, we again can only directly conclude that the MHQoL appears a promising new outcome measure in the context of HTA of mental health interventions, more outcome measures are being developed and validated that may help to broaden the use of HTA in other areas as well. Moreover, while we believe to have reported on important and relevant challenges by the case studies of TAVI and the MHQoL, we cannot and do not claim these were the most relevant or important ones. Finally, it should be acknowledged that our perspective, the samples used, the methods used, and our research scope leaves many research questions unaddressed which are worthy of future research.

7.5 Implications for (future) HTA research

Our findings regarding broadening the scope of HTA highlighted the relevance of additional investigation of the desirability and feasibility of adding profitability as an assessment criterion to the decision-making framework. In that context, it should be noted that, as of 2022, ZIN invites manufacturers to provide additional information, in addition to the information in the normal submission, that may be relevant in evaluating the reasonableness of demanded price during the appraisal phase. Research is needed to further determine whether adding profitability as an

assessment criterion is desirable and feasible. Such research could include investigating the normative views of (Dutch) politicians and their electorate on this topic. A citizen forum may be one way of investigating the views within the Dutch general public. It should be acknowledged in such research that there are alternative ways to address profitability, for example during price negotiations, as part of price regulations, or when possible through price competition.[25] The complex relationships between reimbursement decisions, price negotiations, and market context, as well as the link between profitability and incentives for innovation also require more attention. Additionally, research is warranted into the practical feasibility of assessing "profitability to the manufacturer" as part of the decision-making process. Systematically obtaining information on or estimating profit margins of individual pharmaceuticals is complicated and would require overcoming many hurdles. Developing a definition of what a profit margin exactly is, and developing a process for measuring or estimating profit margins for use in reimbursement decision-making could be a start. Subsequently, policy makers will need a normative framework to appraise the obtained information or estimated profit margins. The development of such a normative framework can build on previous work by Uyl et al. [27] and may incorporate available fair pricing models [10, 52].

Our findings regarding broadening the use of HTA imply the need for assessments of medical devices to more systematically reflect on and deal with the specific methodological challenges related to such devices (i.e., learning effects in their use, incremental innovation, context dependent outcomes). The reviewed economic evaluations of TAVI emphasized this is not yet the case. For instance, specific sensitivity analyses may inform policy makers better on the uncertainty around the assessment. Developing standardized methods that may help in this context needs to be high on the research agenda. These solutions may related to both analytical aspects as

well as appropriate data collection. [128] Developing and sharing best practices may also further this field.

Additionally, to allow policy makers to compare different interventions and evaluations, HTA researchers ideally should use generic outcome measures. However, when these are not sufficiently sensitive or broad to capture relevant outcomes, the use of other outcomes (which may be broader than HRQoL – e.g. measuring wellbeing – or narrower – e.g. domain specific outcome measures), is inevitable. Supplementing conventional generic HRQoL measures with other measures then may be a first step in better informing policy makers. Still, using more than one outcome measure may complicate the decision-making process. The solution to just use one outcome measure, if this is known to be inadequate in certain circumstances, may make decision making easier but also very well lead to wrong decisions. The trade-off in this context is often framed between comparability between interventions and validity of outcomes for specific types of healthcare. Both are required for optimal allocation decisions. It needs noting however that comparability may also be achieved by having sound threshold values connected to different outcome measures. This allows policy makers to appraise presented results in a similar fashion, even if different outcome measures are used. Establishing adequate thresholds for different relevant outcome measures requires time and resources. However, being able to better compare different types of interventions will support optimal allocation of resources across the whole BBP.

This line of research and decision making would benefit from having a core set of generic outcome measures that validly cover most types of interventions. It is expected that this smallest possible set of generic outcome measures will at least include a general outcome measure for health related QoL (probably the EQ-5D-5L), and a general outcome

measure for wellbeing, but it might also include instruments like the MHQoL or instruments measuring outcomes in the context of palliative care. A joint effort from the HTA community to come to such a set (and perform research to validate it in different contexts) is required for this.

This thesis did not deal with a number of relevant aspects related to broadening HTA that also deserve more attention in future research. First, research using a more process oriented perspective on HTA, for example by using the evidence-informed deliberative processes (EDPs) framework [183], can be complementary to our research which followed a relatively technical perspective. The use of the EDPs framework could involve more explicit attention to stakeholder deliberation and other aspects of a legitimate reimbursement decision-making process. The legitimacy of processes is related to the acceptance of reimbursement decisions. This acceptance may be of special concern for non-pharmaceutical interventions which more often than outpatient pharmaceuticals may already be provided to patients, and obtaining public support for disinvestment is challenging. [12, 13] Second, our focus has been on HTA organisations such as ZIN as the users of HTA. Other actors such as health insurance companies and clinical guideline developers were not considered in this thesis as stakeholders who might wish to use HTA – highlighting another type of broadening the use (and perhaps scope) of HTA. Further research could investigate how these actors might (wish to) use HTA and in which form and context. Third, we focused on profitability as potential new assessment criterion. Other highly relevant candidate criteria, as environmental sustainability aspects [184] [185], impact on labor use and impact on socio-economic inequalities, should also be examined as potentially broadening the scope of HTA.

7.6 Implications for policy makers

It is important for policy makers to develop a position on the role of

profitability as a criterion during the process of setting boundaries to the BBP. For example, they could state whether using price negotiations as a policy instrument could be based on an expected high profitability to the manufacturer, even when the cost-effectiveness is acceptable. The importance of an explicit position on the role of profitability followed from our findings that information on profitability in general is absent in both reviewed reimbursement reports as within HTA frameworks. However, at the same time it is expected to influence reimbursement decisions when presented and considered relevant by a majority of policy makers. Currently, information on profitability may be used implicitly in the appraisal of (some) healthcare interventions, which would negatively affect the 'accountability' [186] of the decision-making process by affecting its uniformity and transparency.

Furthermore, policy makers should plan the way forward towards a more systematic consideration of healthcare interventions other than outpatient pharmaceuticals in terms of their eligibility of being funded in the BBP. This way forward should include the optimization of the decision-making framework and underlying HTA methodology. Making such a plan is important since broadening the use of HTA is challenging and even more so in some areas than in others. The varying deviations from the characteristics of outpatient pharmaceuticals of other interventions determine which challenges are likely to be faced and their degree of severity. Policy makers should be aware of those various challenges and could focus on gradually expanding the use of HTA, starting in those areas where challenges are most manageable. Interventions relatively similar to outpatient pharmaceuticals, like inpatient pharmaceuticals and medical devices, would be logical first steps in coming to a broader use of HTA. Following this path of 'least expected challenges' has the advantage of being close to the methods and processes known from the outpatient

pharmaceuticals context and may result in a relatively controlled way forward. However, such an approach also comes with downsides as the area's most like outpatient pharmaceuticals may be least challenging but perhaps at the same time not the most important. As a consequence, extensive opportunities for improving the content of the BBP may remain unaddressed. This also entails the risk that at some point other policy makers will impose cost containment measures that do not balance affordability with quality and accessibility. Therefore, planning the route forward for the use of HTA should also involve the expected benefit of (more extensive) use of HTA when prioritizing areas. For example, within long-term care, representing a large part of Dutch healthcare expenditures and being relatively underrepresented in HTA based decisions, a more systematic consideration of healthcare interventions may potentially lead to more gains in terms of quality, accessibility and affordability of care. Relatively large investments in the development of methodologies and in the modification of the policy context may be required in order to facilitate the aim of a broad use of HTA. This includes establishing a valid general outcome measure for wellbeing, financing HTA research, and developing an adapted appraisal framework. Although these investments may be extensive, they may be largely compensated by improved quality, accessibility and affordability of healthcare, including long-term care. However, whether the route forward mainly follows the line of least expected challenges, the line of highest expected (net) benefits, or a combination of both approaches, is a decision for policy makers to make. For Dutch policy makers making these decisions, this may involve reflecting on the current attention to "Appropriate care" (Passende zorg). Several national stakeholders, including ZIN, are currently aiming to promote the sustainability of the Dutch healthcare system. The four main principles they follow include the requirement that healthcare interventions need to be effective at a reasonable cost. The other three principles are shared decision

making, a positive approach to health, and care provided close to the patient. Use of HTA will promote ”effectiveness at a reasonable cost” by setting boundaries to the content of the BBP according to this principle. Broadening the use of HTA, particularly to areas where this can have a significant impact, will contribute to achieving the goals underlying the concept of “Appropriate care”.

7.7 Conclusion

There are many opportunities to improve the decision-making framework and underlying HTA methodology that are used to set accepted boundaries to the BBP. “Profitability to the manufacturer” can be considered by policy makers as an additional assessment criterion. Moreover, based on the experience with the assessment of outpatient pharmaceuticals, the framework and underlying HTA methodology can be optimized for a more systematic consideration of other healthcare interventions. The latter will require planning and choices, including making a prioritization of which areas to address first which can be based on expected challenges and expected benefits. The resulting optimized framework and methodology will then allow the explicit balancing of affordability with quality and accessibility in parts of the BBP that may be much larger than that of outpatient pharmaceuticals. Doing so will help to develop and steer the healthcare system in line with public values and will contribute to its sustainability. While unknown roads lay ahead in achieving this goal, the destination will be well worth the journey.

References

1. RIVM. Volksgezondheid Toekomst Verkenning 2018 Een gezond vooruitzicht Synthese. 2018. Available from: https://www.vtv2018.nl/.
2. CBS. Zorguitgaven; kerncijfers. Statline [Internet]. 2022. 28/12/2022. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84047NED/barv?ts=1672224733132.
3. van Ewijk CvdH, A.; Besseling, P. Toekomst voor de zorg. Den Haag: Centraal Planbureau (CPB); 2013.
4. WRR. Kiezen voor houdbare zorg. Mensen, middelen en maatschappelijk draagvlak. wrr-Rapport 104, Den Haag: WRR; 2021.
5. Stadhouders N, Kruse F, Tanke M, Koolman X, Jeurissen P. Effective healthcare cost-containment policies: A systematic review. Health Policy. 2019;123(1):71-9.
6. O'Rourke B, Oortwijn W, Schuller T. Announcing the New Definition of Health Technology Assessment. Value Health. 2020;23(6):824-5.
7. Reckers-Droog V, van Exel J, Brouwer W. Willingness to Pay for Health-Related Quality of Life Gains in Relation to Disease Severity and the Age of Patients. Value Health. 2021;24(8):1182-92.
8. Al-Janabi H, Wittenberg E, Donaldson C, Brouwer W. The relative value of carer and patient quality of life: A person trade-off (PTO) study. Soc Sci Med. 2022;292:114556.
9. Marsh K, Ganz ML, Hsu J, Strandberg-Larsen M, Gonzalez RP, Lund N. Expanding Health Technology Assessments to Include Effects on the Environment. Value Health. 2016;19(2):249-54.
10. Moon S, Mariat S, Kamae I, Pedersen HB. Defining the concept of fair pricing for medicines. BMJ. 2020;368:l4726.
11. ZIN. Cost effectiveness in practice. Diemen: Zorginstituut Nederland; 2015.
12. Rotteveel AH, Reckers-Droog VT, Lambooij MS, de Wit GA, van Exel NJA. Societal views in the Netherlands on active disinvestment of publicly funded healthcare interventions. Soc Sci Med. 2021;272:113708.
13. van de Wetering EJ, van Exel J, Brouwer WB. The Challenge of Conditional Reimbursement: Stopping Reimbursement Can Be More Difficult Than Not Starting in the First Place! Value Health. 2017;20(1):118-25.
14. Wang Y, Qiu T, Zhou J, Francois C, Toumi M. Which Criteria are Considered and How are They Evaluated in Health Technology Assessments? A Review of Methodological Guidelines Used in Western and Asian Countries. Appl Health Econ Health Policy. 2021;19(3):281-304.
15. ZIN. Package advice in practice Deliberations for arriving at a fair package. 2017:[26 p.]. Available from: www.zorginstituutnederland.nl.
16. Dunning AJ. Kiezen en delen Advies in Hoofdzaken Van De Commissie Keuzen in De Zorg. Den Haag: Albani; 1991.
17. Skedgel C, Henderson N, Towse A, Mott D, Green C. Considering Severity in Health Technology Assessment: Can We Do Better? Value Health. 2022;25(8):1399-403.
18. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands. Health Policy. 2018;122(6):621-9.
19. ZIN. Ziektelast in de praktijk. Diemen: Zorginstituut Nederland; 2018.

20. ZIN. Assessment of 'established medical science and medical practice' (update 2015). Diemen: Zorginstituut Nederland; 2017.
21. Brouwer W, van Baal P, van Exel J, Versteegh M. When is it too expensive? Cost-effectiveness thresholds and health care decision-making. The European Journal of Health Economics. 2019;20(2):175-80.
22. van Houdenhoven M. Extreem dure medicijnen door verziekt patentsysteem. Gemeenschapsgeld verdient een betere bestemming. https://www.medischcontact.nl: KNMG; 2021.Website: https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/extreem-dure-medicijnen-door-verziekt-patentsysteem.htm.
23. Balderrama F, Schwartz LJ, Longo CJ. When are Pharmaceuticals Priced Fairly? An Alternative Risk-Sharing Model for Pharmaceutical Pricing. Health Care Anal. 2020;28(2):121-36.
24. Berdud M, Drummond M, Towse A. Establishing a reasonable price for an orphan drug. Cost Eff Resour Alloc. 2020;18:31.
25. Gronde TV, Uyl-de Groot CA, Pieters T. Addressing the challenge of high-priced prescription drugs in the era of precision medicine: A systematic review of drug life cycles, therapeutic drug markets and regulatory frameworks. PLoS One. 2017;12(8):e0182613.
26. Trujillo AJ, Karmarkar T, Alexander C, Padula W, Greene J, Anderson G. Fairness in drug prices: do economists think differently from the public? Health Econ Policy Law. 2020;15(1):18-29.
27. Uyl-de Groot CA, Lowenberg B. Sustainability and affordability of cancer drugs: a novel pricing model. Nat Rev Clin Oncol. 2018;15(7):405-6.
28. ZIN. ZorgCijfers Monitor Zorgverzekeringswet en Wet langdurige zorg 4e kwartaal 2021. Diemen: Zorginstituut Nederland; 2022.
29. Drummond MF, Schwartz JS, Jonsson B, Luce BR, Neumann PJ, Siebert U, et al. Key principles for the improved conduct of health technology assessments for resource allocation decisions. Int J Technol Assess Health Care. 2008;24(3):244-58; discussion 362-8.
30. Chamova J. Mapping of HTA national organisations, programmes and processes in EU and Norway2017.
31. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs--same or different? Value Health. 2009;12(4):402-4.
32. Tarricone R, Torbica A, Drummond M. Challenges in the Assessment of Medical Devices: The MedtecHTA Project. Health Econ. 2017;26 Suppl 1:5-12.
33. van Krugten FCW, Feskens K, Busschbach JJV, Hakkaart-van Roijen L, Brouwer WBF. Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. Health Qual Life Outcomes. 2021;19(1):249.
34. Xu RH, Keetharuth AD, Wang LL, Cheung AW, Wong EL. Measuring health-related quality of life and well-being: a head-to-head psychometric comparison of the EQ-5D-5L, ReQoL-UI and ICECAP-A. Eur J Health Econ. 2022;23(2):165-76.
35. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res. 2012;21(1):167-76.

36. Brazier J. Is the EQ-5D fit for purpose in mental health? Br J Psychiatry. 2010;197(5):348-9.
37. House-of-Commons. Health and Social Care Committee Oral evidence: Availability of Orkambi on the NHS, HC 18082019. Available from: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/health-and-social-care-committee/availability-of-orkambi-on-the-nhs/oral/97783.html
38. WHO. WHO guideline on country pharmaceutical pricing policies, second edition2020. Available from: https://apps.who.int/iris/rest/bitstreams/1309649/retrieve.
39. EUnetHTA. HTA Core Model® Version 3.02016 2020 october 5. Available from: http://www.corehta.info/model/HTACoreModel3.0.pdf.
40. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on DEcisionMaking--the EVIDEM framework and potential applications. BMC Health Serv Res. 2008;8:270.
41. Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, et al. From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decisionmaking. Cost Eff Resour Alloc. 2012;10(1):9.
42. Stratil JM, Baltussen R, Scheel I, Nacken A, Rehfuess EA. Development of the WHO-INTEGRATE evidence-to-decision framework: an overview of systematic reviews of decision criteria for health decision-making. Cost Eff Resour Alloc. 2020;18:8.
43. Husereau D, Dempster W, Blanchard A, Chambers J. Evolution of drug reimbursement in Canada: the Pan-Canadian Pharmaceutical Alliance for new drugs. Value Health. 2014;17(8):888-94.
44. Enzing JJ, Himmler S, Knies S, Brouwer WBF. Do Profit Margins of Pharmaceuticals Influence Reimbursement Decisions? A Discrete Choice Experiment Among Dutch Healthcare Decision Makers. Value in Health. 2021;doi: 10.1016/j.jval.2021.08.007.
45. Vogler S, Paris V, Ferrario A, Wirtz VJ, de Joncheere K, Schneider P, et al. How Can Pricing and Reimbursement Policies Improve Affordable Access to Medicines? Lessons Learned from European Countries. Appl Health Econ Health Policy. 2017;15(3):307-21.
46. EXPH. Defining value in "value-based healthcare" Report of the Expert Panel on effective ways of investing in Health. Expert Panel on Effective Ways of Investing in Health (EXPH) [Internet]. 2019. Available from: https://ec.europa.eu/health/sites/default/files/expert_panel/docs/024_defining-value-vbhc_en.pdf.
47. de Lima Lopes G, Nahas GR. Chimeric antigen receptor T cells, a savior with a high price. Chin Clin Oncol. 2018;7(2):21.
48. Prasad V. Immunotherapy: Tisagenlecleucel - the first approved CAR-T-cell therapy: implications for payers and policy makers. Nat Rev Clin Oncol. 2018;15(1):11-2.
49. Brigand T, Prokupkova A, Muller G. CAR-T Cell Therapies: How much for survival?2018 30-4-2019. Available from: https://www.europeancancerleagues.org/wp-content/uploads/2018/06/CAR-T-ECL-Article_Final_20062018.pdf.
50. Paulden M. Recent amendments to NICE's value-based assessment of health technologies: implicitly inequitable? Expert Rev Pharmacoecon Outcomes Res. 2017;17(3):239-42.

51. AR. Paardenmiddel of noodverband? Resultaten prijsonderhandelingen geneesmiddelen. 2020. Available from: https://www.rekenkamer.nl/onderwerpen/gezondheidszorg/documenten/rapporten/2020/04/23/paardenmiddel-of-noodverband-resultaten-prijsonderhandelingen-geneesmiddelen.
52. van den Berg S, van der Wel V, de Visser SJ, Stunnenberg BC, Timmers L, van der Ree MH, et al. Cost-Based Price Calculation of Mexiletine for Nondystrophic Myotonia. Value Health. 2021;24(7):925-9.
53. Sarpatwari A, Avorn J, Kesselheim AS. Accounting for US public funding in drug development: how can we better balance access, affordability, and innovation? BMJ. 2020;371:m3841.
54. Barenie RE, Avorn J, Tessema FA, Kesselheim AS. Public funding for transformative drugs: the case of sofosbuvir. Drug Discov Today. 2021;26(1):273-81.
55. Annett S. Pharmaceutical drug development: high drug prices and the hidden role of public funding. Biol Futur. 2021;72(2):129-38.
56. Loi de financement de la sécurité sociale 2020 Amendement n°4742019. Available from: http://www.assemblee-nationale.fr/dyn/15/amendements/2416/AN/474
57. Neumann PJ, Ganiats TG, Russell LB, Sanders GD, Siegel JE. Cost-effectiveness in health and medicine. New York, NY: Oxford University Press; 2016. Available from: http://dx.doi.org/10.1093/acprof:oso/9780190492939.001.0001.
58. Enzing JJ, Knies S, Boer B, Brouwer WBF. Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? Health Econ Policy Law. 2021;16(4):440-56.
59. EXPH. INNOVATIVE PAYMENT MODELS FOR HIGH-COST INNOVATIVE MEDICINES. Expert Panel on Effective Ways of Investing in Health (EXPH) [Internet]. 2018.
60. Ledley FD, McCoy SS, Vaughan G, Cleary EG. Profitability of Large Pharmaceutical Companies Compared With Other Large Public Companies. JAMA. 2020;323(9):834-43.
61. CBS. Bedrijfsresultaat hoogst in farmaceutische industrie2017 September 16, 2020. Available from: https://www.cbs.nl/nl-nl/nieuws/2017/15/bedrijfsresultaat-hoogst-in-farmaceutische-industrie.
62. NICE. Eculizumab for Treating Atypical Haemolytic Uraemic Syndrome Highly Specialised Technologies Guidance. 2015.
63. ZIN. GVS Rapport 16/08 Lumacaftor/Ivacaftor (Orkambi®). 2016. Available from: https://www.zorginstituutnederland.nl/publicaties/adviezen/2016/05/13/gvs-advies-lumacaftor-ivacaftor-orkambi-bij-cystische-fibrose-cf-bij-patienten-van-12-jaar-en-ouder-die-homozygoot-zijn-voor-de-f508del-mutatie-in-het-cftr-gen.
64. Trapero-Bertran M, Rodriguez-Martin B, Lopez-Bastida J. What attributes should be included in a discrete choice experiment related to health technologies? A systematic literature review. PLoS One. 2019;14(7):e0219905.
65. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: Past, Present and Future. Pharmacoeconomics. 2019;37(2):201-26.
66. Koopmanschap MA, Stolk EA, Koolman X. Dear policy maker: have you made up your mind? A

discrete choice experiment among policy makers and other health professionals. Int J Technol Assess Health Care. 2010;26(2):198-204.
67. Hauber AB, Gonzalez JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health. 2016;19(4):300-15.
68. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Muhlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16(1):3-13.
69. Janssen EM, Hauber AB, Bridges JFP. Conducting a Discrete-Choice Experiment Study Following Recommendations for Good Research Practices: An Application for Eliciting Patient Preferences for Diabetes Treatments. Value Health. 2018;21(1):59-68.
70. Ratcliffe J, Lancsar E, Walker R, Gu Y. Understanding what matters: An exploratory study to investigate the views of the general public for priority setting criteria in health care. Health policy. 2017;121(6):653-62.
71. van de Wetering EJ, van Exel NJA, Rose JM, Hoefman RJ, Brouwer WBF. Are some QALYs more equal than others? The European Journal of Health Economics. 2016;17(2):117-27.
72. Wranik WD, Jakubczyk M, Drachal K. Ranking the Criteria Used in the Appraisal of Drugs for Reimbursement: A Stated Preferences Elicitation With Health Technology Assessment Stakeholders Across Jurisdictional Contexts. Value Health. 2020;23(4):471-80.
73. Mirelman A, Mentzakis E, Kinter E, Paolucci F, Fordham R, Ozawa S, et al. Decision-making criteria among national policymakers in five countries: a discrete choice experiment eliciting relative preferences for equity and efficiency. Value Health. 2012;15(3):534-9.
74. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete-choice experiment with assessment of external validity. Pharmacoeconomics. 2013;31(4):345-55.
75. Wouterse B, ter-Rele H, van-Vuuren D. CPB Policy Brief Financiering van de zorg op de lange termijn. Den Haag: Centraal planbureau; 2016.
76. Rijksoverheid. Rijksbegroting 2018. Available from: http://www.rijksbegroting.nl/2018/voorbereiding/begroting,kst236981_37.html
77. Banta D, Jonsson E. History of HTA: Introduction. Int J Technol Assess Health Care. 2009;25 Suppl 1:1-6.
78. INAHTA. HTA definition.Website: http://www.inahta.org.
79. Luce BR, Drummond M, Jonsson B, Neumann PJ, Schwartz JS, Siebert U, et al. EBM, HTA, and CER: clearing the confusion. Milbank Q. 2010;88(2):256-76.
80. ZIN. Procedure beoordeling extramurale geneesmiddelen. Diemen: Ministerie van Volksgezondheid, Welzijn en Sport, Zorginstituut Nederland; 2016.
81. ZIN. Publicaties 2020. Available from: https://www.zorginstituutnederland.nl/publicaties.
82. Chapman AM, Taylor CA, Girling AJ. Are the UK Systems of Innovation and Evaluation of Medical Devices Compatible? The Role of NICE's Medical Technologies Evaluation Programme (MTEP). Applied Health Economics and Health Policy. 2014;12(4):347-57.

83. van de Ven WP, Schut FT. Universal mandatory health insurance in the Netherlands: a model for the United States? Health Aff (Millwood). 2008;27(3):771-81.
84. Zorgverzekeringswet (Zvw). Dutch Government. Website: https://wetten.overheid.nl/BWBR0018450/2020-01-01.
85. CVZ. Pakketbeheer in de praktijk deel 3. Diemen: College voor Zorgverzekeringen; 2003.
86. Regeling zorgverzekering (Rzv). Dutch Government.Website: http://wetten.overheid.nl/BWBR0018715/2018-07-06.
87. Besluit zorgverzekeringen (Bzv). Dutch Government.Website: http://wetten.overheid.nl/BWBR0018492/2018-07-01.
88. Titlow K, Randel L, Clancy CM, Emanuel EJ. Drug coverage decisions: the role of dollars and values. Health Aff (Millwood). 2000;19(2):240-7.
89. Al-Badriyeh D, Alameri M, Al-Okka R. Cost-effectiveness research in cancer therapy: a systematic review of literature trends, methods and the influence of funding. BMJ Open. 2017;7(1):e012648.
90. Stolk EA, van Donselaar G, Brouwer WB, Busschbach JJ. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. Pharmacoeconomics. 2004;22(17):1097-107.
91. CVZ. Beoordeling stand van de wetenschap en praktijk. Diemen: College voor Zorgverzekeringen; 2015.
92. ZIN. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Diemen: Zorginstituut Nederland; 2015.
93. Paul JE, Trueman P. 'Fourth hurdle reviews', NICE, and database applications. Pharmacoepidemiol Drug Saf. 2001;10(5):429-38.
94. Oortwijn W, Sampietro-Colom L, Habens F, Trowman R. How Can Health Systems Prepare for New and Emerging Health Technologies? The Role of Horizon Scanning Revisited. Int J Technol Assess Health Care. 2018;34(3):254-9.
95. ZIN. Horizonscan Geneesmiddelen. Zorginstituut Nederland; 2020. Website: https://www.horizonscangeneesmiddelen.nl.
96. CADTH. Health Technology Management Program Horizon Scanning Products and Services Processes. Canadian Agency for Drugs and Technologies in Health (CADTH); 2017.Website: https://www.cadth.ca/sites/default/files/pdf/HS_Consolidated_Process_Final.pdf
97. NICE. Healthtech Connect. National Institute for Health and Care Excellence (NICE); 2020.Website: https://www.healthtechconnect.org.uk.
98. Moes F, Houwaart E, Delnoij D, Horstman K. Collective constructions of 'waste': epistemic practices for disinvestment in the context of Dutch social health insurance. BMC Health Services Research. 2019;19(1):633.
99. ZIN. Room for improvement analysis: Chest pain (Suspected) Stable Angina Pectoris 2017.
100. NZa. Monitor inhoud contracten medisch-specialistische zorg 2018 Utrecht: Nederlandse Zorgauthoriteit (NZa); 2019.Website: http://puc.overheid.nl/doc/PUC_297585_22.
101. French-Mowat E, Burnett J. How are medical devices regulated in the European Union? J R Soc

Med. 2012;105 Suppl 1:S22-8.
102. EP. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union. 2017;117(5.5.2017):1-175.
103. Van Norman GA. Drugs and Devices: Comparison of European and U.S. Approval Processes. JACC Basic Transl Sci. 2016;1(5):399-412.
104. Kirisits A, Redekop WK. The economic evaluation of medical devices: challenges ahead. Appl Health Econ Health Policy. 2013;11(1):15-26.
105. van der Sande RL, S.W.J.; Rooijmans, H.G.M. Kennis op de plank? Het nuttig effect van onderzoeken uit het programma van het fonds Ontwikkelingsgeneeskunde. Ned Tijdschr Geneeskd. 2003;147:2390-3.
106. ZIN. Program 'Potentially Promising Care'2020. Available from: https://english.zorginstituutnederland.nl/about-us/tasks-of-the-national-health-care-institute/program-potentially-promising-care
107. Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M. Characterising Uncertainty in the Assessment of Medical Devices and Determining Future Research Needs. Health Econ. 2017;26 Suppl 1:109-23.
108. Tarricone R, Torbica A, Drummond M, Medtec HTAPG. Key Recommendations from the MedtecHTA Project. Health Econ. 2017;26 Suppl 1:145-52.
109. Husereau D, Henshall C, Jivraj J. Adaptive approaches to licensing, health technology assessment, and introduction of drugs and devices. Int J Technol Assess Health Care. 2014;30(3):241-9.
110. Blencowe NS, Boddy AP, Harris A, Hanna T, Whiting P, Cook JA, et al. Systematic review of intervention design and delivery in pragmatic and explanatory surgical randomized clinical trials. Br J Surg. 2015;102(9):1037-47.
111. Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, et al. Challenges in evaluating surgical innovation. Lancet. 2009;374(9695):1097-104.
112. Knapp M, Wong G. Economics and mental health: the current scenario. World Psychiatry. 2020;19(1):3-14.
113. Schawo S, Bouwmans C, van der Schee E, Hendriks V, Brouwer W, Hakkaart L. The search for relevant outcome measures for cost-utility analysis of systemic family interventions in adolescents with substance use disorder and delinquent behavior: a systematic literature review. Health Qual Life Outcomes. 2017;15(1):179.
114. Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, et al. Outcomes of social care for adults: developing a preference-weighted measure. Health Technol Assess. 2012;16(16):1-166.
115. Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Soc Sci Med. 2008;67(5):874-82.
116. Hackert MQN, Exel JV, Brouwer WBF. Valid Outcome Measures in Care for Older People: Comparing the ASCOT and the ICECAP-O. Value Health. 2017;20(7):936-44.

117. Weatherly HLA, Neves De Faria RI, van den Berg B, Sculpher MJ, O'Neill P, Nolan K, et al. Scoping review on social care economic evaluation methods. York UK: Centre for Health Economics; 2017. p. 1-53.
118. Jönsson BO, W; Rutten, F; Wailoo, A. Evaluation Health Technology Assessment Methodology Programme. The Hague: ZonMW; 2015.
119. Fuchs S, Olberg B, Panteli D, Perleth M, Busse R. HTA of medical devices: Challenges and ideas for the future from a European perspective. Health Policy. 2017;121(3):215-29.
120. EUnetHTA. Guideline "Therapeutic medical devices"2015 3th December 2019. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/Therapeutic-medical-devices_Guideline_Final-Nov-2015.pdf
121. Fuchs S, Olberg B, Panteli D, Busse R. Health Technology Assessment of Medical Devices in Europe: Processes, Practices, and Methods. Int J Technol Assess Health Care. 2016;32(4):246-55.
122. Bluher M, Saunders SJ, Mittard V, Torrejon Torres R, Davis JA, Saunders R. Critical Review of European Health-Economic Guidelines for the Health Technology Assessment of Medical Devices. Front Med (Lausanne). 2019;6:278.
123. Erdos J, Ettinger S, Mayer-Ferbas J, de Villiers C, Wild C. European Collaboration in Health Technology Assessment (HTA): goals, methods and outcomes with specific focus on medical devices. Wien Med Wochenschr. 2019;169(11-12):284-92.
124. Tufts Medical Center IfCRaHPS, Center for the Evaluation of Value and Risk in Health. The Cost-Effectiveness Analysis Registry2019 15th March 2019. Available from: http://www.cearegistry.org.
125. EUnetHtA Joint Action WPS, Heintz E, Gerber-Grote A, Ghabri S, Hamers FF, Rupel VP, et al. Is There a European View on Health Economic Evaluations? Results from a Synopsis of Methodological Guidelines Used in the EUnetHTA Partner Countries. Pharmacoeconomics. 2016;34(1):59-76.
126. Sorenson C, Tarricone R, Siebert M, Drummond M. Applying health economics for policy decision making: do devices differ from drugs? Europace. 2011;13 Suppl 2:ii54-8.
127. Taylor RS, Iglesias CP. Assessing the clinical and cost-effectiveness of medical devices and drugs: are they that different? Value Health. 2009;12(4):404-6.
128. Craig JA, Carr L, Hutton J, Glanville J, Iglesias CP, Sims AJ. A review of the economic tools for assessing new medical devices. Appl Health Econ Health Policy. 2015;13(1):15-27.
129. Gialama F, Prezerakos P, Apostolopoulos V, Maniadakis N. Systematic review of the cost-effectiveness of transcatheter interventions for valvular heart disease. Eur Heart J Qual Care Clin Outcomes. 2018;4(2):81-90.
130. Tarricone R, Callea G, Ogorevc M, Prevolnik Rupel V. Improving the Methods for the Economic Evaluation of Medical Devices. Health Econ. 2017;26 Suppl 1:70-92.
131. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
132. Odnoletkova I, Goderis G, Pil L, Nobels F, Aertgeerts B, Annemans L, et al. Cost-Effectiveness of Therapeutic Education to Prevent the Development and Progression of Type 2 Diabetes: Systematic Review. J Diabetes Metab. 2014;5(9).

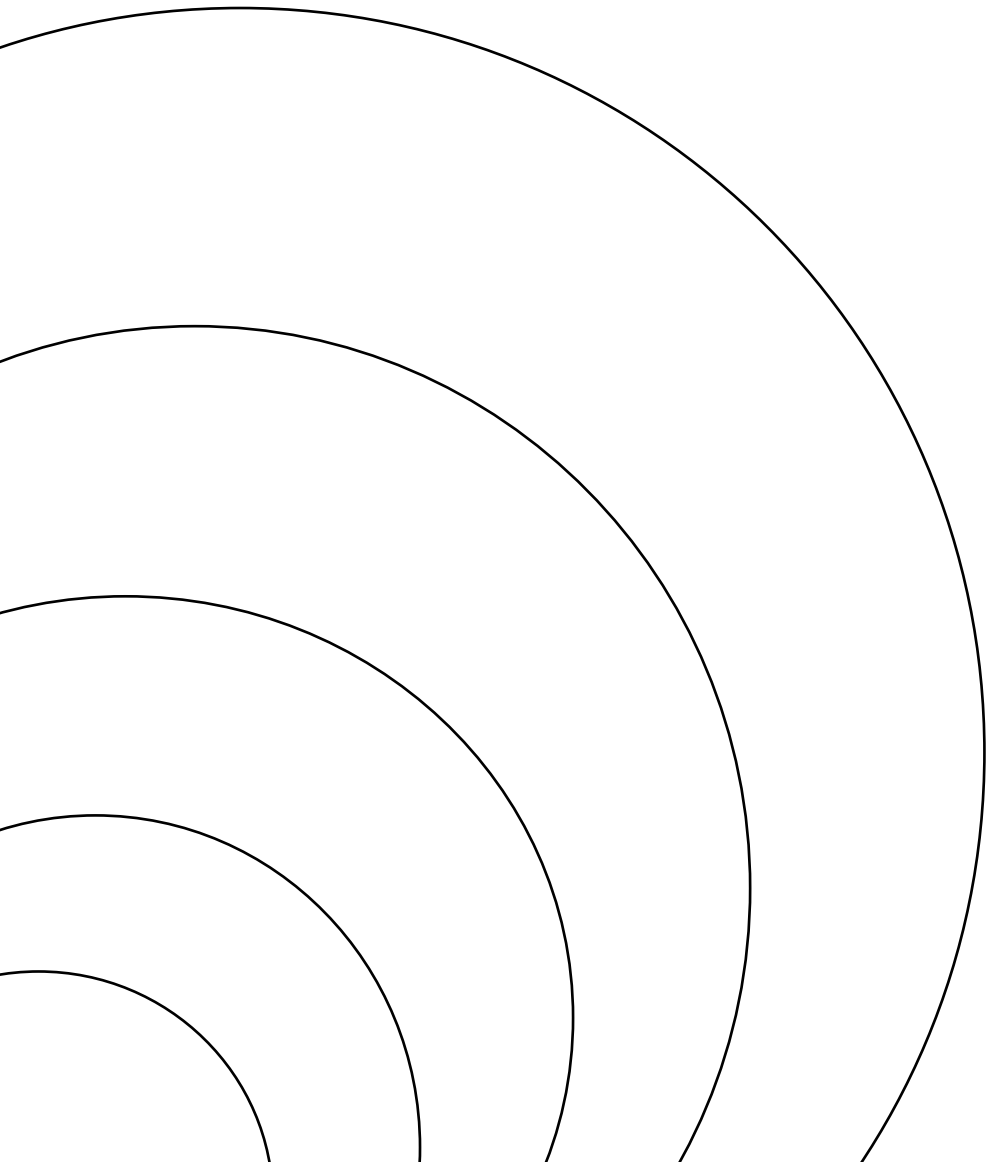
133. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. 2005;21(2):240-5.
134. Sehatzadeh S, Doble B, Xie F, Blackhouse G, Campbell K, Kaulback K, et al. Transcatheter aortic valve implantation (TAVI) for treatment of aortic valve stenosis: an evidence update. Ont Health Technol Assess Ser. 2013;13(1):1-40.
135. Sehatzadeh S, Doble B, Xie F, Blackhouse G, Campbell K, Kaulback K, et al. Transcatheter aortic valve implantation (TAVI) for treatment of aortic valve stenosis: an evidence-based Analysis (part B). Ont Health Technol Assess Ser. 2012;12(14):1-62.
136. Doble B, Blackhouse G, Goeree R, Xie F. Cost-effectiveness of the Edwards SAPIEN transcatheter heart valve compared with standard management and surgical aortic valve replacement in patients with severe symptomatic aortic stenosis: a Canadian perspective. J Thorac Cardiovasc Surg. 2013;146(1):52-60 e3.
137. Gada H, Kapadia SR, Tuzcu EM, Svensson LG, Marwick TH. Markov model for selection of aortic valve replacement versus transcatheter aortic valve implantation (without replacement) in high-risk patients. The American journal of cardiology. 2012;109(9):1326-33.
138. Ribera A, Slof J, Andrea R, Falces C, Gutierrez E, Del Valle-Fernandez R, et al. Transfemoral transcatheter aortic valve replacement compared with surgical replacement in patients with severe aortic stenosis and comparable risk: cost-utility and its determinants. International journal of cardiology. 2015;182:321-8.
139. Reynolds MR, Lei Y, Wang K, Chinnakondepalli K, Vilain KA, Magnuson EA, et al. Cost-Effectiveness of Transcatheter Aortic Valve Replacement With a Self-Expanding Prosthesis Versus Surgical Aortic Valve Replacement. Journal of the American College of Cardiology. 2016;67(1):29-38.
140. Reynolds MR, Magnuson EA, Lei Y, Wang K, Vilain K, Li H, et al. Cost-effectiveness of transcatheter aortic valve replacement compared with surgical aortic valve replacement in high-risk patients with severe aortic stenosis: results of the PARTNER (Placement of Aortic Transcatheter Valves) trial (Cohort A). Journal of the American College of Cardiology. 2012;60(25):2683-92.
141. Neyt M, Van Brabandt H, Devriese S, Van De Sande S. A cost-utility analysis of transcatheter aortic valve implantation in Belgium: focusing on a well-defined and identifiable population. BMJ open. 2012;2(3).
142. Tam DY, Hughes A, Fremes SE, Youn S, Hancock-Howard RL, Coyte PC, et al. A cost-utility analysis of transcatheter versus surgical aortic valve replacement for the treatment of aortic stenosis in the population with intermediate surgical risk. The Journal of thoracic and cardiovascular surgery. 2018;155(5):1978-88.e1.
143. Orlando R, Pennant M, Rooney S, Khogali S, Bayliss S, Hassan A, et al. Cost-effectiveness of transcatheter aortic valve implantation (TAVI) for aortic stenosis in patients who are high risk or contraindicated for surgery: a model-based economic evaluation. Health technology assessment (Winchester, England). 2013;17(33):1-86.
144. Gada H, Agarwal S, Marwick TH. Perspective on the cost-effectiveness of transapical aortic valve

implantation in high-risk patients: Outcomes of a decision-analytic model. Annals of cardiothoracic surgery. 2012;1(2):145-55.
145. Tam DY, Hughes A, Wijeysundera HC, Fremes SE. Cost-Effectiveness of Self-Expandable Transcatheter Aortic Valves in Intermediate-Risk Patients. The Annals of thoracic surgery. 2018;106(3):676-83.
146. Van Gestel R, Muller T, Bosmans J. Does my high blood pressure improve your survival? Overall and subgroup learning curves in health. Health economics. 2017;26(9):1094-109.
147. Alli OO, Booker JD, Lennon RJ, Greason KL, Rihal CS, Holmes DR, Jr. Transcatheter aortic valve implantation: assessing the learning curve. JACC Cardiovascular interventions. 2012;5(1):72-9.
148. Badheka AO, Patel NJ, Panaich SS, Patel SV, Jhamnani S, Singh V, et al. Effect of Hospital Volume on Outcomes of Transcatheter Aortic Valve Implantation. The American journal of cardiology. 2015;116(4):587-94.
149. Lunardi M, Pesarini G, Zivelonghi C, Piccoli A, Geremia G, Ariotti S, et al. Clinical outcomes of transcatheter aortic valve implantation: from learning curve to proficiency. Open heart. 2016;3(2):e000420.
150. Van de Wetering G, Woertman WH, Adang EM. A model to correct for short-run inefficiencies in economic evaluations in healthcare. Health economics. 2012;21(3):270-81.
151. Adang EM. Economic evaluation of innovative technologies in health care should include a short-run perspective. The European journal of health economics : HEPAC : health economics in prevention and care. 2008;9(4):381-4.
152. Barbanti M, Gulino S, Costa G, Tamburino C. Pathophysiology, incidence and predictors of conduction disturbances during Transcatheter Aortic Valve Implantation. Expert review of medical devices. 2017;14(2):135-47.
153. Barbanti M, Gulino S, Costa G, Tamburino C. Optimization and simplification of transcatheter aortic valve implantation therapy. Expert review of cardiovascular therapy. 2018;16(4):287-96.
154. Chevreul K, Brunn M, Cadier B, Haour G, Eltchaninoff H, Prat A, et al. Cost of transcatheter aortic valve implantation and factors associated with higher hospital stay cost in patients of the FRANCE (FRench Aortic National CoreValve and Edwards) registry. Archives of cardiovascular diseases. 2013;106(4):209-19.
155. Fuchs S, Olberg B, Perleth M, Busse R, Panteli D. Testing a new taxonomic model for the assessment of medical devices: Is it plausible and applicable? Insights from HTA reports and interviews with HTA institutions in Europe. Health Policy. 2019;123(2):173-81.
156. Doble B, Blackhouse G, Goeree R, Xie F. Cost-effectiveness of the Edwards SAPIEN transcatheter heart valve compared with standard management and surgical aortic valve replacement in patients with severe symptomatic aortic stenosis: a Canadian perspective. The Journal of thoracic and cardiovascular surgery. 2013;146(1):52-60.e3.
157. Fairbairn TA, Meads DM, Hulme C, Mather AN, Plein S, Blackman DJ, et al. The cost-effectiveness of transcatheter aortic valve implantation versus surgical aortic valve replacement in patients with severe aortic stenosis at high operative risk. Heart (British Cardiac Society). 2013;99(13):914-20.
158. Transcatheter Aortic Valve Implantation for Treatment of Aortic Valve Stenosis: A Health

Technology Assessment. Ontario Health Technology Assessment Series. 2016/12/19 ed2016. p. 1-94.
159. Kodera S, Kiyosue A, Ando J, Komuro I. Cost effectiveness of transcatheter aortic valve implantation in patients with aortic stenosis in Japan. Journal of cardiology. 2018;71(3):223-9.
160. Rehm J, Shield KD. Global Burden of Disease and the Impact of Mental and Addictive Disorders. Curr Psychiatry Rep. 2019;21(2):10.
161. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry. 2016;3(5):415-24.
162. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
163. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002;21(2):271-92.
164. NICE. Guide to the methods of technology appraisal 2013. London: National Institute for Health and Care Excellence; 2013 4 April 2013.
165. van Krugten FCW, Busschbach JJV, Versteegh MM, Hakkaart-van Roijen L, Brouwer WBF. The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. Qual Life Res. 2022;31(2):633-43.
166. Kleinman A. Rethinking Psychiatry: From Cultural Category to Personal Experience. New York: Free Press (Macmillan Publishing Co.); 1988. 454- p.
167. Sabat I, Neuman-Bohme S, Varghese NE, Barros PP, Brouwer W, van Exel J, et al. United but divided: Policy responses and people's perceptions in the EU during the COVID-19 outbreak. Health Policy. 2020;124(9):909-18.
168. Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, et al. Scoring the Icecap-a capability instrument. Estimation of a UK general population tariff. Health Econ. 2015;24(3):258-69.
169. Rohrbach PJ, Dingemans AE, Essers BA, Van Furth EF, Spinhoven P, Groothuis-Oudshoorn CGM, et al. The ICECAP-A instrument for capabilities: assessment of construct validity and test-retest reliability in a general Dutch population. Qual Life Res. 2021.
170. Kroenke K, Spitzer RL, Williams JB, Lowe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics. 2009;50(6):613-21.
171. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53-5.
172. Hoefman RJ, van Exel J, Brouwer WB. Measuring the impact of caregiving on informal carers: a construct validation study of the CarerQol instrument. Health Qual Life Outcomes. 2013;11:173.
173. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. Qual Life Res. 2014;23(3):977-89.
174. Devlin N, Parkin D, Janssen B. Analysis of EQ VAS Data. Methods for Analysing and Reporting EQ-5D Data. Cham (CH): Springer; 2020. p. 51-9.
175. Ohrnberger J, Fichera E, Sutton M. The relationship between physical and mental health: A mediation analysis. Soc Sci Med. 2017;195:42-9.

176. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22.
177. Wailoo A, Alava MH, Pudney S, Barton G, O'Dwyer J, Gomes M, et al. An International Comparison of EQ-5D-5L and EQ-5D-3L for Use in Cost-Effectiveness Analysis. Value Health. 2021;24(4):568-74.
178. M MV, K MV, S MAAE, de Wit GA, Prenger R, E AS. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52.
179. Coppens E, Van Audenhove C, Scheerder G, Arensman E, Coffey C, Costa S, et al. Public attitudes toward depression and help-seeking in four European countries baseline survey prior to the OSPI-Europe intervention. J Affect Disord. 2013;150(2):320-9.
180. Janssen B, Szende A. Population Norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, editors. Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht: Springer Netherlands; 2014. p. 19-30.
181. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. Eur J Health Econ. 2018;19(1):123-52.
182. Hackert MQN, van Exel J, Brouwer WBF. Content validation of the Well-being of Older People measure (WOOP). Health Qual Life Outcomes. 2021;19(1):200.
183. Oortwijn W, Jansen M, Baltussen R. Use of Evidence-Informed Deliberative Processes by Health Technology Assessment Agencies Around the Globe. Int J Health Policy Manag. 2020;9(1):27-33.
184. Pekarsky BAK. The Inclusion of Comparative Environmental Impact in Health Technology Assessment: Practical Barriers and Unintended Consequences. Appl Health Econ Health Policy. 2020;18(5):597-9.
185. Ossebaard HC, Lachman P. Climate change, environmental sustainability and health care quality. Int J Qual Health Care. 2021;33(1).
186. Daniels N, Sabin JE. Accountability for reasonableness: an update. BMJ. 2008;337:a1850.

Samenvatting



De collectieve zorguitgaven zijn de afgelopen decennia sterker gestegen dan de overige collectieve uitgaven. Deze groei roept de vraag op of de zorg op termijn betaalbaar blijft en benadrukt de noodzaak om de uitgaven te rechtvaardigen. Deze uitgaven hadden immers ook aan bijvoorbeeld onderwijs of sociale zekerheid kunnen worden besteed. Er is dan ook behoefte aan doeltreffende maatregelen om de zorguitgaven te beheersen.

Om de collectieve zorguitgaven beter te beheersen, kunnen verschillende beleidsinstrumenten worden gebruikt. Sommige daarvan, zoals budgettering, hebben voornamelijk als doel de betaalbaarheid van de zorg te waarborgen. Zulke maatregelen kunnen echter ongewenste effecten hebben op de kwaliteit en toegankelijkheid van de zorg. Een alternatieve aanpak is om de inhoud van het basispakket te begrenzen aan de hand van de expliciete afweging van kosten en baten van zorginterventies. Veel landen met een collectief gefinancierd gezondheidszorgsysteem bepalen op die manier, aan de hand van criteria die aansluiten bij hun publieke doelen en waarden, welke interventies wel en niet vergoed moeten worden. Om deze, vaak moeilijke en controversiële besluitvorming te ondersteunen en te structureren, zijn besluitvormingskaders ontwikkeld. Deze kaders zijn gebaseerd op een proces dat bekend staat als health technology assessment (HTA). HTA en de daarop gebaseerde besluitvormingskaders worden in de praktijk vaak toegepast en steeds verder verfijnd. Ze worden voornamelijk toegepast voor besluiten over geneesmiddelen. Desalniettemin zijn deze kaders nog niet optimaal of volledig ontwikkeld in alle opzichten. Zo kunnen de gebruikte pakketcriteria, zoals "noodzakelijkheid", verder geoperationaliseerd worden en kunnen aanvullende criteria worden toegevoegd. Bovendien zijn de huidige kaders en onderliggende methodieken mogelijk niet optimaal voor het beoordelen van andere (niet-geneesmiddelen) interventies, zoals chirurgische ingrepen. Deze andere interventies worden vaak zonder expliciete

beoordeling in het basispakket opgenomen, terwijl deze een veel groter deel van de zorguitgaven vertegenwoordigen dan geneesmiddelen. Dit kan problematisch zijn, aangezien een meer systematische beoordeling van deze andere zorginterventies aanzienlijk kan bijdragen aan de benodigde beheersing van de uitgaven in lijn met de doelstellingen van het zorgstelsel.

Het centrale doel van dit proefschrift is om een bijdrage te leveren aan de optimalisering van het besluitvormingskader en de onderliggende HTA-methodologie die gebruikt worden voor beslissingen over opname van interventies in het basispakket. Daarvoor zijn twee belangrijke thema's onderzocht: het verbreden van de getoetste pakketcriteria door "winstgevendheid voor de fabrikant" toe te voegen (hoofdstukken 2 en 3), en het verbreden van het gebruik van HTA naar andere soorten zorginterventies (hoofdstukken 4, 5 en 6).

Hoofdstuk 2 onderzocht of HTA-organisaties de kosten die geneesmiddelfabrikanten maakten in relatie tot de prijs die ze voor een duur geneesmiddel vroegen (winstgevendheid), expliciet meewogen in vergoedingsbeslissingen. Hiervoor werden rapporten, gepubliceerd in vier landen (Australië, Canada, Engeland, Nederland), systematisch verzameld en geanalyseerd. De resultaten hiervan lieten zien dat in HTA-rapporten over dure geneesmiddelen de kosten van fabrikanten in relatie tot de prijs van geneesmiddelen geen systematische aandacht kregen. Dit ondanks de aandacht voor winstgevendheid in de wetenschappelijke literatuur en in publieke debatten. Wel deden HTA-organisaties in enkele rapporten de uitnodiging aan fabrikanten om hoge prijzen te rechtvaardigen op basis van de door hen gemaakte kosten.

Hoofdstuk 3 onderzocht vervolgens of, en in welke mate, beleidsmakers informatie over winstmarges betrokken bij hypothetische

vergoedingsbeslissingen. Hiervoor werd via een online vragenlijst aan verschillende Nederlandse beleidsmakers een aantal vergoedingsbeslissingen voorgelegd. De resultaten toonden aan dat wanneer winstmarge naast de gebruikelijke pakketcriteria (effectiviteit, noodzakelijkheid, kosteneffectiviteit en uitvoerbaarheid) werd gepresenteerd, deze invloed had op de beslissing, waarbij hogere winstmarges de kans op vergoeding verlaagden. De invloed van het criterium winstmarge op de keuzen was in vergelijking met de andere pakketcriteria relatief laag, maar niet verwaarloosbaar. Ook gaf meer dan de helft van de respondenten aan dat winstmarges een rol zouden moeten spelen in de besluitvorming over vergoeding.

In **hoofdstuk 4** werden uitdagingen verkend voor Zorginstituut Nederland (ZIN), een HTA-organisatie, wanneer deze het gebruik van hun besluitvormingskader en de onderliggende HTA-methodologie zou willen verbreden naar een aantal andere zorginterventies. Dit betrof verbreding naar intramurale geneesmiddelen, medische hulpmiddelen, curatieve ingrepen, niet-farmaceutische curatieve geestelijke gezondheidszorg en niet-curatieve zorg (langdurige zorg, waaronder ouderenzorg). Het hoofdstuk belichtte vijf belangrijke kenmerken van extramurale geneesmiddelen die het gebruik van HTA vergemakkelijken en stimuleren. Deze kenmerken waren (I) het gesloten vergoedingssysteem, (II) de afwezigheid van alternatieve beleidsmaatregelen, (III) het bestaan van een systeem van markttoelating, (IV) aanwezigheid van een identificeerbare en aanspreekbare marktpartij en (V) specifieke productkenmerken. Voor elk van de andere soorten zorginterventies werden uitdagingen bij het gebruik van HTA in relatie tot deze vijf kenmerken verkend. Deze uitdagingen hadden betrekking op alle fasen van het besluitvormingsproces. Ze omvatten onder meer de noodzaak van de identificatie van te beoordelen zorginterventies, het niet beschikbaar zijn van bewijs

(onderzoeksresultaten) en verschillende specifieke methodologische uitdagingen. Belangrijk is dat bij sommige soorten zorginterventies meer uitdagingen werden verwacht dan bij andere. Voor een aantal uitdagingen werden oplossingen geopperd, variërend van horizon scanning, de invoering van een "sluis" voor medische hulpmiddelen en overheidsfinanciering van onderzoek, tot het ontwikkelen van alternatieve, meer adaptieve HTA-processen.

Hoofdstuk 5 beschreef een casus over methodologische uitdagingen die spelen wanneer HTA-onderzoek wordt toegepast bij medische hulpmiddelen. Medische hulpmiddelen wijken op meerdere manieren af van geneesmiddelen waardoor het doen van HTA-onderzoek uitdagend is. Zo zijn de kosten en baten van de inzet van medische hulpmiddelen meer contextafhankelijk, zijn ze vaak onderhevig aan incrementele innovatie en is er soms sprake van leereffecten bij het gebruik. Onderzocht werd in hoeverre binnen gepubliceerde economische evaluaties rekening is gehouden met deze specifieke kenmerken. Hiervoor werd Transcatheter Aortic Valve Implantation (TAVI) als casus gebruikt. Geconstateerd werd dat de genoemde kenmerken en hun invloed op de validiteit van de gerapporteerde resultaten in de onderzochte economische evaluaties van TAVI over het algemeen slechts gedeeltelijk werden besproken en zelden kwantitatief werden behandeld. Twee van de vijftien geïncludeerde studies boden een analytische oplossing voor één van de uitdagingen. Deze beperkte aandacht voor de specifieke kenmerken van medische hulpmiddelen binnen de onderzochte economische evaluaties lijkt ongewenst aangezien dit kan leiden tot suboptimale beslissingen.

Hoofdstuk 6 ging over een aspect van de verbreding van de toepassing van HTA-onderzoek naar de geestelijke gezondheidszorg, namelijk adequate uitkomstmeting in die context. Onderzocht werd of een

uitkomstmaat die specifiek voor deze context is ontwikkeld, het Mental Health Quality of Life (MHQoL) instrument, kwaliteit van leven (KvL) betrouwbaar en valide meet. De MHQoL werd ontwikkeld om te worden gebruikt voor het beschrijven en waarderen van de mentale gezondheidsgerelateerde KvL van respondenten. Op basis van bestaande internationale, cross-sectionele gegevens werd een psychometrische evaluatie van de MHQoL uitgevoerd. De resultaten voor zowel de interne consistentie van de MHQoL als voor de constructvaliditeit van de MHQoL, suggereerden dat de MHQoL een betrouwbaar en valide meetinstrument is voor het meten van KvL bij mensen met problemen met hun mentale gezondheid.

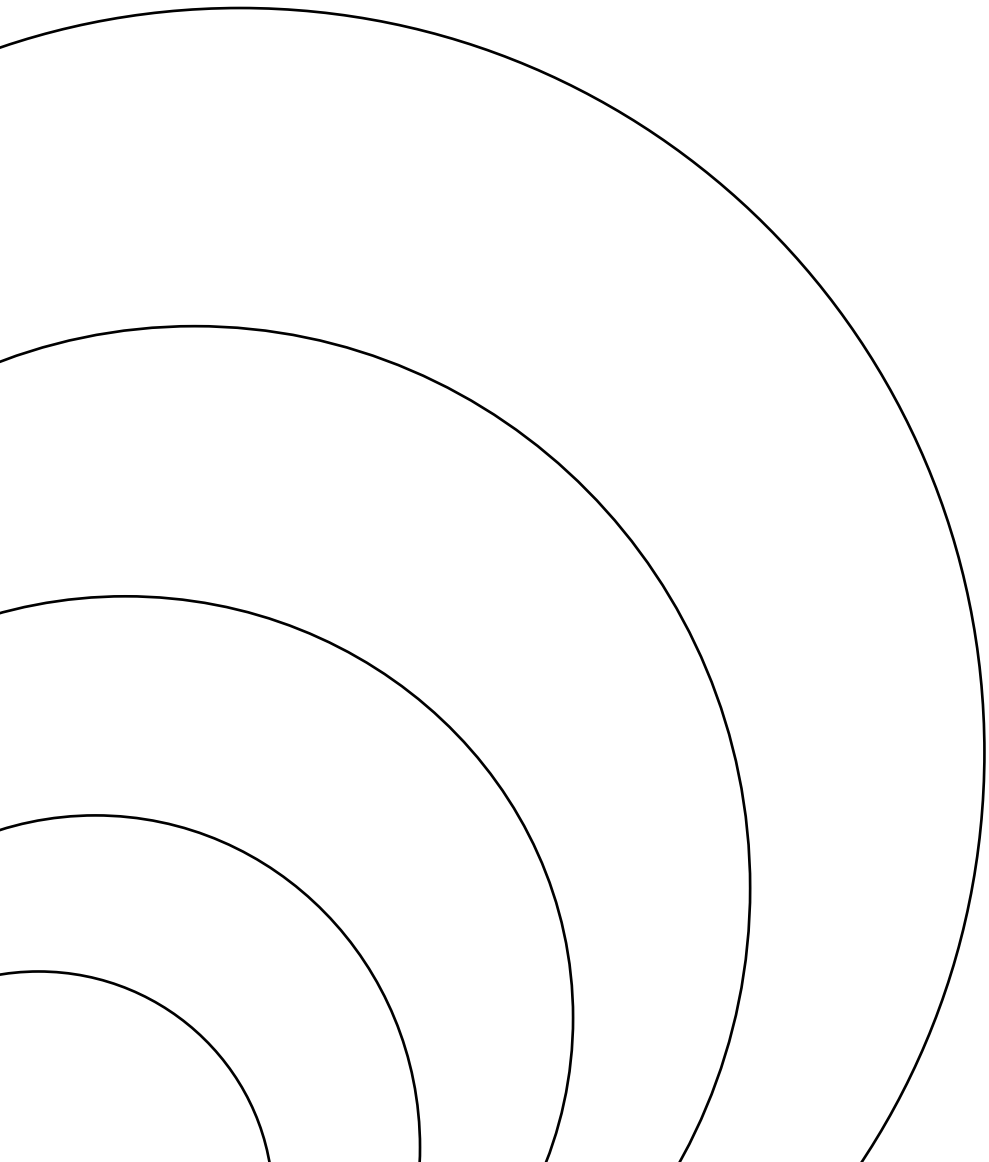
Op basis van de hoofdstukken 2 tot en met 6 trokken we een aantal conclusies. Ten eerste concludeerden we dat "winstgevendheid" niet gangbaar is als pakketcriterium, maar vaak wel als relevant wordt gezien en invloed kan hebben op pakketbeslissingen. Daarmee is “winstgevendheid” voor HTA-organisaties een serieus te overwegen aanvulling op de bestaande pakketcriteria. Onderzoek naar de maatschappelijke wenselijkheid en de praktische haalbaarheid van deze aanvulling is geboden.

Ten tweede concludeerden we dat als het doel van HTA-organisaties en beleidsmakers is om HTA in de toekomst breder en systematischer te gebruiken, het noodzakelijk is om het huidige besluitvormingskader en de onderliggende HTA-methodologie daar geschikt voor te maken. De veelheid en complexiteit van de geïdentificeerde uitdagingen en hun variatie tussen zorgdomeinen, benadrukken de noodzaak van prioritering bij een overgang van de huidige praktijk naar een breed gebruik van HTA in alle zorgdomeinen. Dit temeer omdat sommige uitdagingen bij de verbreding lastiger op te lossen zijn dan andere. Dit proefschrift biedt een basis voor deze benodigde prioritering.

Ten derde concludeerden we dat er een duidelijke relatie bestaat tussen HTA-methodologie en het besluitvormings- en beleidsproces. Als bepaalde methodologische uitdagingen niet adequaat worden opgelost, legt dit meer druk op besluitvormers. Dit benadrukt de noodzaak om beleids- en methodologische ontwikkeling hand in hand te laten gaan.

Dit proefschrift draagt hiermee bij aan de verdere optimalisering van het besluitvormingskader en de onderliggende HTA-methodologie die worden gebruikt bij beslissingen over het opnemen van interventies in het basispakket. Het uiteindelijke doel is het mogelijk maken om voor een groter deel van het basispakket expliciet de relevante criteria te wegen. Dit kan helpen om het zorgstelsel en zorguitgaven zich in lijn met publieke waarden te laten ontwikkelen. De weg daarnaartoe is uitdagend, maar de bestemming is de reis meer dan waard!

Summary



In recent decades, public healthcare expenditures have outpaced other public expenditures, raising concerns about the affordability of healthcare. The growing share of total wealth spent on healthcare necessitates the need to justify these expenditures, as they compete with funding for sectors like education and social security. Consequently, there is a demand for effective measures to control healthcare spending.

Various policy instruments can be employed to control healthcare expenditures. Some, such as budgeting, primarily focus on ensuring affordability of care. However, these measures may inadvertently impact the quality and accessibility of healthcare. Another approach involves setting deliberate boundaries for the basic benefit package (BBP) to control expenditures while considering the broader goals of the healthcare sector, rather than solely prioritizing affordability. In this way, many countries with collectively financed healthcare systems determine which pharmaceuticals are eligible for reimbursement based on criteria aligned with their public goals and values. To support and structure the decision-making process in this regard, frameworks based on health technology assessment (HTA) have been developed. HTA and its associated decision-making frameworks have been widely used and tested, especially in relation to pharmaceuticals. However, there is room for improvement and expansion of these frameworks. For instance, operationalizing common decision criteria like "necessity" requires attention, and the set of considered criteria can be broadened. Moreover, the current frameworks and underlying methodologies may not be optimal for assessing healthcare interventions other than pharmaceuticals, such as surgical procedures, which are often included in the BBP without explicit assessment. This limitation is problematic because a more systematic consideration of these interventions, which collectively account for a significant portion of total healthcare expenditure, could greatly contribute to controlling costs.

The central aim of this thesis is to contribute to the optimization of the decision-making framework and underlying HTA methodology used for determining the inclusion of healthcare interventions in the BBP. The research covers two main areas: broadening the scope of HTA by considering "profitability to the manufacturer" as a criterion (Chapters 2 and 3), and exploring the expansion of the use of HTA beyond pharmaceuticals to other healthcare interventions (Chapters 4, 5, and 6).

Chapter 2 examined whether HTA organisations explicitly considered the costs incurred by manufacturers in relation to the prices they charged for expensive pharmaceuticals. The analysis of assessment reports from four countries (Australia, Canada, England, the Netherlands) revealed that although manufacturers' costs in relation to price were discussed in scientific literature and public debates, they were not systematically addressed in HTA reports regarding expensive pharmaceuticals. However, some reports invited manufacturers to justify high prices by demonstrating high costs.

In **Chapter 3**, Dutch policymakers were presented with hypothetical reimbursement decisions involving different pharmaceutical products with different profit margins. The influence of profit margins on their reimbursement decisions was assessed through an online survey. The results indicated that when profit margin information was presented alongside other decision criteria (effectiveness, disease severity, cost-effectiveness, and budget impact), it influenced the decisions made, with higher profit margins reducing the probability of reimbursement. Profit margin was considered less important than other criteria, but its influence

was not negligible. In addition, more than half of the respondents indicated that profit margins should play a role in reimbursement decisions.

Chapter 4 explored the challenges faced by Zorginstituut Nederland (ZIN), an HTA organisation, in broadening the use of their decision-making framework and underlying HTA methodology for various healthcare interventions. These interventions included inpatient pharmaceuticals, medical devices, curative interventions, non-pharmaceutical curative mental healthcare, and non-curative care (including elderly care). The chapter highlighted five important characteristics of extramural pharmaceuticals that facilitate and encourage the use of HTA for these types of interventions. These characteristics were as follows: (I) a closed reimbursement system, (II) the absence of alternative policy measures, (III) the existence of marketing authorization, (IV) the presence of an identifiable and accountable counterparty, and (V) the product characteristics. The challenges associated with using HTA in absence of (some of) these characteristics were discussed for each type of healthcare intervention. These challenges were found to be relevant at all stages of the decision-making process and included the need for intervention identification, lack of evidence, and specific methodological challenges in performing HTA. Certain types of healthcare interventions were expected to be associated with more challenges than others. The chapter suggested possible solutions for some of the identified challenges, such as horizon scanning, implementing a "lock" for medical devices, public funding for evidence generation, and the development of alternative adaptive HTA processes.

Chapter 5 focused on the methodological challenges that arise when expanding the use of HTA research to medical devices. Medical devices differ from pharmaceuticals in ways that make conducting HTA research

challenging, such as the context-dependent nature of costs and benefits, the potential for incremental innovation, and the occurrence of learning

effects in the early stages of use. The chapter investigated the extent to which these specific characteristics were taken into account in economic evaluations, using Transcatheter Aortic Valve Implantation (TAVI) as a case study. The findings revealed that these characteristics and their impact on the validity of reported results were only partially discussed and rarely quantitatively addressed in the reviewed economic evaluations of TAVI. However, two included studies offered analytical solutions to some of these challenges. The limited focus on the specific characteristics of medical devices within economic evaluations is a clear point of attention, as it may lead to misinformed decisions.

Chapter 6 dealt with one aspect of the broadening the use of HTA research to mental health, namely the adequate outcome measurement in that context. It was investigated whether an outcome measure specifically developed for this context, the Mental Health Quality of Life (MHQoL) instrument, measures quality of life (QoL) reliably and validly. The MHQoL is designed to describe and later value current mental health-related QoL of respondents. Based on existing international, cross-sectional data, a psychometric evaluation of the MHQoL was performed. The results for both the internal consistency of the MHQoL and the construct validity of the MHQoL suggested that the MHQoL is a reliable and valid measure of QoL in people with mental health problems.

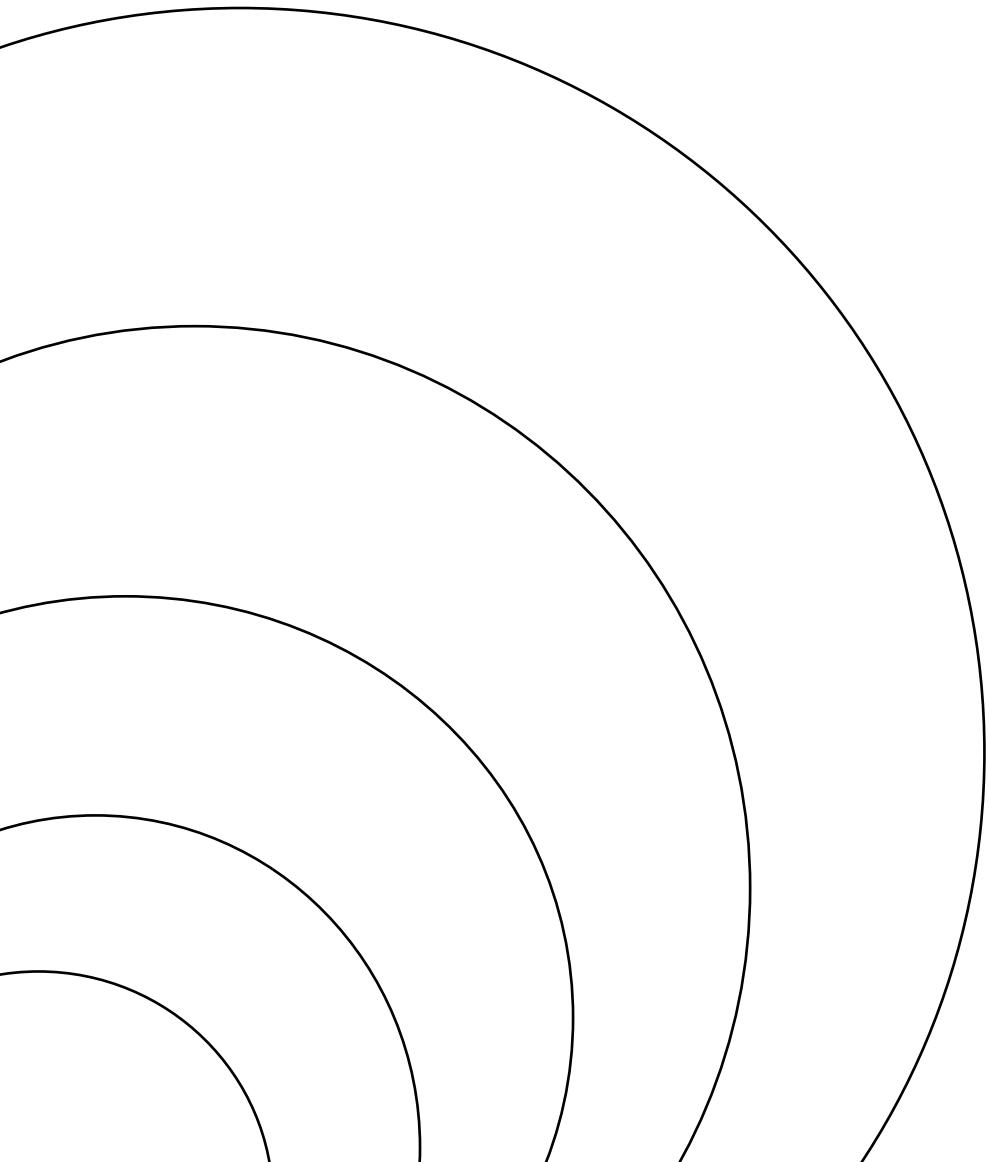
Based on chapters 2 to 6, we drew a number of conclusions. First, we concluded that "profitability" is not a common package criterion, but is often seen as relevant and may influence package decisions. Hence, HTA organisations should therefore seriously consider whether "profitability" should be added to the existing package criteria. Research into the societal desirability and practical feasibility of this addition is required.

Second, we concluded that if the goal of HTA organisations and policy makers is to use HTA more widely and systematically in the future, it is necessary to adapt the current decision-making framework and underlying HTA methodology to allow this. The multiplicity and complexity of the identified challenges and their variation between care domains emphasize the need for prioritization in a transition from current practice to the widespread use of HTA in all care domains. This is emphasized by the fact that some expansion challenges are more difficult to solve than others. This thesis provides a basis for this prioritization.

Third, we concluded that there is a clear relationship between HTA methodology and the decision-making and policy processes. If certain methodological aspects are not adequately addressed, this puts more pressure on decision makers, as the uncertainty surrounding specific estimates will increase and their relevance may decrease. This emphasizes the need for policy and methodological development to go hand in hand.

With these insights, this thesis contributes to the ongoing optimization of the decision-making framework and the underlying HTA methodology used for determining the inclusion of healthcare interventions in the BBP. Ultimately, it should be possible to explicitly weigh the relevant criteria for a larger part of the basic package and thus help to develop the health care system in line with public values, fostering sustainability. Although the path to achieving this goal is challenging, the destination justifies the journey.

Dankwoord



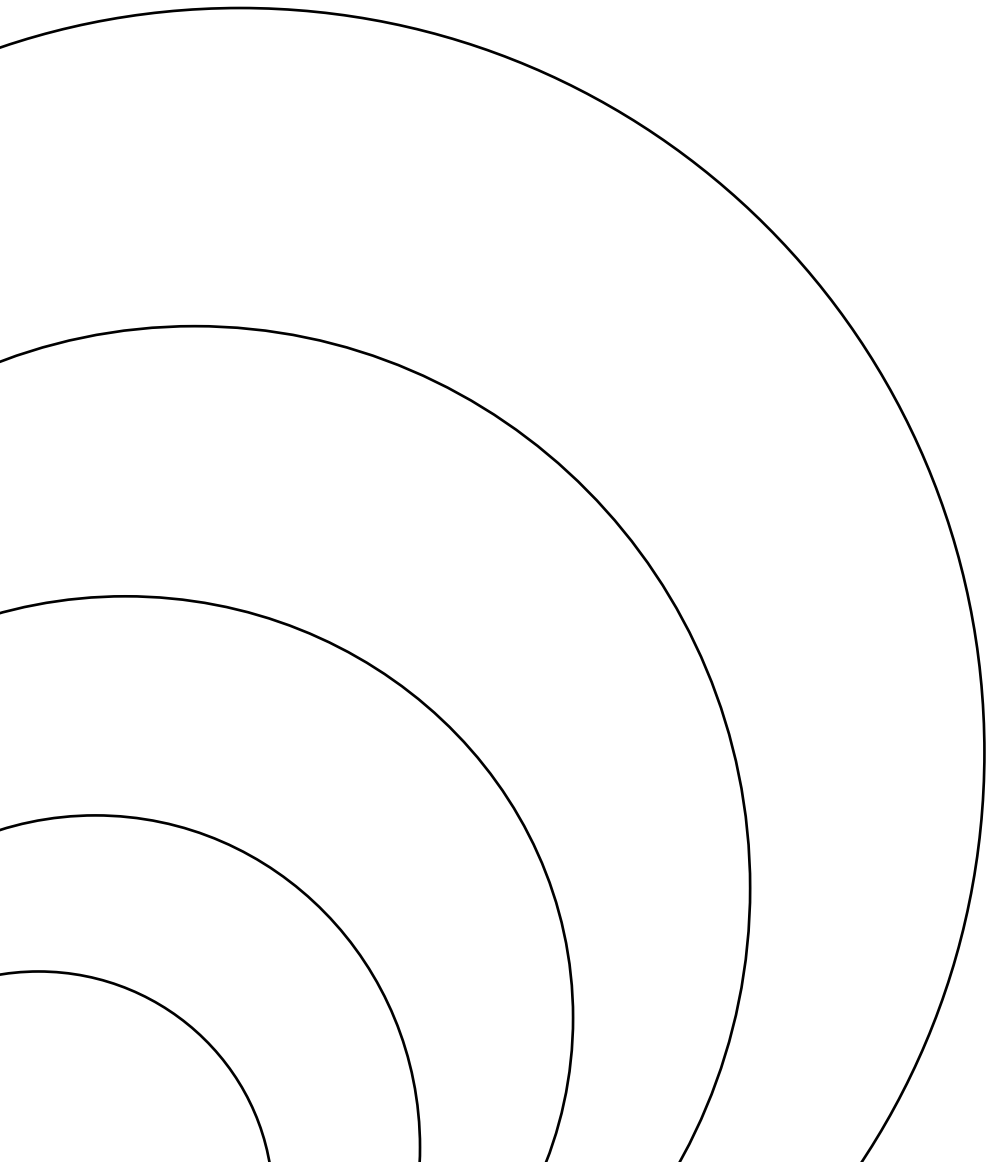
Het is gelukt. Het is volbracht. En dankbaar kijk ik terug. Die dankbaarheid gaat uit naar velen. Naar Sarah Kleijnen, die enthousiast reageerde op mijn wens om te starten met een promotie-traject. Naar Diana Delnoij die het vertrouwen gaf dit binnen de Academische werkplaats Verzekerde Zorg te doen. Naar Saskia Knies, die me vanaf de eerste dag heeft geholpen. Zij vormde samen met promotoren Werner Brouwer en Bert Boer het team waarmee we op onderzoek zijn gegaan. Aan hen drieën zeer veel dank, om wat ze deden en om wie ze zijn. Om hun humor, energie en expertise. Om hun betrokkenheid, relativeringsvermogen en geduld. Om het vinden van het juiste samenspel. Ze zagen me heen en weer rennen van onderwerp naar onderwerp. Om me met een verwijzing naar Top Gun te doen richten op één target per keer. Ze zagen me in worsteling met manuscripten waarvan de versies zich opstapelden. Om me met een slim duwtje weer los te krijgen. Zonder hen was het niet gelukt. Zonder hen was het niet zo'n waardevolle tijd geweest. Zonder de maandelijkse gesprekken, die soms over niets gingen en soms over alles. Bedankt.

Hierbij ook veel dank aan, Vielen Dank an, muito obrigado a, many thanks to, onze hierbij onmisbare coauteurs: Sylvia Vijgen, Sebastian Himmler, Frédérique van Krugten, Sebastian Neumann-Böhme, Iryna Sabat, Jop Engel, Maarten IJzerman, Beate Sander en Rick Vreman. Speciale dank daarnaast aan Vivian Reckers-Droog voor onze samenwerking die hopelijk nog lang zal doorlopen. Dank ook aan Joost Geenen aan wiens proefschrift ik mocht bijdragen. En natuurlijk aan alle respondenten en betrokken Zorginstituut collega's, zoals Jacqueline Zwaap en Angèl Link, zonder wiens hulp we niet zo ver waren gekomen. En niet in de laatste plaats aan de Raad van Bestuur van het Zorginstituut, in het bijzonder Sjaak Wijma, en het management van het Zorginstituut, in het bijzonder Timon Sibma, voor het creëren van de juiste omstandigheden, de lerende organisatie die we zijn. Een organisatie met veel inspirerende collega’s, binnen OWIZ, binnen

de Academische werkplaats Verzekerde Zorg, binnen het KE-team. Te veel om bij naam te noemen. Nou vooruit, eentje dan: Joke Derksen.

En dan is er nog een andere categorie onmisbare mensen. Zij hielpen mij door ruimte te bieden, door te inspireren en door af te leiden. Voorop staan Cathelijne, Neeltje, Teun en Ciske. Promoveren in tijden van corona kon alleen dankzij jullie. Caat, je bent alles voor me. Neeltje, Teun, Ciske, wat een geluk jullie vader te mogen zijn. Daarnaast veel dank aan mijn ouders en aan mijn zussen en (schoon)familie. Gemma, altijd creatief en onderzoekend, bedankt voor je warmte. Tjeerd, vriendelijke rationele rots in de branding, dank voor je praktische raad en daad. Maaike 2.0, je bent geweldig. Geerke, mijn voorbeeld in positief doorzetten. Ita, Kees, Fred, Marijke, bedankt voor wie jullie zijn voor mij en voor onze kinderen. Verder een shout-out richting mijn nichtjes Petra Havinga PhD en Siëlle Gramser PhD, die me recent voorgingen. Ten slotte dank ik mijn vrienden en bekenden. In het bijzonder Laura van Hasselt, voor onze PhD lunches in Krasnapolsky, Gerard van der Werf en Floor Oosting voor de wereldbeschouwingen aan de keukentafel en Nick van Loo, mede voor het helpen vormgeven.

PhD Portfolio



Training

2021

- Philosophy of the social sciences and the humanities, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands

2020

- Health Economic Modelling in R, Decision Analysis in R for Technologies in Health (DARTH) Workgroup

2019

- Professionalism and integrity in research, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands
- Self-presentation, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands

2018

- Academic Writing in English, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands
- Doing the literature review, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands

2017

- How to survive your PhD, Erasmus Graduate School of Social Sciences and the Humanities, Rotterdam, the Netherlands

Teaching

2022

- Practice session: Rationing health care in the Netherlands, bachelor programme in Health Policy & Management, Erasmus University Rotterdam

2021

- Practice session: Rationing health care in the Netherlands, bachelor programme in Health Policy & Management, Erasmus University Rotterdam

2018

- Guest lecture: Managing the basic benefit package, bachelor programme Beleid, Management en Evaluatie van Zorg, Maastricht University

Conferences, symposia

2023

- LolaHESG, Egmond aan Zee. Paper presentation
- NVTAG symposium: Where to go if you don't know? Uncertainty in HTA decision-making, Utrecht

2022

- NVTAG symposium: Fitting, or one-size-fits-all?, Utrecht
- FMS workshop for Slovenian delegacy, Utrecht. Oral presentation
- HTAi Annual Meeting, Utrecht. Oral presentation
- Young-NVTAG symposium HTA for Medical devices, Utrecht. Oral presentation
- LolaHESG, Maastricht. Paper presentation
- NVTAG symposium: 25 years NVTAG Back to the future, online

2021

- LolaHESG, online. Paper presentation

2020

- VGE/NVTAG webinar Participatory Value Evaluation, online
- LolaHESG, online. Paper presentation
- VGE/NVTAG symposium: De vele dimensies van gezondheid, online

2019

- ISPOR Europe, Kopenhagen
- LolaHESG, Almen. Paper presentation

2018

- ZIN/NVTAG Symposium: Verdringing in de zorg, Utrecht
- LolaHESG, Hoenderloo
- ESHPM symposium Pakketbeheer: kennis over keuzes, Rotterdam. Oral presentation

2017

- VGE/NVTAG symposium Value Based Health Care en kosteneffectiviteit: keerzijden van dezelfde euro?, Leiden

Miscellaneous

- ZorginstituutForum. Oral presentation (2023)
- Member of the NVTAG prize jury (2021 – 2023)
- Contributor to the ISPOR Good Practices Report on Critical Appraisal of Systematic Reviews With Costs and Cost-Effectiveness Outcomes (2020)
- Member of the NVTAG board (2017 – 2023)
- Member of the Research Network HTA NL (Academische Werkplaats Verzekerde Zorg) (2017 – 2023)

Scientific publications

Enzing JJ, Himmler S, Knies S, Brouwer WBF. Do Profit Margins of Pharmaceuticals Influence Reimbursement Decisions? A Discrete Choice Experiment Among Dutch Healthcare Decision Makers. Value Health. 2022;25(2):222-9.

Enzing JJ, Knies S, Boer B, Brouwer WBF. Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? Health Econ Policy Law. 2021;16(4):440-56.

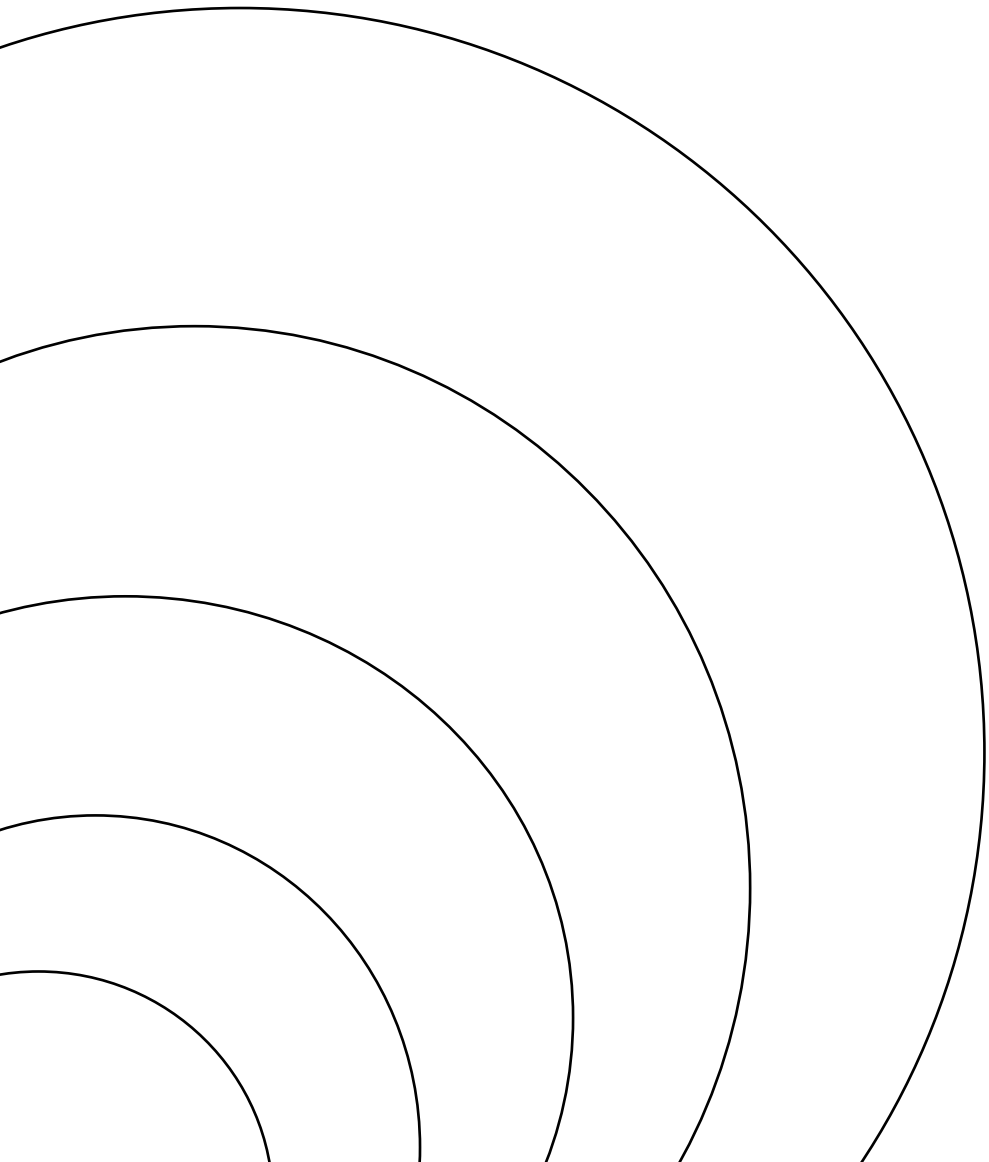
Enzing JJ, Knies S, Engel J, MJ IJ, Sander B, Vreman R, et al. Do Health Technology Assessment organisations consider manufacturers' costs in relation to drug price? A study of reimbursement reports. Cost Eff Resour Alloc. 2022;20(1):46.

Enzing JJ, van Krugten FCW, Sabat I, Neumann-Bohme S, Boer B, Knies S, et al. Psychometric evaluation of the Mental Health Quality of Life (MHQoL) instrument in seven European countries. Health Qual Life Outcomes. 2022;20(1):129.

Enzing JJ, Vijgen S, Knies S, Boer B, Brouwer WBF. Do economic evaluations of TAVI deal with learning effects, innovation, and context dependency? A review. Health Policy and Technology. 2021;10(1):111-9.

Geenen JW, Boersma C, Enzing J, Klungel OH, Hovels AM. Real options analysis to inform reimbursement decisions. (*Chapter 5.2 of Geenen, JW. Affordability, Cost-Effectiveness and Uncertainty: An Integrated Approach for Decision-Making on Medicines. Dissertation. Utrecht University Repository. 2019*)

About the author



Joost Enzing is a Dutch health economist who holds a master's degree in business economics from University of Groningen. He began his career as an information technology consultant. Presently, and for several years now, he serves as a policy advisor at the National Health Care Institute (ZIN), which has an advisory role regarding the content of the Dutch basic benefit package. At ZIN, Joost contributes to the utilization of economic evaluations in policy, conducts budget impact analyses, and assists in the development of health technology assessment methodologies. Notably, he was a member of the project team responsible for the Dutch guideline for economic evaluations in healthcare. Alongside his work at ZIN, Joost has been working on his thesis at the Erasmus School of Health Policy & Management (ESHPM), part of Erasmus University Rotterdam. Additionally, Joost holds a board position in the NVTAG (Dutch Association for Technology Assessment in Healthcare), an association dedicated to promoting the use of health technology assessment in the Dutch healthcare system.