Do we value rarity?

An explorative study on the WTP for rarity of diseases

by

Sophie van der Wild - 388119

Supervised by Linda de Vries

A Master Thesis

Submitted to Erasmus School of Health Policy and Management

Erasmus University Rotterdam

In Partial Fulfillment of the Requirements

for the Degree of Health Economics, Policy and Law with the specialization in Health Economics

September 20, 2021

Rotterdam

Words: 17,322

Acknowledgements

The time has finally come: In front of you lies my Master's thesis "Do we value rarity? An exploration on the WTP for rarity of diseases". It has been written to fulfill the graduation requirements of the Master of Science in Health Economics, Policy & Law and its specialization Health Economics at the Institute of Health Policy and Management of the Erasmus University Rotterdam.

I started this Master Program in September 2017, but dropped out in March 2018 because the passion and excitement within me was lacking. After becoming ill and needing to recover from my mental health problems from early 2019 onwards, I noticed I was missing my classes and this program. In September 2019, I tried combining my therapy together with this Master, which was quite challenging at times. Now, in September 2021, it is all coming together. I can finally say I am finishing my Master's degree and thereby closing such a big chapter in my life.

As the basis for this thesis originally stemmed from my interest in the valuation of different diseases and their treatments, seeing one of my friends actually having a rare disease was truly an eye opener. Especially when I found out how little is actually known about these types of diseases and their possible treatments.

I found the realization of the research and analyses discussed in this thesis truly challenging. In truth, I could not have achieved half of my thesis without the guidance of my supervisor Linda de Vries. I would like to thank her as she has been available and willing to answer my questions at any time, greatly enabling me to make progress until the very end. Even though some times were hard on me, our meetings always gave me new energy and your confidence in me was something I did not know I really needed. We did it! Words cannot express how grateful I am to have had you as my supervisor.

Also, my parents, siblings (in special my sister Carlijn) and closest friends deserve a particular note of thanks: as always, but most importantly in the last 2 years, your unwavering love and understanding have greatly supported me to follow this thesis through to its successful conclusion. Your help and energy, in its many forms, kept me going.

I hope you enjoy your reading.

Abstract

Background

No explicit research has been done on what the role of rarity of diseases is in society's WTP, while often a lot of (media) attention is drawn to negative reimbursement decisions. This attention implies there is a social opinion on the value of this type of care.

Objective

This thesis aimed to investigate whether there is a difference in willingness to pay (WTP) in the Netherlands for rare diseases and non-rare diseases (common and frequent occurring). This was done in a societal perspective.

Methods

Respondents valued different hypothetical treatments resulting in health changes by means of a contingent valuation survey. The health changes varied in terms of rarity of the disease to obtain the WTP for rarer and more frequently occurring diseases. The WTP as an increase in the monthly basic health insurance was obtained using a two-step elicitation procedure.

Results

150 respondents participated in the study. The mean WTP per QALY differed from €769,776 per QALY for rare diseases, to €38,889 per QALY for common diseases and €5,437 per QALY for frequent occurring diseases. Raw mean WTP's per treatment per month are €8.30; €8.74 and €11.34 respectively. These differences are significant. Regression analyses showed that there is a significant relationship between WTP, age and living in the Randstad (big cities). Given information on available alternative treatments for diseases is not of significant influence in the scenarios rare and common, but is significant for frequent occurring diseases.

Conclusion

This first exploration on what the effect of rarity is on the WTP for a QALY in society results in a higher WTP per QALY for rare diseases than for common and frequent occurring diseases. The influence of rarity on WTP may be further investigated in future research.

***Keywords*:** WTP, QALY, Social Perspective, rarity, rare diseases, cost-effectiveness, ICER threshold, decision-making, HTA

Table of contents

Contents

List of abbreviations

CV: Contingent Valuation

EE: Economic Evaluation

EQ-5D: EuroQol 5D (5 dimensions)

HTA: Health Technology Assessment

QALY: Quality Adjusted Life Year

QOL: Quality of Life

SP: Stated Preference

VAS: Visual Analogue Scale

WTP: Willingness to Pay

Chapter 1 Introduction

To support decision-making around the implementation of new interventions within healthcare, policy-makers often request economic evaluations (Ryen & Svensson, 2014). The main goal of the decision-making is typically to maximize the population's health from a certain budget. Economic evaluations enable researchers to gain insights in finding optimal allocations. These evaluations have therefore become important, because the health care budget is limited and consequently health care interventions cannot all be implemented. This ultimately comes down to deciding how and which resources should be allocated (Drummond, 2005).

In recent years, the most common-used economic evaluations have been cost-effectiveness analyses (CEA) and its sub-form, cost-utility analyses (CUA) (Bobinac, 2012). Often in CUAs, the Incremental Cost-Effectiveness Ratio (ICER) is calculated with Quality Adjusted Life Years (QALYs) as the primary outcome measure (Ryen, 2014). The advantage of QALYs is that the benefits are comparable amongst each other, and they combine length together with quality in one measure. The value of a QALY can lie between 'zero' (death) and 'one' (perfect health)[1]. As a result, it is easier to compare alternatives based on QALYs with different health effects rather than natural effects- as used in CEA, especially since the alternatives have the same outcome measure (Weinstein et al, 2009).

When using the QALY in economic evaluations, the ICER is calculated for comparison. This ICER represents the amount of costs per QALY gained of an intervention relative to an existing comparator. Interventions with the lowest cost per QALY under a certain value, which serves as a threshold, would be labelled as cost-effective (Evans et al., 2004). For CEA to be a proper decision-making tool, it has to have this threshold value for the systematic and worldwide criterium of decision making will drop otherwise and decisions will lack empirical foundation (Ryan et al., 2004; Ghandjour, 2020). Since its implementation, there has been much debate on the implicit and explicit cost-effectiveness threshold, but there is (still) no consensus on how high the threshold should be or what it should be based on (Claxton et al, 2010; Bennett et al, 2018). In short, the decision rule would look like the following: (cost differential / outcome differential) < threshold.

Seeing that there is no consensus, the threshold does not have one value: They differ per country and are based on different elements. For example, thresholds in Sweden and the Netherlands are based on severity but differ between €80,000 - €135,000 and €20,000 and €80,000 respectively (Svensson et al., 2015; Versteegh et al, 2019). England maintains a threshold of £20,000- £30,000 (Rawlins& Culyer, 2004). However, even though based on certain aspects of diseases, some find that these thresholds still lack empirical foundations (Cameron et al., 2018; Ghandjour, 2020), which is why it is needed to

[1] Negative QALYs also exist, indicating a health state which is worse than death, but this will not be taken into account in this thesis.

investigate the monetary value of QALYs (Ryan et al., 2004). Possible ways to set the threshold are exhausting a certain budget, or reviewing decisions on previous cost-effective interventions (Mccabe et al, 2008). Another frequently used method is the maximum willingness to pay (WTP) for a QALY (Robinson et al., 2013). WTP is often used in contingent valuation (CV) studies and presents respondents with hypothetical scenarios on a specific intervention (or even health care program) with the goal to evaluate this intervention. To give valuation to the intervention or program that is not on the market yet, a (representative) sample of society is asked what maximum it is willing to pay. Hence, respondents have to think about what they would be willing to sacrifice of their own income to have benefits of a health plan, if it were available (Drummond et al, 2005; Quevedo et al, 2009). This leads to a direct monetary valuation.

Different sorts of research on which phase of the disease people are in have been conducted, but been given different values (Ryen et al., 2014; Pennington et al., 2015; Eichler et al., 2004). Likewise, on WTP for the severity of the disease within the Netherlands specifically (Versteegh et al, 2019). However, it remains difficult to find one specific value of a QALY. There is evidence showing non-consistency in the WTP for a QALY, depending on size, duration but also type of health gain (Smith and Richardson, 2005; Appleby, 2007; Pinto Prades et al., 2009). This means the valuation of QALYs need to be taken into consideration. Whenever a disease has multiple aspects that a QALY is based on, or aspects that a threshold might not be based on, it becomes complicated due to the absence of specific guidelines (Mccabe et al., 2008). A category of diseases that can find hinder on the current valuation of QALYs and the threshold, is rare diseases[2]. In past research was concluded that no European consensus has been made on how interventions (Orphan Medicinal Products, otherwise known as OMP[3]) should be valued and how inequalities of access to these treatments (OMPs) have previously been observed. Ultimately, this leads to relatively scarce knowledge of the valuation of rare diseases and therefore potentially suboptimal allocations (Guttierez et al, 2015).

Extra difficulties for rare diseases come with the definition, differing per region: In the Netherlands, where this study will be focused on, a disease is called rare when affecting no more than 5 in 10,000 people (Mccabe et al., 2008). Amongst rare diseases are conditions that can be "*severe, chronic, progressive and also life-threatening, with multiple medical, psychological and social consequences*" (Nivel, 2003). Apart from their definitions, interventions for rare diseases could differ in quality in the clinical trials. This is mostly due to the small number of individuals suffering from these diseases, especially when compared to more common diseases. (Drummond et al., 2007; Richter et al., 2015). Consequently, in the Netherlands, the societal policy-making might have limitations in the validation. Still, no explicit research has been done on the societal WTP for rare diseases, while

[2] Richter et al. found in their study in 2015 that terms such as 'rare disease' and 'orphan disease' (or drugs) are used interchangeably, so that will happen as well in this thesis
[3] Orphan drugs

often a lot of (media) attention is drawn to negative reimbursement decisions. Ultimately, this attention implies there is a (social) opinion on the value of this type of care.

To summarize: it is difficult to find one specific value of a QALY and there is evidence showing non-consistency in the WTP for a QALY. Nonetheless, present WTP studies are attempting to make this more measurable. Especially within rare diseases there might be a need for outcomes of this topic. This is what makes these types of diseases interesting for this study. Not only because there has not been done a societal valuation research within the scope of WTP, but also because there is no consensus on how to value rarity. Because of ongoing innovations, what is considered as an orphan disease changes and therefore the amount of identified rare indications will increase substantially (Orphanet, 2019). It is also estimated this amount will rise to approximately 303 new diseases each year (Aitken & Kleinrock, 2017). Worldwide, it is estimated that rare diseases combined affect approximately 400 million people (WHO, 2021). Also, the number of orphan drugs is expected to increase in the future, especially considering the interest that has been awoken in the orphan drug market, mainly due to the Orphan Drug Act (Wellmann & Zhou, 2010; Joppi et al, 2013;)[4]. This is what makes it extra crucial to have more insights on the differences in valuation of rare and non-rare diseases.

This thesis can therefore contribute to the first findings of societal valuation of rarity of diseases within the Netherlands. For example, if (multiple) studies imply that people are willing to pay more for rare diseases, an adjustment in the policy can be made for rare diseases since it is currently only set for the severity of the disease. The social values may be particularly informative in the context of collectively funded healthcare systems, which is the way the Dutch healthcare is being funded.

The aim of this thesis is to find out what role the rarity of the disease plays within the willingness to pay. Hence, the research question will be as follows:

"How is the WTP for a healthcare intervention affected by the rarity of the disease that is treated?"

There is currently no consensus on valuation of interventions in rare diseases. Prices are higher, but are people willing to pay more for a disease that does not occur very often? Research into the influence of rarity of diseases on the WTP can give some first ideas on how consensus can be made. In this study, a questionnaire will be made and held amongst the general Dutch population to assess what rarity of a disease does to the WTP where different scenarios come to the surface. The WTP of rare, common and frequent diseases are compared to one another.

[4] The Orphan Drug Act, implemented in 1999 in the EU, is a law providing financial incentives for pharmaceutical companies to develop drugs for rare diseases. For precise information on what the incentives are, you can find more information on the website of the EMA (see references)

The construction of this Master Thesis is as follows: Chapter 2 provides a theoretical framework on previous research on WTP-studies, decision-making in the Netherlands and the effect of the use of different perspectives in these types of studies. Chapter 3 is a description of how data was collected and the analyses used to look at the given data. In chapter 4, the results of the WTP scenarios are given together with some sub-analyses and, finally, chapter 5 concludes with a discussion of the results and possible limitations. This chapter will be closed with recommendations for future research.

Chapter 2 Theoretical framework

2.1 Dutch decision making

As briefly touched upon in the introduction, current decision making in health care is often based on HTA, where economic evaluation is a part of. The difference in threshold values in the Netherlands is based on the severity. The severity lies between 0 and 1; these values are based on the proportion of normal Quality Adjusted Life Expectancy (QALE) lost due to the disease at hand (Van de Wetering et al., 2013). The first threshold, €20,000 per QALY is for a severity between 0.1 and 0.4; €50,000 for a severity between 0.41 and 0.7; €80,000 for a severity between 0.71 and 1 (Zwaap et al., 2015). In practice, these thresholds are more used as a reference guide for the discussion on cost-effectiveness, which means some drug will be implemented even though they are not considered cost-effective (Franken et al, 2015). By comparing the ICER of a treatment with the threshold it can be assessed whether the health gains of a new intervention are better than the effects of other interventions in the healthcare system, in order to compensate for these additional costs. However, in the decision making for cost-effectiveness, the impact of disease rarity on data uncertainty is not taken into consideration (Blonda et al., 2019). Multi-Criteria Decision Analysis (MCDA) has since been introduced. This incorporates a (flexible) set of value factors and involves perspectives of multiple stakeholders in valuation of the drug meaning multiple criteria are involved. Therefore, decision-making should be made 'easier' (Marsch et al., 2018). Nevertheless, Schey et al. assessed preferences for criteria in MCDA, where they found that respondents agreed on the importance of criteria captured in the 'regular' HTA. Moreover, they found that several disease-related and drug-related criteria should be included in the MCDA for the assessment of orphan drugs, since this is currently lacking. Other research shows that its successful implementation relies on decision-makers' openness toward transparency and a pragmatic approach, while allowing the flexibility for continuous improvement (Baran-Kooiker et al., 2018). This gives some criticism on the implementation. That is why there is still ongoing research on how to optimize the MCDA (Marsch et al., 2018). Since this study is only the first exploration what the general public is willing to pay on this subject, whatever this entails for cost-effectiveness and further steps are not taken into consideration.

Currently, standard economic evaluations treat the orphan diseases/drugs equal to non-orphan diseases/drugs. Although many countries across the world have implemented specific legislation to stimulate the development of orphan drugs, the standards for regulatory approval of orphan drugs are still the same as those for other drugs (Wästfelt et al., 2006).

2.2 Concept of rarity

The aim of this study is to find out if the monetary valuation of a QALY varies with the rarity of the disease that is treated. As mentioned earlier in the introduction, rarity does not have one definition, but it depends on jurisdictions whether or not a disease is called rare (Mccabe et al., 2005). In most European countries, a disease is rare when affecting less than 1 in 2.000 individuals (Medic et al., 2016). However, in the United States there have to be less than 200.000 individuals overall affected (Jayasundara, 2017). In this study, the current European guidelines for rarity will be implemented since the focus of the study is the Dutch population, where these guidelines are indicated. Rare diseases are often genetic, chronic and progressive (Nguengang Wakap et al., 2020). Within rarity are also diseases neglected by doctors taken into account, but the term in this thesis mostly focuses on diseases that affect small numbers of individuals, where patients have a severe condition (Aronson et al., 2006). One other factor that is of importance in this study concerning rarity, is the lack of treatments or even, whenever a treatment is available, alternatives for treatments (Medic et al., 2016). This is mostly due to the sample size of patients in studies, where the sizes are not big enough to actually effectively hold clinical trials (Richter et al., 2015). Besides this, because there is so little knowledge on the course and treatment effects of the natural development of the disease, identification and measurement of these treatment effects are also hampered (Philips, 2013; Wastfelt et al., 2006). Whenever treatments are available (which is increasing), the lives of thousands of patients are being improved. This ultimately means it is important to keep developing orphan drugs (Richter et al, 2015).

2.3 Previous literature on Willingness-to-pay

There are different ways to measure the Willingness-to-pay. Often used are a stated preference (SP) study or a revealed preference (RP) study (Ryen & Svenssion, 2014). Within SP studies, hypothetical scenarios together with (possibly hypothetical) alternatives are described, and these scenarios are flexible. It is also more based on what people say than what they would do in real choices, what is typical for RP. Within SP, contingent valuation (CV) is a frequent used method, where WTP-questionnaires are often implemented (Khan et al., 2014; Ginsburgh, 2017). A different option could be a discrete choice experiment. In this study the use will be made of WTP in a contingent valuation form especially since different literature also made use of this CV form (for example Gyrd-Hansen, 2003; Bobinac et al., 2010; King et al., 2005). As mentioned in the introduction, a hypothetical market is created through which the attitude of the respondent can be measured (Smith, 2003). Five basic forms of a contingent valuation can be distinguished: open-ended, bidding game, payment scales, discrete-choice experiment and a discrete-choice with follow-up (Klose 1998 & Smith 2000). In this study, the use of payment-scales and open-ended questions will be used, followed by an open-ended WTP (see Chapter 3 Methods). Since no explicit studies have been done on the influence of rarity on WTP, some

other known WTP-studies will be discussed below. Through the text is incorporated what that could entail for rare diseases, a hypothesis on the influence of rarity will be given at the end.

2.3.1 Perspectives

The used perspective in the study is of importance, since it could give different outcomes (Claxton et al, 2010). Three often used perspectives are the individual perspective, the societal perspective and the socially-inclusive perspective. The individual perspective shows how much an individual is willing to give up to improve their own health (Dolan et al., 2003). The societal perspective shows what an individual is willing to pay for the health of someone else- with the chance they affect health gain within the entire society as well-without any self-interest (Bobinac, 2012). Here, there is usually a collectively healthcare system funded by all members of society (such as the NHS) (Dolan et el., 2003). The socially inclusive individual perspective (SII) is basically a combination of individual's self-interest as well as the concern for others, so individuals may be affected themselves as well. This perspective fits for example well with the context of a social insurance system, based on solidarity (Bobinac, 2012). Social valuations may be lower or higher than individual valuations: lower because an individual may be willing to pay much more for themselves but not always for another. It could be higher, for instance, when equity is taken into consideration and society is willing to pay more for health improvements in certain groups (in this case people with a rare disease whether it be children or adults) than individual budgets of these patients would allow (Smith, 2005).

Furthermore, individual values are expected to differ much more that societal values for the QALYs: for example, in relation to end-of-life treatments (Olsen et al., 2004). Here, individual valuations may be high because of diminishing marginal utility of income. It could also be the case, on the basis of equity considerations, that society is willing to pay more for health improvements in certain groups (e.g., severely ill children) than individual budgets of these patients would allow (Bobinac, 2012). If this theory is applied to rare diseases, which are often severe and amongst children, the societal perspective would be preferable and therefore used in this study. Within healthcare however, most researchers have to date focused on finding WTP values from the individual perspective and not from a social perspective (King et al., 2005; Gyrd-Hansen et al., 2003; Bobinac et al., 2010). Because this study focuses on rare diseases and therefore not many people will actually endure the disease burden, it is crucial not to focus on potential treatment for respondents themselves, but on broader objectives (which will avoid influence of coping and adaptation in valuations) (Gyrd-Hansen, 2003; Olsen et al., 2004). This automatically excludes the individual perspective as the right perspective for this study. Since the Dutch insurance system is based on solidarity, you would expect the socially-inclusive perspective to be the right perspective. The downfall to this perspective would be that respondents will have to be part of the patient population. What is of importance in this study is the rarity, entailing as

mentioned above that the respondents will not actually endure the disease burden and thus not be part of the patient population. Therefore, a complete social perspective is used in this study.

Within these perspectives, there is also the option to choose between an ex-ante and an ex-post perspective. The difference is that in an ex-ante perspective, there is an uncertainty: the disease is not yet present and in ex post, the disease has already come to expression, where there is no component of risk anymore (Van der Star & Van den Berg, 2011).

2.3.2 Quality of Life and Life Extension

Mason et al. (2008) did an empirical review on early WTP studies, where they found that the type of QALY influences the Willingness-to-pay. For example, life-extending QALYs are valued higher than Quality-of-life (QoL) QALYs (Ryan et al, 2014). Often in rare diseases, there is no treatment available yet, which means the goal is to extend lives as much as possible rather than curing patients (NIVEL, 2003). If the above-mentioned statements are joined together, it would be likely that QALYs for rare diseases are higher valued than non-rare diseases, due to the lack of treatments. Severity also plays a positive role in the WTP (Versteegh et al., 2019; Svensson, 2015), meaning orphan drugs could be valued higher since they are more often more severe than non-rare diseases. Pinto-Prades and colleagues found a higher value of QALY gain for patients at the end-of-life than those of temporary health conditions, so again it can be concluded that the severity is involved in the valuation (2015). Also, Pinto-Prades found that QoL improving treatments in the end-of-life were valued higher than life-extending end-of-life treatments. So, a conclusion might be made that the consideration QoL or LE also depends on the phase in life that patients are in.

Jayasundara and colleagues did a study in 2017 concerning the cost-effectiveness in rare diseases (cancers). They state that the uncertainty around the cost-effectiveness of orphan drugs is larger than regular drugs, as is it harder to apply (standard) economic techniques. Because of a small patient population, costs made in the research and development phase need to be covered by a significant small group; especially compared to 'regular' drugs. This will result in higher costs per patient. Because some types of interventions are reimbursed, the expectation is that the willingness to pay is also higher, but that has not explicitly been researched. Medic et al (2016) have conducted a study in multiple countries (France, Germany, Italy, Norway, Spain, Sweden and United Kingdom), where their conclusion is that payers value rarity in pricing. However, since this study was about price setting, they would be a step further than the WTP measurement. But, if payers value rarity in pricing, the hypothesis could be made that the WTP for rarity would also be higher.

2.3.3 Equity

In his literary review in 2014, Koonal Shah found that the NHS has equity objectives, such as improving health for the poorest fastest, or for the sickest fastest. Even the sacrifice of health gains in order to achieve a more equitable health distribution is in order (Koonah Shah, 2014). However, equity

objectives are not only present in the NHS, they are present in each country that incorporates HTA. In this case, we can trace the subgroups back to the group that has a rare disease through characteristics of diseases.

Another study was done on valuing QALYs in relation to equity considerations (via a discrete choice experiment) in 2014 where was found that a considerable proportion of the public was willing to pay more per QALY for severe diseases. However, they found that the age of the patients had no impact on the WTP (Van de Wetering et al., 2014). This is controversial, since earlier research has shown that the ages of the patient is also of influence on equity basis, possibly even more than the influence of severity of disease (Dolan & Tsuchiya, 2005; Stolk et al, 2005; Olsen, 2013). A possible explanation could be that the weights based on absolute shortfall in different countries is not the same. For example: Norway implicitly prioritizes younger over older patients, and, conversely, the weights based on proportional shortfall in the Netherlands and end-of-life considerations in England may implicitly prioritize older over younger patients (Reckers-Droog, 2021). Since rare diseases are often linked to young patients, this could either take a positive turn for the WTP on rare diseases as well as negative, depending on the country. There is no consensus on what is socially acceptable on diverging from the QALY maximization rule and therefore important to keep researching this topic (2021). Some evidence of support is found for giving priority to the patient with shorter remaining life expectancy. Substantial preference for quality-of life improvement over life extension was observed, which would entail that the people would be willing to pay less for rare diseases (based on the QoL/ LE section) (Reckers-Droog, 2021). Because there is contradiction in the literature, the hypothesis for rare diseases could go either way and therefore, more research is needed.

Gyrd-Hansen and colleagues researched a social valuation of health changes in Denmark in 2002. In their study, the main conclusion was that the public has a strong inclination to give priority to those in a more severe health state provided their expected benefits are large enough to bring them to the health level where their rival patients are without treatment. In here, relieving patients of extreme problems are valued higher than relieving minor ailments (Gyrd-Hansen, 2002). As clarified with the definition of rarity, these diseases are often more severe than more common diseases as well as not treatable. This would entail that, if we would incorporate the literature into this study, treatments for patients with rare diseases would be valued higher than the more common diseases. However, he states that social values are not solely a function of the incremental changes in health and suggests that individual QALY gains cannot be directly interpreted as a measure of social value, so that has to be taken into account (Gyrd-Hansen, 2002).

2.3.4 Altruism

Following on the part of equity is altruism. This is especially important in this study, since it is highly unlikely respondents will develop a rare disease themselves. This due to the definition of rarity, which

includes genetics and the expression at a young age (Nguengang Wakap et al., 2020). Not many (empirical) studies have been done thus far concerning altruistic preferences concerning health care., but some WTP studies have been done with the societal perspective in general, which might be linked to altruism in health care.

Existing evidence has shown that other's people health does matter to individuals, which ultimately means that altruism exists in healthcare (Jacobsson et al., 2005). In Sweden- together with support of the population through contributions, campaigns and fundraisers- the cancer society finances 75% of Swedish cancer research project. Though cancer is a frequent occurring disease, lots of cancers are still unknown and therefore can also be rare. Jacobsson also states that the preferences for altruism are relatively higher for severe health states compared to preferences for 'selfishness'. This means that when the severity of the disease increased, the relative increase was also higher concerning attention for others than for themselves. Through this conclusion, a link could be made between altruism and equity. Jacobsson also advised more attention and resources towards severe health states compared to the milder ones, which would entail that towards rare diseases, especially because of the severity of these conditions.

As concluded in the previous section, severely ill people are favored to allocate resources for. If we apply this theory to rare diseases, these patients would be favored since rare diseases often have a high severity. However, if the comparator would be a drug for a frequent occurring severe disease, this could entail that the WTP for rare diseases would be lower. Therefore, it is of importance to look further than severity alone, due to the different expressions and situations severity can be in. Although the QALY estimate only incorporates the change in health status and the duration of the effect, there is evidence to suggest that initial health state prior to intervention has significant impact on the overall valuation of a health care program (Gyrd-Hansen, 2004). This would entail that people who are worse-off, which people with rare diseases mostly are, the impact of the valuation would be more significant. For example: Lin et al. (2017) used open-ended format to estimate to WTP for newborn screening test for spinal muscular atrophy which showed that people expressed a willingness to pay for spinal muscular atrophy screening even without an available therapy (median: $142; mean: $253). Willingness to pay however increased with treatment availability and respondent income (Frew et al, 2004; Lin et al, 2017).

In a different study, the difference between private WTP and altruistic WTP was measured. They found that the mean value for a private WTP was higher than the mean value for an altruistic WTP. However, in their study they also found that the more severe the disease was the higher the WTP became, even when the scenario was altruistic (Javan-Noughabi et al., 2017). So, it might be concluded that altruism is valued less than own health gain, but severity in combination with altruism will increase a WTP. They also researched what determinants were of influence on the altruistic WTP and tested that

the monthly income of respondents is of significant influence, that is in all different health states. Since they conducted their research in different health states, they also concluded sex is of influence in some health states (Javan-Noughabi et al., 2017).

2.3.5 Risk-aversion

A different subject that could have influence on the WTP is risk-aversion. If rarity goes up, the risk of people getting the disease will go down. Especially since patients mostly get rare diseases at a very young age, there is a relatively small chance of respondents getting this disease and therefore have little risk. So, risk-aversion and rarity are at complete opposite with each other: people want to pay more in order to get the treatment, but the risk is not there for themselves in this study. That is why the WTP could go down in this study. An example here is Covid-19. In 2020, when there was a break out through the entire world, there was little to no knowledge about this disease. Consequently, in a way you can say that when you are less risk averse, and if you know you are not much at risk for a disease, a lot of people are less altruistic (in this example due to violating the social distancing rule or stay-at-home-order) (Kluwe et al., 2021). You can also lead this back to rare diseases: if there are so few people having a rare disease, respondents might be less inclined to pay for this disease.

If a disease is frequent occurring, you see that people who are more risk averse, and people who perceive a greater health risk are generally more likely to be willing to pay (and/ or willing to pay more) for a given reduction in that risk (Hunter et al., 2012). Alternatively, if this is turned around and it is taken from the perspective of individuals who do not perceive the health risk, they probably will be less inclined to pay for reduction of the risk. Again, this would entail people being less inclined WTP for risk reduction in rare diseases.

2.3.6 Background variables

Aizuddin et al. did a literary review on factors that influence WTP and found that significant factors are age, education, income, household size and living in rural/ urban areas (2012). In this study, these background characteristics will also be taken into account. Different studies have done research on the WTP and in most of them, these background characteristics were taken into account (Bobinac, 2012; Van de Wetering, 2016). Xiong et al did research in China what factors positively influenced the WTP on ecological environment. In a way, this can be connected to altruism, since the effect of ecological environment goes beyond self-interest. The factors that significantly influence their WTP include the educational background, work type, residential location (2021). The past year, research has been done concerning the WTP for Covid vaccines (Harapan, 2020; Catma 2021, Qin, 2021; Cerda & Garcia, 2021), where they conclude that men have a higher WTP than women. This will be copied into this study.

2.3.7 Rarity

Based on the previous sections, a hypothetical theory can be made on what the influence of rarity on WTP would be. Out of these previous studies, together with some characteristics of rare diseases, it could be suggested that valuation of rare diseases would be higher than more common diseases, but this has never explicitly been researched. However, the weight people lie on valuing equity and/ or frequency, it could also be that people value rare diseases lower than more frequent occurring diseases. What is also researched is that the higher the price, the lower the WTP for the intervention (Wolff, 2020). Of course, the costs of orphan drugs are much higher than for non-orphan drugs due, amongst others, to the fact of low number of patients. Even though this study was done on shellfish quality and environmental conditions, a link could still be made to the rare diseases where people are not willing to pay for something that is extremely expensive (let alone that they do not have to use).

Study would be a first step in exploring the differences between the WTP of common diseases or the rarer diseases and if this hypothesis would be true. Below, a table with elements influencing the WTP is made together with what that would entail for the influence of the rarity on the WTP.

Factor	Influencing WTP	Influencing Rarity	Explanation
QoL / Life extension	+/ -	+/-	Since not many treatments for life improvement are available within rare diseases, the main goal is life extension. Since there is a greater WTP for life extension than QoL, this would mean rarity has a positive influence on the WTP.
Equity	+	+	Rare diseases are harder to treat and there is often no treatment (as said before). With the equity principle, people are more willing to invest in these diseases to give patients an equal chance.
Altruism	+	+	Respondents will not likely endure the disease burden in the future if not already, since rare diseases express at a young age. Altruism is present in health care, so respondents will be likely to pay for financing of treatment (if the treatment is present).
Risk-Aversion	+	-	Risk-averse people are more willing to pay. However, rare diseases are not frequent occurring so there is less risk. Therefore, people maybe not willing to pay for these diseases.

Table 1 Possible factors influencing WTP on rarity

The following hypothesis would be tested for this study:

H_0: The impact of rarity of diseases on the WTP = 0.

H_1: Rarity of diseases has an impact on the WTP.

Chapter 3 Methodology

This chapter describes the methods that are applied in this thesis. The main objective is to measure the differences in Willingness to Pay for a QALY in rarity of diseases. As mentioned in the theoretical framework, rarity is defined as: "*uncommon and serious conditions which are defined (…) as life-threatening or chronically debilitating conditions with a prevalence of no more than five in 10.000 people*". Furthermore, within rare diseases there is a high level of unmet need, due to no or a limited choice of therapeutic options (Medic et al, 2016). These aspects formed the basis of this research.

A contingent valuation (CV) amongst a sample of the Dutch population was performed to reveal the WTP for a QALY based on different scenarios with specific characteristics. In the scenario's, hypothetical changes in health are described as being the effect of healthcare interventions. The diseases in these scenarios varied in rarity and availability of other options, which will be elaborated later on. In the setting of a societal perspective, respondents were asked to state the maximum WTP for a hypothetical health gain that could accrue other people, thus with the presence of altruism since respondents are no part of the group of beneficiaries (Diener et al. 1998, Klose 1999, Smith 2003). Please note that, even though altruism or solidarity is present in this study, it is not the focus of this research and thus no further elaboration on the measure of solidarity will be given.

In the previous chapter, it became clear that the societal perspective was chosen for this study because of the nature of rarity: The individual perspective is not appropriate, since the respondents do not value the health gain for themselves. Within the societal perspective there are two options: socially inclusive perspective and a complete societal perspective. Since the respondents will not be part of the patient population, it might be useless to have any self interest in this study. This means the socially inclusive perspective is not of importance. That is why the completely societal perspective is most applicable to this theme.

The question(s) in a WTP study could either be framed from an *ex-ante* or an *ex-post* perspective. In the ex-ante perspective respondents are assumed not yet to have fallen ill, but might be at risk of ever needing healthcare in the future. In WTP studies carried out from the ex-post perspective on the other hand, respondents are asked to value health benefits derived from a treatment they will in fact utilize or have already utilized (Olsen & Smith, 1999; Dolan et al., 2003; Gyrd-Hansen, 2005). An advantage of the ex-ante perspective is that it could activate both option and externality values, meaning values from people who have little probability of using this type of health care in the future can be derived (Olsen et al, 2004). The ex-post perspective enables researchers to elicit the so-called use values (option values), from people who actually need the health care. Because the respondents are no part of the patient population, the need for externalities is present but for use values not so much, leaving the ex-ante perspective is the most appropriate perspective.

3.1 Design of the survey instrument & scenarios

In order to develop the WTP survey, a series of steps were undertaken. First, a literature review was held to determine the best ways to elicit WTP. The literature review on the questionnaire was done with multiple studies; Bobinac (2012); Reckers-Droog (2019) were commonly used, since these studies relate the most to the study at hand. These researchers also used the societal perspective in their study on WTP which meant there was some overlap. The questionnaire for this study contained five sections, to give a general overview of the substantive questionnaire, the following figure has been made:

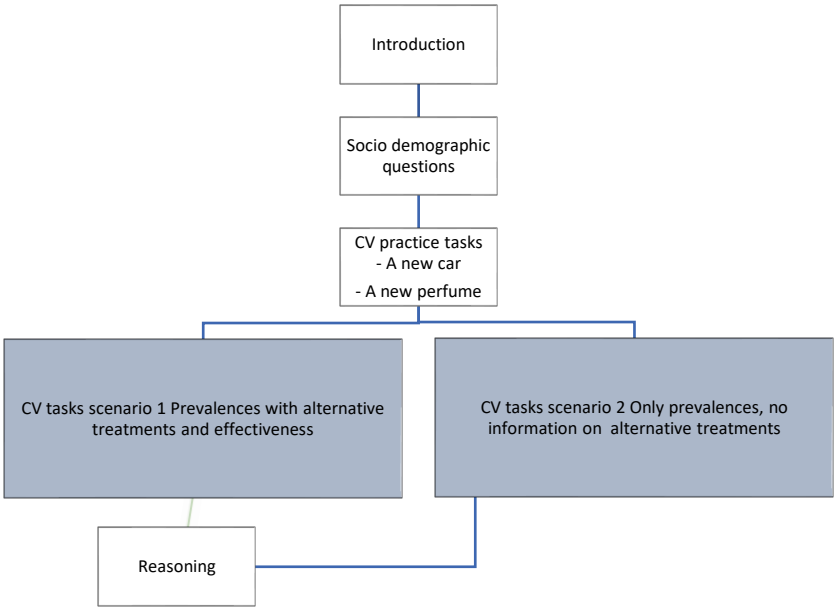

Figure 1: Lay-out of the questionnaire. Since the CV tasks are the core of the study, these are highlighted in a different color.

At the beginning of the survey, respondents were informed about the purpose and some background information was given on how decision-making in health care works concerning diseases of frequent occurrence as well as diseases with rare occurrence. After the introduction, respondents were asked about their demographic characteristics such as age, (household) income, where they live etc. Their current health was also measured through an EQ-5D-5L and by a self-perceived health through a VAS scale where 0 was 'death' and 100 'perfect health'.

The EQ-5D-5L was introduced in 2009 to enhance sensitivity and provide respondents with the opportunity for a more detailed and accurate picture of their health. The EQ-5D-5L uses the same five dimensions as the EQ-5D-3L but has two extra levels of severity in each dimension, which means it is elaborated. This results in more possible health states (Versteegh et al., 2016). The five levels are related to the dimensions ´mobility´, ´self-care´, ´usual activities´, ´pain/ discomfort´ and ´anxiety/depression´ and are 'Not /no problems', 'Slight problems', 'Moderate problems', 'Severe problems' and 'Unable to / extreme'.

Before the actual WTP study started, a short introduction was given to what was expected of the respondents followed by two CV warm-up questions in relation to non-health items to familiarize themselves with the concept of WTP as well as the concept of CV. These questions concerned a new car and a new perfume by own choice. After the introduction, when the real WTP study started, respondents had to fill in questions based on three scenarios. Because this study focuses on the effects of rarity on the WTP, all other circumstances had to be the same, meaning that the description of the diseases was the same. The health states that were given to the respondents were also based on the Dutch EQ-5D tariff (Versteegh et al., 2016). In each scenario, patient groups start out in the health state I- 43332- (0.524 QALY) and would go to the health state II -55233- (0.305 QALY) without treatment. The health states that are presented are purely hypothetical. Since focus in this study is rarity of the disease, this already implicates an often-bad condition to start with. Though rare diseases are progressive, it would seem more realistic if in a year, which this study is about, it has not yet progressed to a state where the QALY would be almost as low as 0, especially because the same health states need to be projected for a non-rare disease. Even though the health states are hypothetical, they are somewhat based on characteristics of an actual rare disease, Duchenne Muscular Dystrophy (RareDiseases.org, 2020). Appendix I gives the description of the scenario and the first question. The difference between the two health states is 0.219 QALY. This overall matches the differences in health gains in different studies (Bobinac et al., 2012; Lawerman-van de Wetering et al, 2016; Reckers- Droog et al., 2019).

In the scenario description- and in the questions-, it was always mentioned whether the disease was rare, relatively common or frequent occurring to make sure there was one line in defining rarity that would be clear to all respondents. Within the scenarios, different prevalences were given, of which none of the respondents would be part. Two health states were shown, one where a patient population was in and the other was the state they would go to if no treatment was given. Health deterioration could be avoided by taking a medicine without any side effects for a year, once a month. The use of an increased monthly basic health insurance premium was implemented to facilitate a payment vehicle. The basic health insurance is indeed how the Dutch residents contribute to the collectively funded health care system (Wammes, 2020).

The prevalences are 5 in 10,000 people in scenario 1; 100 in 10,000 people in scenario II and 1,000 in 10,000 people in scenario III. Scenario 1, rare, has been chosen with the definition of rarity in mind. The frequent occurring diseases, scenario III, was based off a frequent occurring disease. Influenza is a disease which is well known. This illness occurs frequently, it affects approximately 10% of the (world) population annually (Clayville,2015). Consequently, having the scenarios presented in these numbers might help the respondents getting an idea of the differences between the diseases and what it is they are paying for. Which leaves scenario II, which is presented in a ‘middle’ part.

In order to measure the influence of rarity in a different manner, two question formats were used to elicit two possibly different values of QALYs -in which the groups were randomly distributed: one group only had the specific characteristic of a rare disease (amongst others the availability of alternative treatments and how effective those are), which the other group did not have. In this way, the difference could be measured within the blocks where rarity is measured through prevalence, but also between the blocks where rarity is measured through alternative treatments. See the table below.

Group respondents (50% each)	*1st WTP question (rare)*	*2nd WTP question (common)*	*3rd WTP question (frequent)*
Situation 1	Prevalence = 5/10.000 No alternative treatments	Prevalence = 100/10.000 Some alternative treatments, though less effective	Prevalence = 1.000/10.000 Lots of alternative treatments, though less effective
Situation 2	Prevalence = 5/10.000	Prevalence = 100/10.000	Prevalence = 1.000/10.000

Table 2 Elaboration of the situations in the study presented to the respondents

The WTP for prevention of health deterioration was elicited by applying a two-step contingent-valuation procedure through payment scales (Donaldson et al., 1995; Donaldson et al., 1997a; Donaldson et al., 1997; Olsen and Donaldson, 1998; Gibb et al., 1998; Bobinac, 2012; Droog-Reckers, 2014). This was followed by a bounded open-ended question. In the first step, a payment scale was presented with a range from €0 to €24 in increase of basic health insurance premium per month for a year with unevenly distributed intervals within this range. This amount was chosen due to earlier research in which the payment vehicle was also an increase in basic health insurance premium (Droog-Reckers et al, 2014). See the appendix for an example question. Respondents were asked to indicate the increase in monthly health care premium they were certainly willing to pay for the duration of one year followed by the same payment scale asking them to indicate the amount they were certainly not willing to pay for a year. To remind them what the situation was, they were told after the amount '*for a disease that occurs within …… people*'. They were also always reminded they were not part of the population group, and that they had to keep their own income in mind (as a proxy for their ability to pay). Within payment scales, range bias can occur. This entails respondents might adjust their choices to the scale or categories from which they can choose, which results in higher choices than the respondents intend to (Bateman & Jones, 2003). The WTP will be higher when the number of categories increases (Whynes,

2004). To reduce range bias, every question in the sample had the same range of categories (van der Star & van den Berg, 2011).

In the second step, respondents were asked to indicate the maximum increase in monthly health-insurance premium they would be willing to pay within the payment range obtained in the first step. Besides their own monthly income, respondents also had to keep in mind that the stated WTP would be mandatory to pay for all adults living in the Netherlands. This maximum WTP was used for the calculation of the WTP for a QALY, since it is closest to their actual WTP. Again, to keep respondents from going back to the description of the scenario at hand, they were reminded in the questions what the main characteristics of that scenario were.

Soeteman et al (2017) did research on the impact of payment scales on the WTP for health gains and found that the two-step approach has several reasons for being more beneficial. With this approach, the WTP is more direct and precise than the range you present to the respondents. In this way, the respondent is less directed towards an answer which makes the respondents more thoughtful about their answer. Besides this, a two-step approach provides multiple valuations per respondent. This means more tests can be done on the mean WTP, but also the minimum WTP and maximum WTP in which differences between those 2 steps can be noticed. Ex post, the respondents were asked how certain they were if they would actually pay that amount of money in reality (if they were willing to pay more than €0). Via this question, the responses could be collected to reveal if their WTP was true (Smith, 2006; Blumenschein et al., 2001; Poe et al., 2002). Please note that this is never a robust answer of actually 'true', but the more certain respondents are of their WTP, the less likely it is for this amount to change (Smith, 2006). Respondents could choose out of options as to how certain they were about their stated WTP.

When respondents indicated they were not willing to pay more than €0 extra premium per month, a follow-up question was asked why. They could either give their reason by filling in an open-text or choose from several given options. The specific options are mentioned in the appendix. Through these answers, a distinction could be made between true zero WTP and protest zero WTP. Amongst the true zero valuations belong 'I can't afford to pay more than €0, Treatments are not worth more than €0 to me and I believe the treatment is worth more than €0 to me, but I would rather spend my money on something else'. 'I believe the basic health insurance premium is already too high', 'patients should pay for their own treatment' and 'For me, the value of health and health care cannot be expressed in monetary terms' are protest answers. (Bobinac et al., 2013; Reckers-Droog et al., 2013). For respondents with protest answers, it was important to look at what else they responded to the questions in order to decide if they can be left in de study or if they have to be excluded (Reckers-Droog et al., 2013).

After the CV tasks with scenarios were finished, respondents were asked if they could explain what the reason was that they would pay the same or a different amount for a disease occurring less frequent or more frequent. This was an open text question, to give the respondent the space to really tell what their reasoning was. Lastly, respondents were asked which factors they take into account in determining the amount of money they are willing to spend extra on their health insurance for a (rare) disease, if they think it is important that more attention is given to treatments for rare diseases and if so, why (see appendix). Finally, they were asked to which extent they are familiar with rare diseases. These questions were asked to potentially explain if results turn out as not expected through the given theory. Scenarios were also presented in a random order to reduce order bias as much as possible, however, it is not possible to completely eliminate order bias this way (Bateman and Jones, 2003).

A pilot was held amongst 22 people (11%). Overall, no big alternations had to be incorporated. In fact, respondents found the extra CV exercise to be helpful to clarify what was expected of them. The only aspect a couple of respondents were missing was some elaboration on certain concepts, such as 'intervention'. Therefore, concepts were assimilated in the text.

3.2 Analyses

3.2.1 Exclusion criteria

65 out of the 216 respondents did not complete the survey. Also, one respondent was above the age of 75 and since this study is focused on the Dutch general public between the age of 18-75, these respondents were excluded from the study.

Income

Income was asked in two ways: household income and personal income. Respondents who indicated their household consisted out of one person, did not get the question what the household income was, but only what their personal income was. Answering the personal income was optional, respondents had the option 'prefer not to say' due to privacy considerations. The data showed that the household income had more responses than the personal income, so the missing values from the household income were complemented with the personal income. In this way, the amount of missing values was reduced (and the personal income variable was taken out of the analysis). This still gave 7 missing values, but a sensitivity analysis showed that it had no difference on the outcome, so they were not excluded from the study. This leads to a total amount of 150 respondents.

3.2.2 Utilities

To see what the health state of the sample was, the measure of the EQ-5D was used instead of the VAS-scale. ED-5D tariffs are preferred, because they are derived using choice-based technique (Gyrd-Hansen, 2003; Bobinac et al., 2010). Also, the EQ-5D tariff is based on specific statements, which are

objective and the same for everyone, whereas the VAS-scale could be very subjective due to valuation of certain limitations one might have (Gyrd-Hansen, 2003). Therefore, in this study the EQ-5D is the base since these are also used in the calculations, where the VAS scale was used as a check.

The amount of QALYs gained (or not lost) due to the implementation of the (orphan) drug is calculated by subtracting health state II (55333) from health state I (43332).

Calculation QALY gains (prevention of loss) in the study

Below, table 3 gives a quick overview of the calculations of the QALY gains.

Scenario	*Calculation*	*# Patients*
- Rare	(5/10,000) * 17,475,415	8,737
- Common	(100/10,000) * 17,475,415	174,754
- Frequent	(1,000/10,000) * 17,475,415	1,747,541
Scenario	*Calculation*	*Total QALY gain*
- Rare	(0,524-0,305) * 8,737	1,913.56
- Common	(0,524-0,305) * 174,754	38,271.16
- Frequent	(0,524-0,305) * 1,747,541	382,711.59

Table 3 Calculations of total QALY gain

Since there are three scenarios in this study, all three have their own amount of total QALY gain depending on the occurrence of the disease. Therefore, the QALY gain was multiplied by the amount of people getting the disease. In the survey, the amount of people at risk of getting the disease is taken over the entire Dutch population and this was 17,475,415 people in January 2021 (CBS, 2021).

3.2.3 Calculation of WTP

Even though the amount of people getting the disease is given in table 3, in a societal perspective the payers are the amount of people paying for the increasing in basic health insurance. This means the amount of people paying for the treatment is not equal to the number of patients getting the treatment, but the total amount of adults (18 and over) living in the Netherlands. As mentioned in the Chapter before, this is due to the fact that rare diseases affect a small proportion of the population, while the entire population above 18 years old in the Netherlands is obligated to pay for health insurance. Using the study population in this way has also been done by Bobinac in 2010. CBS statistics showed the Dutch population consisted out of 14,190,874 people in January 2021 (2021). Calculation of the WTP is different in the societal perspective than in the individual perspective. For the calculations, the total amount of QALYs gained has to be known first. This amount is calculated, as seen above, amongst the patient population. In order to calculate the maximum WTP per QALY per year, the mean WTP had to be multiplied by twelve. This amount was subsequently multiplied by the total amount of adults (18 and over) living in the Netherlands. This results in the following formula:

Raw mean WTP	14,190,874 * (WTP *12)
WTP per QALY	Max WTP / total QALY gain

Table 4 Calculation of WTPs

The health gain in the scenarios was different per scenario. In order to calculate the WTP values per QALY, these different values are needed. To see what the differences are between the different versions, calculations were also done on the two groups separately. Two scenarios were created for the calculation of WTP before creating one value:

1. Scenario prevalences together with alternative treatments
2. Scenario only prevalences
3. (Scenarios combined together, mean WTP)

3.2.4 T-tests

The main question of this study what the influence of rarity is on the WTP. This means that a test is needed to find out if there are significant differences between the scenarios. The following hypothesis would have to be tested:

H_0: $\mu_1 = \mu_2 = \mu_3$, where μ_1 = *scenario rare*, μ_2 = *scenario common and* μ_3 = *scenario frequent*. This was tested with paired t-tests, since the study was amongst independent samples where different WTP scenarios were valued by the same respondent. This led to the following three hypotheses:

$H_0: \mu_1 = \mu_2$ $\quad$ $H_1: \mu_1 \neq \mu_2$

$H_0: \mu_2 = \mu_3$ $\quad$ $H_1: \mu_2 \neq \mu_3$

$H_0: \mu_1 = \mu_3$ $\quad$ $H_1: \mu_1 \neq \mu_3$

In this test, the raw WTP values are used the way Bobinac has done in her study (2010). It could be distorting to use WTP per QALYs in these tests, since the total amount of QALYs gained is different per scenario. Consequently, this would lead to outcomes with different proportions, which must be adjusted with more calculations. Using the raw WTP per month leads to the same proportions and more accurate results.

In order to execute the t-test, some assumptions need to be fulfilled. The first is that the outcomes have to have a continuous or ordinal scale (Stata, 2017), where the difference of the two logWTPs has to have a normal distribution. This was done through a Skewness/Kurtosis test for Normality, which showed that the difference between the dependent variables was not normally distributed. X^2 was 0.0000, which is not greater than 0.05, implying it is not significant at a 5% level. To overcome this, the dependent variables were transformed into log variables. Because the log of zero is impossible, 'one' was added to the value of the raw WTP variable. When this was done, the histogram

showed a normal distribution that was skewed to the left for the variables logWTP_1 and logWTP_2, as well as the difference between these two variables. The data also has to be random sample, and the final assumption is homoskedasticity of variance. At first, WTP scenario 1 (rare) and scenario 2 (common) have a similar variance (respectively 59.27 and 53.74), but the variance of scenario 3 (frequent) was much larger (218.51) but after the log transformation they showed homoskedasticity of variance. After transformations to fulfil the assumptions were done, the parametric t-tests were performed on the log-transformed data (Bobinac et al., 2012). Because the influence of rarity could go both ways (higher WTP or lower WTP), a double-sided p-value was maintained.

Not only was there difference in the questionnaire between the scenarios (different QALY gain per scenario), but there was also a different measure in rarity within the scenarios by including alternative treatment (equal QALY gains per scenario). Within scenarios, it was tested if respondents reported significant differences if they were given extra information on the rarity of the disease (for equal QALY gains). Because these two scenarios contain independent respondents, a two-sample t-test was held. Log WTPs are being used once more. Assumptions here (and that are met) are that the data has to be normal, and the variances have to be equal.

The following hypotheses are tested for each scenario (rare, common and frequent occurring):

H_0: $\mu_a = \mu_b$ H_1: $\mu_a \neq \mu_b$

Where μ_a is the version with alternative treatment given and μ_b is the version with only prevalences given.

3.2.5 Subgroup analysis

In this study, the mean WTP has been conducted. However, there are different income groups. Empirically, it shows that higher income groups are often willing to pay more than lower income groups (Georgiou et al., 2012). Therefore, this is also tested in the sub analysis. The income groups were recoded to 4 new groups:

- < €3000 per month
- €3000- €5000 per month
- €5000- €7500 per month
- > €7500 per month

These groups were based on CBS statistics, where most common income groups are shown (2021). Based on the theory, the expectation is that the higher income groups have a higher WTP. Hence, the null hypothesis that there is no difference in WTP between the different income groups will be tested through a one-way ANOVA test. A couple of assumptions also need to be tested: Apart from the dependent variable needing to be continuous, the independent variable has to be categorical. There also has to be independence of observation. There was no relationship between the observations in each

group or between the groups themselves, since the sample was random. It is also not allowed to have significant outliers, which was made sure of by looking at a box plot. The dependent variable should be (approximately) normally distributed. Lastly, there needs to be homogeneity of variances. This was tested with Bartlett's test for variances, where prob>chi2 = 0.605.

After executing the ANOVA test, pairwise comparisons of means were done if there was significance. This was done to see which groups are significant amongst each other.

3.2.6 Validity

To obtain reliable and valid data, the size of the sample has to be large and diverse enough. Internal validity represents the causality of a study and can be guaranteed by having enough measure results (Swanborn, 2006). Construct validity is also described as the measure of what was intended to measure. When performing a survey, the respondents should have a good understanding of the survey, or the measurement is incorrect the results will not reflect reality. The questionnaire had been tested through a pilot (>10%) and this showed that respondents knew what was expected of them and how they had to value the situations. Reliability is often realized by reproducing methods from previous studies. That is also what is done in this study, as told in the beginning of the methods.

3.2.7 Regressions

To investigate the theoretical validity of the dataset, multiple regression analyses were done with the maximum WTPs as the dependent variable. Regressions were executed to find out the effect of other variables instead of solely the version that was presented to the respondent. It is important that this is tested, since it could be that certain background characteristics may be of significant influence in one scenario (say rare), but not in the other. In order to run the regressions, some assumptions have to be tested. A first assumption in order to do a linear regression is the linearity of parameters. The dependent variable has to be related to the independent variables and the error, through the following formula:

$$y = \beta_0 + \beta_1 + u$$

The second assumption is that the residu has to have a normal distribution. This was done through a Skewness/Kurtosis test for Normality, and a Jarque-Bera normality test. Also, a histogram was graphed. When plotting the histogram, it was noticed that with the raw WTP value the residu was not normally distributed. However, when plotting the residu with the logWTP value, it did show a normal distribution. Therefore, the dependent variable was the logWTP. This means the log-linear regression has the following formula:

$$\text{Log}(y+1) = \beta_0 + \beta_1 + u$$

The data also has to be homoscedastic. To test for heteroskedasticity, the residuals were plotted by fitted value. Also, Cameron & Trivedi's decomposition of IM-test and Breusch-Pagan / Cook-Weisberg test for heteroskedasticity were performed (Stata, 2017; Bobinac, 2010; van de Wetering, 2016). These showed that the data is homoscedastic. The following assumption is non collinearity. Even though the regressions showed no multicollinearity between variables, it did show collinearity within some variables. For example, the variable 'region'. The collinearity was likely due to the fact that the different options of regions did not add a lot of information (most respondents came from the Randstad), therefore it was turned into a dummy variable for living in the Randstad. This was also done for education, which was recoded into 'low', 'middle' and 'high' (CBS,2021). In the variable income '€0-3,000' was the reference category, as was 'single' for marital status.

Model I has the background characteristics age, bruto house income, sex, education, region, marital status as independent variables. Income was taken into account even though the subgroup analysis was not significant, since income remains an important aspect within WTP studies (King et al, 2005). For the exact reason that income was not significant in the ANOVA, it is left out of the regression in the second model. At the same time, this model has some additional factors, including the effect of own health and knowing someone who has a severe disease. Since the theory stated that household size is also of significant influence, this is also tested in the regression. Model III exists of all variables together, including having children and working. These variables often lead to altruism, working means getting paid to provide for yourself/ your family. When you have children, you become responsible for them, which means you think of them first (Freund, 2014; Long & Krause, 2017).

3.3.8 Sensitivity analysis

To test the validity of the WTP values and how robust these are, sensitivity analyses were also run. For the main question of this study is whether there are differences between mean WTP of the different types of rarity of diseases, the sensitivity analyses were done with by t-tests. In the analyses the (protest) zero-valuations were excluded, even though they were seven percent of the sample, to see what would happen to the outcome. Outliers were also excluded. Insensitivity would mean that the WTP is not responsive to the outliers and non-outliers, as well as zero valuations.

All analyses were performed with Stata 16.0 for Windows

Chapter 4 Results

4.1 Data

216 respondents partook in this study. However, 65 responses were incomplete and 1 respondent was above the age of 75, which leads to the total of 150 responses. Table 5 shows the summary statistics. Ultimately, the sample would be a perfect reflection of the Dutch general public, though this was not possible due to time and financial restraints. As is shown in the table, the results indicate that this sample is representative of the Dutch general public in terms of age, sex shows there are more females than males though the percentages are somewhat higher than the statistics (CBS, 2021). The same goes for the household members. The average age of the respondent is 42.9 years old, with a fairly good health.

Table 5 summary statistics of the sample

N= 150	Mean	Std. Dev.	Min	Max	CBS statistics
Age	42.93	15.97317	21	74	42.22
Age group					
- < 30	51 (34.00%)				
- 31-45	37 (24.67%)				
- 46-65	43 (28.67%)				
- 66-75	19 (12.67%)				
Sex					
- Female	89 (59.33%)	.4928573			50.32%
- Men	61 (40.67%)				49.68%
Household	2.38	1.120188	1	6	2,14
Education					
- Lower school	2 (1.33%%)				7.35%
- Middle school	4 (2.67%)				18.28%
- Middelbaar beroepsonderwijs	9 (6.00%)				27.03%
- VWO	13 (8.67%)				9.94%
- HBO	54 (36.00%)				23.06%
- WO	68 (45.33%)				14.39%
Province					
- Friesland	1 (0.67%)				1.21%
- Overijssel	1 (0.67%)				3.43%
- Gelderland	2 (1.33%)				4.42%
- Utrecht	14 (9.33%)				8.14%
- Noord-Brabant	2 (1.33%)				9.36%
- Noord-Holland	18 (12.00%)				29.32%
- Zuid Holland	112 (74,67%)				37.61%
Region					
- Randstad	114 (6%)				
- West-Nederland	30 (20%)				75.57%
- Noord-Nederland	1 (0.67%)				4.55%
- Oost-Nederland	3 (2.00%)				8.08%
- Zuid-Nederland	2 (1.33%)				11.80%
Household income (N=131)					
- €0-500	7 (4.90%%)				

- €1,000-2,000	11 (7.69%)				
- €2,000-3,000	18 (12.59%)				
- €3,000-4,000	22 (15.38%)				
- €4,000-5,000	17 (11.89%)				
- €5,000-7,500	43 (30.07%)				
- €7500-10,000	12 (8.39%)				
- > €10,000	1 (9.09%)				
Marital status					
- Married	64 (42.67%)				39.47%
- Single	45 (30%)				48.95%
- Divorced	6 (4.00%)				7.88%
- Widow	1 (0.67%)				4.92%
- Cohabitant**	34 (22.67%)				unknown
Employment					
- Employed	119 (79.33%)				68,7%
- Voluntary work	3 (2.00%)				NA****
- (Early) Retirement	20 (13.33%)				18.39%
- Student	8 (5.33%)				13.01%
Children (N=122)					
- Yes	65 (53.28%)				32.5%***
- No	57 (46.72%)				67.5%
Know someone ill					
- Yes	39 (26.00%)				
- No	111 (74.00%)				
EQ-5D	0.917	0.0995482	0.492	1	N.A.
VAS scale health state to QALY	0.815	0.1249031	0.40	1	N.A.
Version					
- Prevalence alternative treatment	77 (48.67%)				
- Only prevalence	73 (51.33%)				

** cohabitant was implied when they had roommates, but it could also with a partner.
*** CBS only had statistics on whether children still lived at home
**** No information was available on volunteering on a 'full-time' basis

Even though the table shows the sample is representative for age and sex, it is not generalizable for, for example, where respondents live. Most of the respondents came from the province Zuid-Holland (South-Holland), where the rest of the Dutch population was underrepresented. This is most likely due to the residence of the researcher, who also studies in the province of South-Holland. Her contacts were also approached to participate in the study, which could explain the relatively high number of residents here. The same counts for the education of the sample, which is not representative for the Dutch population.

4.2 Practice WTP tasks

In the practice CV questions, it turned out that respondents were willing to pay €13,394 for a new car by choice and €52.68 for a new perfume. This indicates that the goal of the study was clear to the respondents, since these answers are reasonable. Please note that, even though these estimates imply that respondents understood what was expected of them, it is different to place a monetary value on health instead of non-health items due to, among others, the lack of market prices for health(care) (Ryan et al., 2004).

4.3 WTP values

There is a difference between the raw WTP and the WTP per QALY. First, the raw means will be displayed of what individuals are willing to pay for per month for the QALY gain of 0.219 for people having the disease.

Situation	**Group 1 (prevalence + alternatives)**	**Group 2 (prevalence)**	**Difference**	**Total (both groups together)**
Mean WTP **rare**	€ 8.45 (€101.40)	€8.16 (€97.92)	€0.29 (€3.48)	€8.30 (€99.60)
Mean WTP **common**	€8.42 (€101.04)	€9.03 (€108.36)	€0.61 (€7.32)	€ 8.74 (€104.88)
Mean WTP **frequent**	€9.49 (€113.88)	€13.10 (€157.20)	€3.61 (€43.32)	€11.34 €136.08

Table 6 Raw mean WTP for different scenarios in the study. Between brackets is the raw mean WTP per year.

T-tests

Paired t-tests on the raw mean WTP showed that there are differences between scenario rare and common, scenario rare and frequent occurring but also between scenario common and frequent occurring:

- The difference between the WTPs of rare diseases compared and common diseases is statistically significant on a 5% significance level (p=0.018%)
- The difference between the WTPs of rare diseases and frequent occurring diseases is statistically significant on a 5% level (p=0.001)
- The difference between the WTPs of common diseases compared to frequent occurring diseases is statistically significant on a 5% level (p=0.001).

Therefore, the null hypothesis that the means of scenarios rare, common and frequent occurring are equal can be rejected.

Apart from the differences between the scenarios, a two-sample t-test was done to see what the differences are between the raw WTPs within the versions. These tests showed that even though there are differences within the scenarios, they are not significant for each scenario:

- The difference between the WTPs of the versions with only prevalences and prevalences with alternative treatments for the scenario rare is not significant on any level (p=0.912).
- The difference between the WTPs of the versions with only prevalences and prevalences with alternative treatments for the scenario common was not significant on any level (p=0.303).
- The difference between the WTPs of the versions with only prevalences and prevalences with alternative treatments for the scenario frequent occurring was significant on a 10% significance level (p=0.06).

The null hypothesis that the mean of the version prevalences and alternative treatments is equal to the mean of the version with only prevalences can be rejected for scenario frequent occurring diseases. It cannot be rejected for scenario common and can also not be rejected for scenario rare.

Table 7 shows the average WTP per QALY for the different scenarios.

Situation	**Group 1 (prevalence + alternatives)**	**Group 2 (prevalence)**	**Total (both groups together)**
Mean WTP per QALY **rare**	€751,978	€719,051	€738,629
Mean WTP per QALY **common**	€37,465	€40,224	€38,889
Mean WTP per QALY **frequent**	€4,223	€5,929	€5,437

Table 7 Raw mean WTP per QALY per scenario.

When comparing the two groups, it becomes clear what the influence is of the additional information about the disease treatment given to the respondent. For example, for rare diseases the WTP is higher when stated that no alternative treatment is possible (+ €32,927). It also becomes clear that the WTP for more common diseases and frequent occurring diseases will be less when there are alternative treatments available (-€2,759 and -€824 respectively), even though the survey explicitly mentioned that the alternatives are less effective than the intervention.

Some respondents gave a zero WTP answer. The most zero valuation answers came from respondents that are against an increase in monthly health insurance. 2% indicated that drugs are worth more than €0 to them, but they would rather spend their money on something else and 2.5% indicates that the value of health and healthcare is not monetary for them. However, since the amount of total zero valuation answers is approximately 7% of all responses, and therein even less for the protest

answers, these observations will not be taken out of the main analyses. Nevertheless, to check for robustness, they are excluded from the sensitivity analyses.

4.4 Subgroup analysis

Table 9 shows what the difference is in raw WTP per month between the different income groups. In between brackets, the WTP per QALY is presented.

Income groups	Mean	Frequency	Percentage	Std. dev.
€0- €3,000	€7.24 (€643,952.10)	36	26.72%	6.69 (595,596)
€3,000- €5,000	€8.79 (€782,668.40)	39	27.48%	7.97 (702,851.50)
€5,000- €7,500	€8.93 (€756,116.90)	43	27.48%	8.03 (735,218.50)
> €7,500	€8.23 (€703,032.50)	25	18.32%	6.80 (609,824.20)

Table 9 Average WTP for different income groups WTP scenario rare.

Income groups	Mean	Frequency	Percentage	Std. dev.
€0- €3,000	€7.24 (€32,197.64)	36	26.72%	5.95 (29779.83)
€3,000- €5,000	€9.26 (€39,133.46)	39	27.48%	6.51 (35142.61)
€5,000- €7,500	€10.59 (€37,805.89)	43	27.48%	7.26 (36760.97)
> €7,500	€10.48 (€35,151.66)	25	18.32%	9.61 (30491.24)

Table 10 Average WTP for different income groups WTP scenario common.

Income groups	Mean	Frequency	Percentage	Std. dev.
€0- €3,000	€11.77 (€3,219.76)	36	26.72%	16.84 (15573.52)
€3,000- €5,000	€10.55 (€3,913.35)	39	27.48%	6.88 (3514.26)
€5,000- €7,500	€13.88 (€3,780.59)	43	27.48%	16.56 (3676.10)
> €7,500	€13.27 (€3,515.17)	25	18.32%	19.23 (3049.12)

Table 11 Average WTP for different income groups WTP scenario frequent occurring.

As shown in the table, and as expected from the theory, the higher the income group the higher the WTP. Income has a concave, if the income is that high, the growth of the increasing WTP stops, even though the group is still willing to pay a higher amount. However, none of the ANOVA tests conducted on the raw WTPs was significant, meaning there is no significant differences between the different income groups.

4.5 Regression

In table 12, the results of the multiple regressions are shown for the different scenarios. In model I, the results are for the rare diseases, model II for the common diseases and model III represents the frequent occurring diseases. In the appendix, regressions on the WTPs regarding primarily background characteristics can be found as well as models with interaction variables.

Table 12 Regression results raw log-WTP

DV = logWTP	Model I Rare			Model II Common			Model III Frequent occurring		
	Coef.	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*
Female (yes)	.0614	.1645	0.709	.1043	.1564	0.506	-.0185	.1782	0.917
Age	.0179	.0058	0.002*	.0118	.0065	0.070**	.0030	.0074	0.682
Randstad (yes)	.1372	1.467	0.351	.0869	.1702	0.610	.0169	.1939	0.930
Education (yes) High	2.050	1.321	0.123	.3206	.1513	0.036*	.3995	.1725	0.022*
Marital status (ref. = single)									
- Married	-.342	.2406	0.163	-.1460	.2199	0.508	-.1577	.2506	0.530
- Cohabitant	-.403	.2250	0.075**	-.2037	.2057	0.324	-.1511	.2344	0.520
- Widow	-.083	.9679	0.675	.2553	.8847	0.773	.3595	1.008	0.722
- Divorced	-1.291	.5742	0.026*	-.9216	.5249	0.081**	-.7399	.5982	0.218

Bruto household income (ref. = €0-€3,000)									
€3,000- €5,000	.0250	.2239	0.911	.1177	.2047	0.566	-.0108	.2333	0.963
€5,000- €7,500	-.2160	.2434	0.376	.0572	.2225	0.798	-.0349	.2535	0.890
> €7,500	-.0361	.2781	0.897	.2147	.2542	0.400	.0949	.2897	0.744
Working (yes)	.0667	.2208	0.763	-.2219	.2019	0.274	-.3341	.23006	0.149
Cons.	.7081	.6290	0.088**	.6674	.5912	0.261	1.2423	.6737	0.067**
Adj. R^2	0.0709			0.0472			0.0211		
Prob> F	0.0291*			0.0972**			0.0827**		
N	143			143			143		

As can be seen in the table, all models are statistically significant. Model for rare diseases at a 5% level, the models for common diseases and frequent occurring diseases at a 10% level. The model for rare diseases showed the highest adjusted R^2 (7.1% of variance), however in all models the explained variance is relatively low. Few variables are significant. As expected, age is positively associated with WTP and also significant at a 5% level. Living in the cities in the Randstad also significantly leads to a higher WTP. Since the subgroup analysis showed that income groups are not significantly related to the WTP, it is no surprise to see that there is no significance in the regression regarding this variable. Since the dependent variable is log transformed, the coefficients are calculated through the following formula: (exp (Coef.) -1) *100% (Ford, 2018).

The following results can be taken from the model:

For the scenario rare:

- Ceteris paribus, when respondents are a year older, their WTP for rare diseases increases with 1.8%. This is significant at a 5% level (p= 0.002)
- Ceteris paribus, respondents who are cohabitant compared to being single, their WTP for rare diseases decreases with 33.2%. This is significant at a 10% level (p=0.075).
- Ceteris paribus, respondents who are divorced compared to being single, their WTP for rare diseases decreases with 72.5%. This is significant at a 5% level (p=0.026).

For the scenario common:

- Ceteris paribus, when respondents are a year older, their WTP for common diseases increases with 1.19%. This is significant at a 10% level (p=0.070).

- Ceteris paribus, respondents who have a high education compared to other types of education, their WTP for common diseases increases with 37.80%. This is significant at a 5% level (p=0.0036).
- Ceteris paribus, respondents who are divorced compared to being single, their WTP for common diseases decreases with 60.2%. This is significant at a 10% level (p=0.081).

For the scenario frequent occurring:

- Ceteris paribus, when respondents have a high education compared to other types of education, their WTP for frequent occurring diseases increases with 49%. This is significant at a 5% level (p=0.022).

Explanations

What was also asked of the respondents is if they could explain why they would pay a certain amount in the rare situation and a different (or the same) amount in the situation with common or frequent occurring diseases. The frequent given answers are presented in the table below:

Question	Answers
Can you explain the difference (or equal amount) of what you are willing to pay for the scenarios rare and frequent occurring?	- Due to high prevalence - Choosing the most effective treatment - Rare diseases are too expensive - A different solution must be found than raising the health care premium - They are all important
Can you explain the difference (or equal amount) of what you are willing to pay for the scenarios rare and common?	- Equal treatment for every patient - Choosing the most effective treatment - They are all important
Can you explain the difference (or equal amount) of what you are willing to pay for the scenarios common and frequent occurring?	- They are all important - A different funding is needed than a premium increase if it involves this many people.
What do you take into consideration when deciding what you are WTP	- How long this disease will last - Depending on life expectancy of the patients - Solidarity

	- What the work load will be on health care personnel - The available alternatives - The lifestyle of the patient. - My own income - Chances that I get the disease

Table 13 Explanations given by respondents for given valuations.

Out of the survey, it seemed that respondents were willing to pay more for others, or were at least more willing to pay, whenever those people would be known to them (family/friends). Even though the common and frequent occurring diseases have alternatives that are less efficient, they do have alternatives. Since the rare scenarios did not have any alternative treatments available, respondents were willing to pay more in order to give every patient a chance.

What is surprising in the regression is that, for every scenario, people who are employed have a negative relationship with WTP for scenario common and frequent occurring. When looking at the given explanations, frequent occurring answers were 'more willing to help greater amount of people in need'. This seems contradictive.

4.6 Sensitivity analysis

After exclusion of the outliers and protest answers, the means of the raw WTPs are as presented in the following table:

Situation **N = 138**	**Group 1 (prevalence + alternatives)**	**Group 2 (prevalence)**	**Difference**	**Total (both groups together)**
Mean WTP **rare**	€ 9.01 (€18.02)	€8.61 (€103.32)	€0.40 (€4.80)	€8.79 (€105.48)
Mean WTP **common**	€9.27 (€111.24)	€9.46 (€113.52)	€0.19 (€2.28)	€ 9.37 (€112.44)
Mean WTP **frequent**	€10.53 (€126.36)	€13.73 (€164.76)	€3.20 (€38.40)	€12.22 (€146.64)

Table 14 Sensitivity analyses: raw mean WTP values after exclusion.

Not surprisingly, all the mean WTPs have increased. This is due to the fact that a mean will automatically go up if the zero valuations are excluded. Paired t-tests showed that the significance levels of the differences between all scenarios did not change the scenarios rare and common diseases did not change. The difference between the valuation of common diseases and rare diseases is statistically significant on a 5% significance level (p= 0.0065), as is the difference between the frequent occurring

and rare diseases (p= 0.0001) and between the frequent occurring and common diseases (p= 0.0001). Where the difference in mean between the versions of scenario rare became larger, the differences in scenarios common and frequent occurring diseases became smaller. What is also noticeable is that there is one difference in significance: the two-sample t-tests showed that the significance of the difference between the versions in scenario frequent occurring diseases has disappeared compared to the inclusion of all variables (p=0.1865). Scenarios rare and common diseases stayed insignificant (p= 0.9049 and 0.5947 respectively).

The WTP per QALY seems to be insensitive for the exclusion of the outliers and zero valuations. As shown in the table below, the WTP per QALY for the scenario frequent occurring diseases has stayed the same for both groups, even though the mean had increased a bit. The means for scenarios rare and common diseases have increased a bit, but it seems unlikely that the WTP per QALY is sensitive for exclusion of these values.

Situation (N=138)	**Group 1 (prevalence + alternatives)**	**Group 2 (prevalence)**	**Total (both groups together)**
Mean WTP per QALY **rare**	€ 801,813	€766,216	€782,768
Mean WTP per QALY **common**	€41,268	€42,073	€41,692
Mean WTP per QALY **frequent**	€4,686	€6,109	€5,437

Table 15 Sensitivity analyses: WTP per QALY after exclusion.

Chapter 5 Discussion

There is currently no consensus on how high a cost-effectiveness threshold should be. Surprisingly as well, no explicit research has been done on what exactly the influence of is rarity on society's WTP. This study was a first exploration on the influence of rarity on the WTP, leading to what the differences are of valuation of rare diseases, more common diseases and/or frequent occurring diseases.

The raw mean WTPs per month are €8.30 for rare diseases; €8.74 for common diseases and €11.34 for frequent occurring diseases. Calculated to WTP per QALY, this will result in €738,629 for rare diseases, €38,889 for common diseases and for frequent occurring diseases €5,437. The fact that respondents are willing to pay more for rare diseases is not strange, since the American Medical Journal (AMJ) reported in 2019 that societies are generally much more willing to pay for rare disease therapies. Especially rare diseases that otherwise do not have an available therapy (2019). The overall means of the WTP per QALY would result in €271,106.55. This amount is higher than current thresholds for the Netherlands, where thresholds differ between €20,000 and €80,000 (Versteegh et al.,2016). This would entail that the overall WTP is not cost-effective.

Looking at the raw WTP per month, on average, people were willing to pay more for frequent occurring diseases than rare diseases while the WTP per QALY is highest for rare diseases and lowest for frequent occurring diseases. This is most likely due to the amount of QALYs the maximum WTP had to be divided by. These were not the same for all scenarios, which is different from other studies, where often the total amount of QALYs to gain is equal: to achieve an equal QALY gain in different scenarios there is often an alternation of prevalence, where small health gains have a bigger prevalence than large health gains (Bobinac, 2012; Van de Wetering, 2016). Due to the methodology in this study, it was simply not an option to equalize the amount of QALYs, due to the fact that- besides the prevalence- everything else e.g., characteristics of the disease, beginning and ending health states, had to be the same. Nonetheless, the results do have an implication despite being an overestimation of the WTP per QALY. It becomes more difficult for respondents to make a calculation on what they are willing to pay. However, there is a difference in WTP per month, which is not that big, while the differences in WTP per QALY are substantial. Respondents recognize there is a difference between 5 out of 10,000 people becoming ill or 100 out of 10,000. The fact, they are willing to pay around the same amount per month could mean that (this characteristic of) rarity plays a role in their WTP. The results are robust since the sensitivity analyses did not give any substantial changes in the outcomes.

Limitations

This research was not without limitations, that need to be mentioned. First, it is hard to imagine health care as an asset. People can easily determine what they are willing to pay for an asset, but a treatment or other intervention is often beyond comprehension (Ryan et al., 2004). Also, in this study, the sample was relatively small. This automatically makes it more difficult for results to be significant or even generalizable, though the budget of a student is not as far reaching to ensure that the number of respondents is substantial. On the other hand, it is known in other studies that income and personal characteristics are not of significant influence, so this result is not entirely divergent (Ryan et al, 2004; King et al., 2005; Bobinac, 2012; Van de Wetering, 2016).

Secondly, the sample was not completely representative This makes it harder for conclusions to be made. Though the sample was representative for age and sex, it was not for, amongst others, education, employment and region of living. Especially region of living is an important factor due to differences between rural and urban areas (Asgary et al, 2004). Whenever there is only response from say an urban area, the results might be overestimated for the people living in rural areas, since they are often less willing to pay for healthcare premiums. In this study, main results came from respondents living in the Randstad, where the rest of the Netherlands was underrepresented. Often to get a more accurate result, more research is needed. However, Richiardi and colleagues state that it is not always needed to have a complete representative sample (2013).

What also could have been the case is that the respondents did not entirely get the goal of the questionnaire. In the study, practice questions were done in an individual perspective. Even though the pilot showed it was helpful for respondents, maybe it could have been better to use a societal perspective for these questions as well. It also seemed as if the respondents did not completely understand the goal, because of their explanations on what they take into consideration. A lot of the respondents mentioned lifestyle as an aspect they take into consideration for their WTP. This is surprising, since rare diseases are often genetic and thus the lifestyle has little influence on the manifestation of the disease (Boycott, 2013). This makes it questionable how much the public knows about rare diseases, the course of disease and available treatments. Particularly, because nothing was said about lifestyle in the survey. This could imply respondents took their own, different aspects into consideration for the valuation of the interventions.

There could also be some influence from the construct validity: in the survey, it was mentioned that their WTP for increase in basic health insurance premium was for one year. When this time period is extended, say every year from now on, respondents may be less willing to pay. This as a result of the fact that most people switch their annual health plan due to a lower price elsewhere, and because the total amount of monthly payments will go up (CBS, 2021). This may imply that the mean is overestimated. A different aspect is that the age of patients is not taken into consideration, which is a

relevant characteristic of rare diseases. Again, this could not be done due to the expected variation in scenarios which was attempted to avoid. However, this could also entail that the mean WTPs are underestimated, since earlier studies have shown that people are more willing to pay for younger patients (Droog-Reckers et al., 2013). The age of the beneficiary is an important equity consideration that ought to be included in the social valuations of publicly provided healthcare services. Equity is a difficult aspect when researching the differences between rare diseases and non-rare diseases. Often, characteristics are completely different, especially the age when the disease manifests. In rare diseases, most cases happen among (very) young people, and they are mostly genetic. Even though non-rare diseases can also be genetic, they could come to light at a later age. Because the goal was to look at the influence of rarity, every other aspect had to be the same.

Another methodological limitation in this study is the option of hypothetical bias. It could be that hypothetical situations are valued lower than when there is an actual situation presented (Baker et al., 2014). Enabling uncertainty, which is often done in an ex-ante perspective, will likely lead to higher WTP estimates than taking the ex-post perspective, which was also researched by Dolan et al (2003). Nonetheless, in this study the prevalence of the diseases are given to the respondents and respondents themselves are not part of the patient population, which would reduce the hypothetical bias in itself. Lastly, the used payment scale could also be a limitation. Even though an interval scale was used with a realistic range, the scale could still have influenced the decision of the respondent. This could be enlarged when people are unsure of their answers, meaning guiding the respondent towards a specific direction (Reckers-Droog et al., 2014).

In societal decision-making, solidarity and altruism play a significant role in determining the WTP of health (Olsen et al, 2004; Jacobsson et al, 2005; Smith, 2006). This is also why the theoretical framework discusses that possible influence. Since they play a significant role in WTP, it may mean that these factors are more important than other factors. However, this is something that could not be measured in this study, since there were no questions regarding the survey involving ‘How important do you find solidarity?’ etc. So, no weights could be placed on certain aspects.

Dolan et al. did a review on the perspectives of WTP studies (2003). The use of the societal perspective in this thesis could be questioned by some, especially since the use of the socially-inclusive-individual perspective seems more in tune with the Dutch healthcare system. However, the definition of social value of a QALY was here interpreted as the amount of own consumption respondents are willing to pay for a health gain in society, by increasing the health care premium. It could be, of course, that there are different definitions for a social value (Van de Watering, 2016). Therefore, the need seems likely for further research in this area.

Still, even though the socially inclusive perspective seems more in tune with the Dutch healthcare system, the complete social perspective seems the best possible perspective in this study. Self-interest

is not applicable, as it is ruled out to get an isolated effect of rarity. There would also be a difference in valuation if the perspective was ex-post. People would probably not be willing to pay, or at least less than they would now, if the disease was already manifested. Uncertainty could lead to higher valuations, so that is something to also have in mind (Okada, 2010). This also leads back to the theoretical framework, where risk-aversion is taken into consideration.

Recommendations for future research

The goal of this study was to research whether QALYs gained in different rarity categories of diseases are valued differently. In this research, 3 scenarios and 2 versions were used. Per scenario, which only differed in prevalence, one group of respondents had prevalence and available alternatives; the other group only saw prevalence of diseases. Rarity can be measured in other ways, for example with certain characteristics of (rare) diseases (life expectancy, age of patients, etc.). In this way, the valuation is of different rarity characteristics in the diseases can be measured as well. The same goes for ultra-rare diseases, which is a subcategory within rare diseases. What also may be useful for further research is if respondents are given more information on the course of disease in general (so that discounting can be applicable as well). This way, proportional shortfall weights could be considered.

This study also only had the transformation of health state I to health state II if patients did not receive treatment. In future research and in further stages, it could be of value if there is variation in the beginning health states to see if the state that patients start out in also has influence on the respondents' WTP. Subsequently, more common diseases often progress less, so the second health state might be of different values as well (Bobinac, 2010). Of course, respondents have a higher risk at getting a non-rare disease, so this could also lead to a different perspective in the study design. It could be that, for some questions, self-interest might be taken into account leading to a socially-inclusive perspective.

Also, the use of equal total QALY gains could lead to more proportional outcomes, which is often the case in other studies (Bobinac, 2010; Reckers-Droog, 2014; etc.) For that reason, a different design might be incorporated in research. Here, the focus was on the differences between rare and non-rare diseases, but in the future the use of regression analyses could be the focus point and further elaborated. In that way, the emphasis is more on what influences these different valuations, which fell out of the scope of this study.

What is also not considered in this study is what further policy implications could be when the WTP turns out to be not cost-effective. For now, more research needs to be done on the differences in valuation of rare diseases and non-rare diseases before something more can be said about future policy. Even though research in the Netherlands is focused on aspects beyond the QALY (MCDA), literature so far has shown it needs further optimalization (Baran-Kooiker et al.,2018; Lasalvia et al.,2019). Still, it is important for future research to be based on these multiple aspects. A dynamic threshold for cost-

effectiveness might be a step in the right direction, where different QALYs are valued in different ways. Therefore, future research should be focused on this area.

Conclusion

The main question of this thesis was: "How is the WTP for a healthcare intervention affected by the rarity of the disease that is treated?"

There is a significant difference between the valuation of the mean WTP for rare, common and frequent occurring diseases. On average, respondents are willing to pay more premium per month for frequent occurring and common diseases than for rare diseases. However, the WTP for rare diseases showed a higher mean and the WTP for common and frequent occurring diseases showed a lower mean when information on the availability of alternative treatments was presented. This is probably because there are hardly any treatments available for rare diseases, while there are treatments available for the non-rare diseases. This might implicate that respondents value alternatives higher than effectiveness. When the WTP is calculated per QALY, respondents are willing to pay more for rare diseases than for common and frequent occurring diseases, where results are robust.

Respondents indicated they find it important to have information on the prevalence and the efficacy of the intervention that is being analyzed. Also, through the explanations of the respondents it might be concluded that more factors need to be considered in the survey. As the search towards a more optimal valuation of a QALY continues, future research might be of contribution on the different weights respondents place on certain aspects of rare diseases with their different characteristics.

Reference list

Aizuddin, N. et al. BMC Public Health 2012, 12(Suppl 2): A37 http://www.biomedcentral.com/1471-2458/12/S2/A37

Appleby, J., Devlin, N. & Parkin, D. (2007). *NICE's cost effectiveness threshold*. BMJ 2007; 335:358–9

Asgary, A., Willis, K., Taghvaei, A. A., & Rafeian, M. (2004). Estimating Rural Households' Willingness to Pay for Health Insurance. *The European Journal of Health Economics, 5*(3), 209–215. http://www.jstor.org/stable/3569979

Baran-Kooiker A, Czech M, Kooiker C. Multi-Criteria Decision Analysis (MCDA) Models in Health Technology Assessment of Orphan Drugs-a Systematic Literature Review. Next Steps in Methodology Development? Front Public Health. 2018 Oct 15; 6:287. doi: 10.3389/fpubh.2018.00287. PMID: 30374435; PMCID: PMC6197072.

Bleichrodt, H., Diecidue, E. & Quiggin J. (2004) Equity weights in the allocation of healthcare: The rank-dependent QALY model. J Health Econ. 2004;23(1):157–171

Blonda, A., Huys, I., Denier, Y., Triest, J., Kiani, P., & Simoens, S. (2019). Pro97 Value Assessment Frameworks for Orphan Drugs: a comparative analysis across European jurisdictions. *Value in Health, 22*, S858–S859. https://doi.org/10.1016/j.jval.2019.09.2427

Bobinac, A., van de Wetering, L., Van Exel, N. J. A. & Brouwer, W. B. F. (2016) Equity-dependent social willingness to pay for a QALY. Valuing and Refining Outcome Measures for Economic Evaluations in Health Care [dissertation]. Rotterdam: Erasmus University Rotterdam;

Bobinac, A., van Exel, N. J. A., Rutten, F. F. H. & Brouwer, W. B.F. (2012) Inquiry into the relationship between equity weights and the value of the QALY. Value Health. 2012;15(8):1119–1126.

Bobinac A, van de Wetering, E. J., Van Exel, N.J.A. & Brouwer W.B.F. (2015). Equity-dependent social willingness to pay for a QALY.

Bobinac, A. (2019) Mitigating hypothetical bias in willingness to pay studies: postestimation uncertainty and anchoring on irrelevant information. Eur J Heal Econ. 2019;20(1):75–82.

Botelho A, Pinto LC. Hypothetical, real, and predicted real willingness to pay in open-ended surveys: experimental results. Appl Econ Lett. 2002;9(15):993–6.

Boycott, K., Vanstone, M., Bulman, D. (2013). Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat Rev Genet* **14,** 681–691 (2013). https://doi.org/10.1038/nrg3555

CBS. (2020) Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83005ned /table?fr omstatweb

CBS. (2020) Available from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide

CBS. (2021) Available from: https://www.cbs.nl/nl-nl/visualisaties/inkomensverdeling

CBS. (2021) available from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461bev/table?ts=1628195467736

Claxton, K., Walker, S., Palmer, S. & Sculpher, M. (2010) Appropriate perspectives for health care decisions. CHE Research Paper54, Centre for Health Economics, University of York.

Clayville L. R. (2011). Influenza update: a review of currently available vaccines. *P & T : a peer-reviewed journal for formulary management*, *36*(10), 659–684.

Coast, J. (2004). Is economic evaluation in touch with society's health values? BMJ (Clinical research ed.), 329(7476), 1233-1236.

Collins, M. & Latimer, N. (2013) NICE's end of life decision making scheme: impact on population health. BMJ. 2013;346(7905):1–5.

Denis, A., Mergaert, L. & Fostier C. (2010) A comparative study of European rare disease and orphan drug markets. Health Policy. 2010; 97:173–179.

Department of Health. Tackling health inequalities: a programme for action. London: DH; 2003.

Dolan, P., Edlin, R. & Tsuchiya, A. (2008) The relative societal value of health gains to different beneficiaries. HEDS Discussion Paper 08/12

Dolan P., Olsen, J. A., Menzel, P. & Richardson, J. (2003) An inquiry into the different perspectives that can be used when eliciting preferences in health. Health Econ. 2003;12(7):545–551

Dolan P, Tsuchiya A. (2005) Health priorities and public preferences: The relative importance of past health experience and future health prospects. J Health Econ. 2005;24(4):703–714. 40.

Draborg, E., Gyrd-Hansen, D., Bo Poulsen, P. & Horder, M. (2005) International comparison of the definition and the practical application of health technology assessment. *Int J Technol Assess HealthCare.* 21:89-95.

Drummond, M. F. (2008) Challenges in the economic evaluation of orphan drugs. Eurohealth. 14(2):16–7.

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G.W.: Methods for the economic evaluation of health care programmes. Oxford University Press, Oxford (2015)

Ford, C. (2018). *Interpreting Log Transformations in a Linear Model*. University of Virginia Library Research Data Services + Sciences. From: https://data.library.virginia.edu/interpreting-log-transformations-in-a-linear-model/

Franken, M., Le Polain, M., Cleemput, I. & Koopmanschap, M. (2012) Similarities and differences between five European drug reimbursement systems. *Int J Technol Assess Health Care.* 28:349-357.

Franken, M., Stolk, E., Scharringhausen, T., de Boer, A. & Koopmanschap, M. (2015) A comparative study of the role of disease severity in drug reimbursement decision making in four European countries. *Health Policy.* 119:195-202.

Freund AM, Blanchard-Fields F. Age-related differences in altruism across adulthood: making personal financial gain versus contributing to the public good. Dev Psychol. 2014 Apr;50(4):1125-36. doi: 10.1037/a0034491. Epub 2013 Sep 23. PMID: 24059256.

Frew, E. J., Wolstenholme, J. L. & Whynes, D. K. (2004) Comparing willingness-to-pay: bidding game format versus open-ended and payment scale formats. Health Policy. 2004;68(3):289–98.

Ginsburgh, V. (2017). Contingent Valuation, Willingness to Pay, and Willingness to Accept. 10.1007/978-3-319-47458-8_26.

Green, C., & Gerard, K. (2009). Exploring the social value of health-care interventions: a stated preference discrete choice experiment. Health Economics, 18(8), 951-976.

Hackenberger B. K. (2019). Rare, rarer, it still has not happened. *Croatian medical journal*, *60*(6), 565–569. https://doi.org/10.3325/cmj.2019.60.565

Hagemans, M. L. C., Winkel, L. P. F, Hop, W. C. J., Reuser, A. J. J., Van Doorn, P. A. & Van der Ploeg, A.T. (2005) Disease severity in children and adults with Pompe disease related to age and disease duration. *Neurology.* 64:2139-2141.

Hunter, P.D., Hanley, N., Czajkowski, M., Mearns, K., Tyler, A.N., Carvalho, L., Codd, G.A.: The effect of risk perception on public preferences and willingness to pay for reductions in the health risks posed by toxic cyanobacterial blooms. Sci. Total Environ. 426, 32–44 (2012)

Jacobson, F., Carstensen, J. & Borgquist L. (2005). Caring externalities in health economic evaluation: how are they related to severity of illness? Health Policy 73 (2005): 172–182

Jayasundara, K., Krahn, M., Mamdani, M., Hoch, J. S. & Grootendorst P. (2017) Differences in Incremental Cost-Effectiveness Ratios for Common Versus Rare Conditions: A Case from Oncology. Pharmacoecon Open. 2017 Sep;1(3):167-173. doi: 10.1007/s41669-017-0022-7.

Joppi, R., Bertele, V. & Garattini, S. (2013) Orphan drugs, orphan diseases. The first decade of orphan drug legislation in the EU. *Eur J Clin Pharmacol*. 69:1009-1024.

Kanters, T. A., Hollak, C.E.M., Van der Ploeg, A. T., Rutten-van Mölken, M. P. M. H. & Hakkaart, L. Evaluation of orphan drugs in the Netherlands: can HTA offer guidance to healthcare policy?

Kanters, T.A., Hakkaart, L., Rutten-van Mölken, M. P. M. H. & Redekop, W. K. (2015) Access to orphan drugs in western Europe: can more systematic policymaking really help to avoid different decisions about the same drug? Expert Rev Pharmacoecon Outcomes Res 15(4):557-559

King, J.T., Tsevat, J., Lave, J.R. & Roberts, M. S. (2005) Willingness to pay for a quality-adjusted life year: implications for societal health care resource allocation. Medical Desicion Making 2005; 25:667–77

Kluwe-Schiavon, B., Viola, T. W., Bandinelli, L. P., Castro, S. C. C., Kristensen, C. H., Costa da Costa, J.,et al. (2021) A behavioral economic risk aversion experiment in the context of the COVID-19 pandemic. 16(1): e0245261 https://doi.org/10.1371/journal.pone.0245261

Lancsar E, Gu Y, Gyrd-Hansen D, et al. (2020) The relative value of different QALY types. J Health Econ. 2020; 70:102303.

Lasalvia, P., Prieto-Pinto, L., Moreno, M., Castrillón, J., Romano, G., GarzónOrjuela, N. & D Rosselli (2019) International experiences in multicriteria decision analysis (MCDA) for evaluating orphan drugs: a scoping review, Expert Review of Pharmacoeconomics & Outcomes Research,19:4, 409-420, DOI: 10.1080/14737167.2019.1633918

List, J.A. & Gallet, C.A. (2001) What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?, *Environmental & Resource Economics*, **20**, (3), 241-254

Lin, P.J., Yeh W-S, Neumann PJ. Willingness to pay for a newborn screening test for spinal muscular atrophy. Pediatr Neurol. 2017; 66:69–75

Long, M.C. & Krause, E. (2017) Altruism by age and social proximity. PLoS ONE 12(8): e0180411. https://doi.org/10.1371/journal. pone.0180411

Malinowski, K. P., Kawalec, P., Tra˛bka, W., Czech, M., Petrova, G., Manova, M., et al. (2019). Reimbursement legislations and decision making for orphan drugs in central and eastern European countries. Front. Pharmacol. 10, 487. doi:10. 3389/fphar.2019.00487

McCabe C, Tsuchiya A, Claxton K, Raftery J. (2007) Assessing the economic challenges posed by orphan drugs: a comment on Drummond et al. Int J Technol Assess Health Care. 23(3):397-401; author reply 401-4. doi: 10.1017/s0266462307071012. PMID:c17579945.

Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., et al. (2020). Estimating cumulative point prevalence of rare diseases: analysis of the orphanet database. *Eur. J. Hum. Genet.* 28 (2), 165–173. doi:10.1038/s41431-019-0508-0

NICE. Social value judgements: principles for the development of NICE guidance. second ed; 2008, http://www.nice.org.uk/media/ C18/30/SVJ2PUBLICATION2008.pdf

Nivel (2003) from: https://www.nivel.nl/sites/default/files/bestanden/rare-diseases-bibliography.pdf

Ollendorf, D. A., Chapman, R. H., and Pearson, S. D. (2018). Evaluating and valuing drugs for rare conditions: No easy answers. Value Health 21 (5), 547–552. doi:10.1016/j.jval.2018.01.008

Olsen, J.A. (2013) Priority preferences: "end of life" does not matter, but total life does. Value Health. 2013;16(6):1063–1066.

Oostenbrink, J. B., Bouwmans, C. A. M., Koopmanschap, M. A. & Rutten, F. F. H. (2004) *Manual for costing studies: Methods and standard unit costs for economic evaluations in health care [in Dutch: Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg]*. Amstelveen, the Netherlands: College voor zorgverzekeringen.

Pennington, M., Baker, R., Brouwer, W., Mason, H., Hansen, D. G., Robinson, A., et al. (2015). Comparing WTP values of different types of QALY gain elicited from the general public. Health Economics, 24(3), 280-293.

Phillips, M. I. (2013) Big pharma's new model in orphan drugs and rare diseases. *Exp Opin Orphan Drugs.* 1:1-3.

Rankin, J. & Robinson A. (2018) Accounting for protest zeros in contingent valuation: a review of literature. HEG Work Paper; 2018, 18–01.

Reckers-Droog, V. T., van Exel, N. J. A. & Brouwer, W. B. F. (2019) Equity weights for priority setting in healthcare: severity, age, or both? Value Health. 2019;22(12):1441–1449.

Richardson J., Iezzi A., Maxwell A., Chen G. (2018) Does the use of the proportional shortfall help align the prioritisation of health services with public preferences? Eur J Heal Econ. 2018;19(6):797–806. 4

Rowen, D., Brazier, J., Mukuria, C., Keetharuth, A., Hole, A. R., Tsuchiya, A., et al. (2014). Update: Eliciting societal preferences for weighting QALYs according to burden of illness, size of gain and end of life.Policy Research Unit in Economic Evaluation of Health & Care Interventions (EEPRU) Research Report.

Ryan M. (2004) Discrete choice experiments in Health care. BMJ. 2004;328(7436):360-1. DOI: 10.1136/bmj.328.7436.360

Soares Santos, A., Guerra-Junior, A. A., Godman, B., Morton, A. & Ruas, C. M. (2018) Cost-effectiveness thresholds: methods for setting and examples from around the world, Expert Review of Pharmacoeconomics & Outcomes Research, 18:3, 277-288, DOI: 10.1080/14737167.2018.1443810

Schey, C., Postma, M. J., Krabbe, P., Topachevskyi, O., Volovyk, A., & Connolly, M. (2020). Assessing the Preferences for Criteria in Multi-Criteria Decision Analysis in Treatments for Rare Diseases. *Frontiers in public health*, *8*, 162. https://doi.org/10.3389/fpubh.2020.00162

Schwappach, D. L. B. (2002) Resource allocation, social values and the QALY: a review of the debate and empirical evidence. Health Expectations 2002; 5:210–22.

Shiroiwa, T., Sung, Y., Fukuda, T., Lang, H., Bae, S., & Tsutani, K. (2010). International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ, 19, 422-437.

Shah, K. K., Tsuchiya, A. & Wailoo, A. J. (2014) Valuing health at the end of life: an empirical study of public preferences. Eur J Heal Econ. 2014;15(4): 389–399.

Shi L, Gao Z, Chen X. The cross-price effect on willingness-to-pay estimates in open-ended contingent valuation. Food Policy. 2014; 46:13–21.

Soeteman, L., van Exel, J. & Bobinac, A. (2017) The impact of the design of payment scales on the willingness to pay for health gains. Eur J Heal Econ. 2017;18(6):743–760.

Sousa N, Costa T, Monteiro-Soares M, Rocha Gonçalves F, Azevedo LF. Bias in valuation of health care benefits in metastatic prostate cancer: a contingent valuation of willingness to pay. ASCO Annual Meeting. 2017.

Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: an empirical inquiry into social value. Health Policy. 2005;74(3):343–355. 41.

Wastfelt, M., Fadeel, B. & Henter, J. (2006) A journey of hope: Lessons learned from studies on rare diseases and orphan drugs. *J Intern Med.* 260:1-10.

Wellman-Labadie O, Zhou Y. (2010) The US orphan drug act: Rare disease research stimulator or commercial opportunity? *Health Policy.* 95:216-228.

Wolff, E., Larsson, S. & Svensson, M. (2020) Willingness to Pay for Health Improvements Using Stated Preferences: Prevention Versus Treatment. Value in Health, Volume 23, Issue 10, 2020;1384-1390,

Wetering, Van de, E. J., Stolk, E. A., Van Exel, N. J. A. & Brouwer, W. B. F. (2013) Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *Eur J Health Econ.* 14:107-115.

Yun, M., Lee, S.H., Kang, H.G.: Analysis of the relationship between risk perception and willingness to pay for nuclear power plant risk reduction. Sci. Technol. Nucl. Install. (2016). https://doi.org/10.1155/2016/6293758

Zwaap, J., Knies, S., Van der Meijden, C., Staal, P. & Van der Heiden L. (2015) *Cost-effectiveness in practice [in Dutch: Kosteneffectiviteit in de praktijk]* No. 2015076142. Diemen, the Netherlands: Zorginstituut Nederland.

Appendix I: Questionnaire

'Imagine there is a patient population having a rare disease. The health state they are currently in is described in 'health state I' (left). Without treatment, the health of these patients will deteriorate and result in 'health state II' (right). This disease occurs within 5 out of 10.000 people per year in the Netherlands. At the same time, there is no other treatment available to these patients.

Health state I	**Health state II**
I have severe problems in walking about	I am unable to walk about
I have moderate problems washing or dressing myself	I am unable to wash or dress myself
I have moderate problems doing my usual activities	I have moderate problems doing my usual activities
I have moderate pain or discomfort	I have moderate pain or discomfort
I am slightly anxious or depressed	I am moderately anxious or depressed

There is a painless medicine (without any side effects) which ensures that the patient will not go to the worse health state during a year and will stay in the better health state I.

The financing of the medicine will be done through an increase in the monthly insurance premium for all Dutch residents for one year. This (increase in) premium is paid out-of-pocket. Your income would be the same in this whole period to your current income.

Imagine you would have to pay for this drug. Please go through amounts below, from lowest to highest, and choose the highest amount that you would certainly be willing to pay extra per month in health insurance premium for a drug that ensures that the health condition does not deteriorate in a rare disease without alternative treatment.

(For example, if you are sure you would like to pay $14 a month for this drug, but not sure if you would be willing to pay $16 a month for it, choose $14).

Remember that you are not part of the risk group.

€0	€0.50	€1	€1.50	€2	€2.50	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	Meer
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

You have indicated that you want to pay €4 extra per month for at least 12 months in health care premiums for a medicine that ensures that the health condition does not deteriorate in the event of a rare disease without alternative treatment. Would you like to go down the amounts below, from low to high, and choose the lowest amount that you definitely wouldn't want to pay per month for this drug? (For example, if you're sure you wouldn't want to pay $20 a month for this drug, but might be willing to pay $18 a month for it, choose $20).

Note that this amount will be equal to or higher than in the previous question and remember that you are not part of the risk group.

--

You have indicated that you want to pay 4 euros per month for at least 12 months, but certainly not 8 euros per month extra in premium for a medicine that ensures that the health condition remains the same. Would you indicate the amount between 4 and 8 euros that is closest to what you are willing to pay for this drug?

--

You have indicated that you want to pay a maximum of 6 euros per month in extra premium for the medicine. How confident are you that you would actually pay this amount if you were asked to do so at this time?

- Certainly
- Probably
- Maybe, maybe not
- Probably not
- Certainly not

--

Appendix II: Questionnaire (Dutch)

Dit was het tweede onderdeel. We gaan nu door naar het derde onderdeel, welke uit 3 scenario's bestaat. In elk scenario wordt een gezondheidstoestand beschreven die kan verslechteren op het moment dat er geen interventie (behandeling) plaatsvindt. In dit onderzoek is de interventie een medicijn. Belangrijk om te onthouden in de vragenlijst, is dat uw familie, vrienden en/of kennissen onderdeel kunnen zijn van de patiëntenpopulatie alsmede onbekende mensen. Uzelf behoort niet tot deze risicogroep.

Stelt u zich voor dat er een patiëntengroep is die een zeldzame ziekte heeft. De gezondheid van deze patiëntengroep is omschreven in gezondheidstoestand I (links). Zonder interventie zal de gezondheid van deze patiënten verslechteren naar gezondheidstoestand II (rechts).
Deze ziekte komt in Nederland voor bij 5 op de 10.000 personen per jaar. Er is tevens geen andere behandelmogelijkheid voor de patiëntengroep.

Gezondheidstoestand I	*Gezondheidstoestand II*
Ik heb ernstige problemen met lopen	Ik ben bedlegerig
Ik heb matige problemen mijzelf te wassen/ aan te kleden	Ik ben niet in staat mijzelf te wassen/ aan te kleden
Ik heb matige problemen met mijn dagelijkse activiteiten	Ik heb matige problemen met mijn dagelijkse activiteiten
Ik heb matige pijn/ andere klachten	Ik heb matige pijn/ andere klachten
Ik ben een beetje angstig/ somber	Ik ben matig angstig/ somber

Er is een medicijn (zonder bijwerkingen) dat ervoor zorgt dat de patiënt niet naar de slechtere gezondheidstoestand II gaat gedurende een jaar. De patiënten blijven dan in de betere gezondheidstoestand I.

Het medicijn wordt betaald door een tijdelijke verhoging van de maandelijkse verzekeringspremie voor alle Nederlanders, gedurende 1 jaar. De premie betaalt u uit het eigen inkomen. Uw inkomen is in de gehele periode gelijk aan uw huidige inkomen.

← →

Wilt u de onderstaande rij met bedragen afgaan, van laag naar hoog, en het hoogste bedrag kiezen dat u zeker wel per maand extra zou willen betalen aan zorgverzekeringspremie voor een medicijn wat ervoor zorgt dat de gezondheidstoestand niet verslechtert bij een zeldzame ziekte zonder alternatieve behandeling?
(Bijvoorbeeld: als u zeker bent dat u €14 per maand zou willen betalen voor dit medicijn, maar niet zeker of u er €16 per maand voor over zou hebben, kiest u €14).

Onthoud dat u geen onderdeel uitmaakt van de risicogroep.

€0	€0.50	€1	€1.50	€2	€2.50	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	Meer
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

← →

U heeft aangegeven dat u zeker 12 maanden €8 per maand extra wilt betalen aan zorgpremie voor een medicijn dat zorgt dat de gezondheidstoestand niet verslechtert bij een zeldzame ziekte zonder alternatieve behandeling. Wilt u de onderstaande bedragen afgaan, van laag naar hoog, en het laagste bedrag kiezen dat u zeker niet per maand zou willen betalen voor dit medicijn? (Bijvoorbeeld: als u zeker bent dat u geen €20 per maand zou willen betalen voor dit medicijn, maar er misschien wel €18 per maand voor over zou hebben, kiest u €20).

Let op dat dit bedrag gelijk aan of hoger zal zijn dan in de vorige vraag en onthoud dat u geen onderdeel uitmaakt van de risicogroep.

€0	€0.50	€1	€1.50	€2	€2.50	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	Meer
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

← →

U heeft aangegeven dat u zeker 12 maanden €8 per maand, maar zeker geen €12 per maand extra aan premie wilt betalen voor een medicijn dat zorgt dat de gezondheidstoestand gelijk blijft. Wilt u het bedrag aangeven (tussen €8 en €12) dat wat u bereid bent om te betalen voor dit medicijn het best benadert?

← →

U heeft aangegeven dat u maximaal 10 euro per maand aan extra premie wilt betalen voor het medicijn. Hoe zeker bent u ervan dat u dit bedrag daadwerkelijk zou betalen als het op dit moment van u gevraagd zou worden?

- ○ Zeker wel
- ○ Waarschijnlijk wel
- ○ Misschien wel, misschien niet
- ○ Waarschijnlijk niet
- ○ Zeker niet

← →

U heeft aangegeven dat u niet bereid bent meer dan €0 te betalen voor een medicijn dat zorgt dat de gezondheidstoestand niet verslechtert. Wat is uw belangrijkste reden hiervoor?

- ○ Ik kan niet meer dan €0 betalen
- ○ Medicijnen zijn niet meer dan €0 waard voor mij
- ○ Ik vind de medicijnen meer dan €0 waard, maar geef mijn geld liever uit aan iets anders
- ○ Ik vind dat de basis zorgverzekering al te hoog is
- ○ De waarde van zorg en gezondheid is voor mij niet in geld uit te drukken
- ○ Anders, namelijk:

← →

Appendix III: Regression models per scenario

In the results section, the main model is presented. Here, the additional different models per scenario are placed.

Rare

DV = logWTP rare	Model I			Model II			Model III		
	Coef.	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*
Female (yes)	.0614	.1645	0.709	.0662	.1543	0.669	.1136	.2170	0.602
Age	.0179	.0058	0.002*	.0193	.0066	0.004	1.052	.4221	0.014*
Randstad (yes)	.1372	1.467	0.351	.2287	.1784	0.202	.1857	.2291	0.519
Education (yes)									
High	2.050	1.321	0.123	.1604	.1974	0.418	.2102	.2960	0.479
Marital status (single = ref.)									
- Married	-.342	.2406	0.163	-.2598	.2350	0.271	-.7062	.4011	0.081**
- Cohabitant	-.403	.2250	0.075**	-.4270	.2104	0.044*	-.4801	.3358	0.156
- Widow	-.083	.9679	0.675	-.0464	.9388	0.961	-.7806	1.168	0.506
- Divorced	-1.291	.5742	0.026*	-.7258	.4220	0.088**	-1.756	1.160	0.133
Bruto household income (0-300 = ref.)	.0250	.2239	0.911				.2472	.2884	0.393
3000-5000	-.2160	.2434	0.376				.0917	.3017	0.762
5000-7500	-.0361	.2781	0.897				.0590	.3437	0.864
>7500									
Working (yes)	.0667	.2208	0.763	0.1000	.1961	0.612	-.2821	.2635	0.287
Children (yes)							.0706	.3152	0.823
Household				-.1018	.0711	0.155	-.2015	1.001	0.841
Chronic sick person in environment				.2129	.1682	0.208	.0473	.2379	0.843
Cons.	.7081	.6290	0.088**	1.055	.4593	0.023*	.6163	1.813	0.236
Adj. R^2	0.0709			0.0626			0.0593		
Prob> F	0.0291*			0.0436*			0.0716**		
N	143			150			143		

Common

DV = logWTP common	Model I			Model II			Model III		
	Coef.	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*
Female (yes)	.1043	.1564	0.506	.0339	.1406	0.810	.2560	.3913	0.154
Age	.0118	.0065	0.070**	.0136	.0060	0.025*	.5617	.1906	.0812**
Randstad (yes)	.0869	.1702	0.610	.1400	.1626	0.391	.0056	.2015	0.978
Education (yes)									
High	.3206	.1513	0.036*	.5073	.1798	0.005*	.3034	.2577	0.242
Marital status									
(single = ref.)	-.1460	.2199	0.508	-.0851	.2141	0.691	-.2867	.3538	0.420
- Married	-.2037	.2057	0.324	-.2277	.1916	0.237	-.1570	.2958	0.597
- Cohabitant	.2553	.8847	0.773	.4480	.8553	0.601	-.0281	1.028	0.978
- Widow	-.9216	.5249	0.081**	-.5142	.3845	0.183	-2.612	.7528	0.001*
- Divorced									
Bruto house									
income (0-300 =	.1177	.2047	0.566				.2008	.2656	0.451
ref.)	.0572	.2225	0.798				.2569	.2794	0.360
3000-5000	.2147	.2542	0.400				.3960	.3161	0.213
5000-7500									
>7500									
Working (yes)	-.2219	.2019	0.274	-.1039	.1786	0.562	-.2456	.2373	0.303
Children (yes)							.0943	.2843	0.741
Household				-.0743	.0648	0.253	.0200	.8977	0.982
Chronic sick				.1141	.1532	0.458	.1003	.2102	0.634
person in									
environment									
Cons.	.6674	.5912	0.261	1.199	.4185	0.005*	-.9352	1.627	0.567
Adj. R^2	0.0472			0.0559			0.0912		
Prob> F	0.0972**			0.0589**			0.0546**		
N	143			150			143		

Frequent occurring

DV = logWTP frequent	Model I			Model II			Model III		
	Coef.	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*	*Coef.*	*Std. Error.*	*P>[t]*
Female (yes)	-.0185	.1782	0.917	-.0604	.1596	0.706	.1460	.2034	0.475
Age	.0030	.0074	0.682	.0049	.0068	0.470	.0036	.4152	0.855
Randstad (yes)	.0169	.1939	0.930	.0392	.1846	0.832	.0111	.2158	0.959
Education (yes)									
High	.3995	.1725	0.022*	.6124	.2042	0.003	.4203	.2792	0.135
Marital status (single = ref)	-.1577	.2506	0.530	-.1797	.2431	0.461	-.1770	.3821	0.644
- Married	-.1511	.2344	0.520	-.2161	.2176	0.322	-.0190	.3151	0.952
- Cohabitant	.3595	1.008	0.722	.4258	.9711	0.662	.1789	1.098	0.871
- Widow	-.7399	.5982	0.218	-.4188	.4365	0.339	-1.808	.8054	0.027*
- Divorced									
Bruto house income									
(0-300 = ref.)	-.0108	.2333	0.963				.1368	.2830	0.630
3000-5000	-.0349	.2535	0.890				.2218	.2980	0.459
5000-7500	.0949	.2897	0.744				.2605	.3355	0.439
>7500									
Working (yes)	-.3341	.23006	0.149	-.2770	.2028	0.174	.1733	.9554	0.856
Children (yes)									
Household				-.0494	.0736	0.503	.3127	.1903	0.104
Chronic sick person in environment				-.07434	.1740	0.670			
Cons.	1.2423	.6737	0.067**	1.912	.4751	0.000	1.406	1.728	0.418
Adj. R^2	0.0211			0.0182			0.018		
Prob> F	0.0827**			0.2594			0.4575		
N	143			150			143		