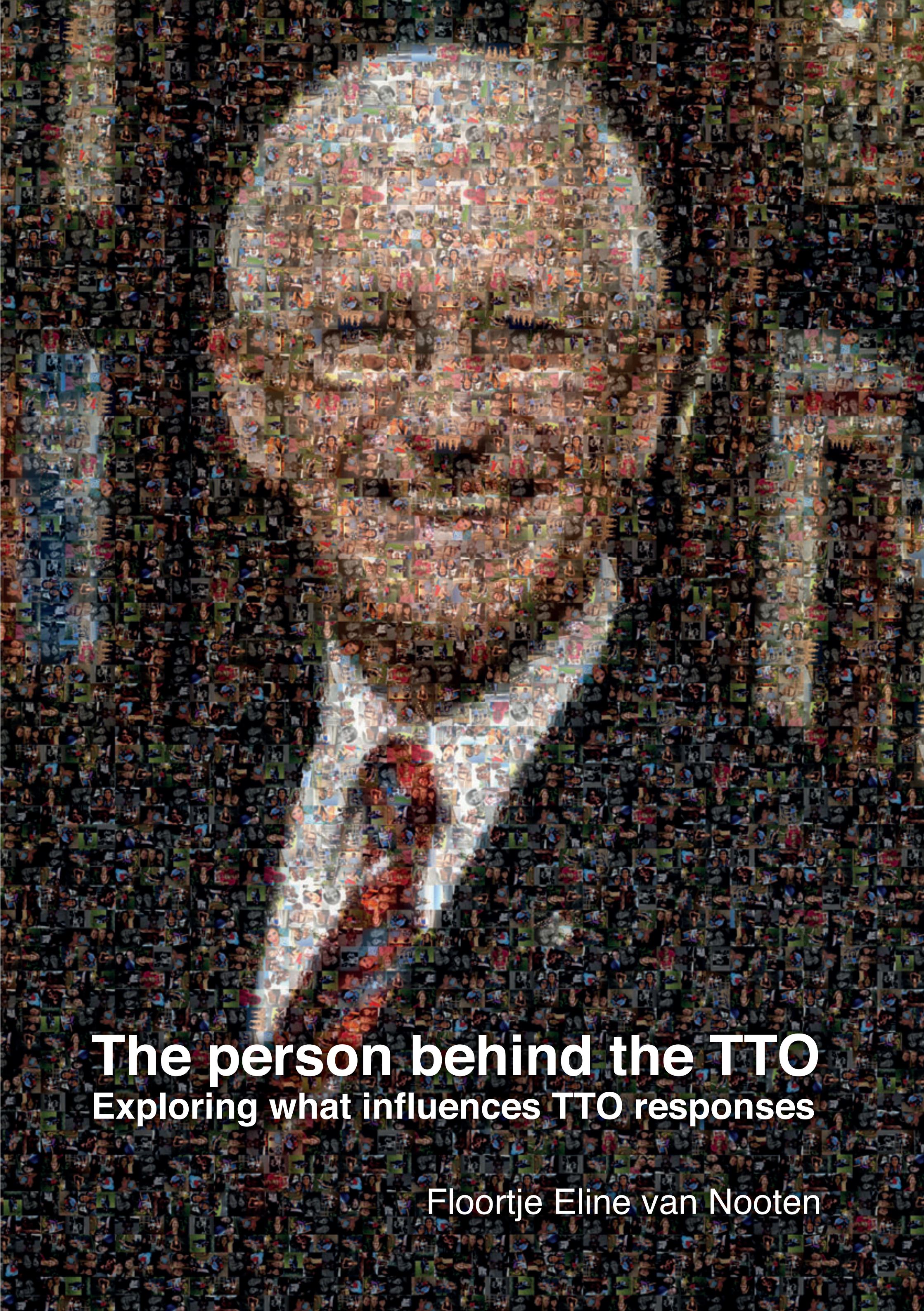


The person behind the TTO

Exploring what influences TTO responses

Floortje Eline van Nooten

The Person Behind the TTO

Exploring what influences TTO responses

Floortje Eline van Nooten

ISBN: 978-94-6375-280-0

Cover: Photos provided by George Torrance and family and friends of Floortje van Nooten

Printed by Ridderprint BV, Ridderkerk, the Netherlands

The Person Behind the TTO

Exploring what influences TTO responses

De persoon achter de TTO

Wat beïnvloedt TTO antwoorden?

Proefschrift

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

Prof. dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op donderdag 21 maart 2019 om 15:30 uur

door

Floortje Eline van Nooten

geboren te Heilbronn, Duitsland

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotoren

Prof. dr. W.B.F. Brouwer

Prof. dr. N.J.A. van Exel

Overige Leden

Prof. dr. C.D. Dirksen

Prof. dr. M.P.M.H. Rutten-van Mölken

Dr. G.A. de Wit

“***Het is een dag als vandaag dat ik mijn zonnepet draag, met mijn zonnenpet op en een lach op mijn kop is het een dag als vandaag dat ik mijn zonnepet draag***”

(uit een aflevering van de radiostrip “Koek en ei”, zoals vastgelegd op de grammofoonplaat “Alles op een plaat”, 1959)”

CONTENT

1.

Introduction

1.1 Introduction

For health policymakers and societies alike, the continuous introduction of new health technologies appears to be a mixed blessing. On the one hand, these technologies have likely contributed to a substantial increase in the average length and quality of life in the past decades. For example, between 2005 and 2015 life expectancy at birth in the Netherlands increased from 77.2 to 79.7 years for men and from 81.6 to 83.1 years for women (CBS, 2016a). Furthermore, it is estimated that by 2030 life in good health will have further increased by 2 to 3 years (CBS, 2014) On the other hand, the innovations that have likely contributed to this success, have also contributed to a substantial increase in healthcare expenditures. For example, during the same period (2005 – 2015) in the Netherlands, healthcare expenditure per capita increased from €4,115 to €5,628 (CBS, 2016b). Although it must be emphasized that not all health improvements resulted from increased healthcare expenditures and that the latter were not only due to new technologies (Douven *et al.*, 2006), the continuous introduction of innovative medical technologies does raise an important challenge.

Despite the growth in expenditure, healthcare budgets ultimately are limited. Hence, in order for society to spend these limited resources optimally, choices have to be made as to which (new) treatments to fund. If an important goal is to generate as much health as possible with the available budget, preferably those technologies are funded that yield most health per invested Euro. This ensures spending the healthcare budget in the most efficient way. In order to assist health care policy makers to attain such optimal spending of scarce resources, economic evaluations of health interventions are increasingly performed (Gold et al., 1996). Such an evaluation compares the costs and (health) benefits of some intervention to a relevant comparator (e.g., care as usual) to provide evidence of value-for-money of that intervention (Gold et al., 1996). Several countries, including Australia, Canada, the Netherlands, and the United Kingdom, systematically use economic evaluations to inform funding decisions, as described in their respective guidelines (e.g. CADTH, 2006; NICE, 2009; PBAC, 2008 and ZiN, 2016). While in using economic evaluations the focus has long been on evaluating pharmaceuticals, they are increasingly used in other areas as well, including surgery, physiotherapy, social care and prevention (Blankers *et al* 2012; Carroll *et al.*, 2014; Clarke *et al.*, 2016; Fobelets & Pil, 2015; Fobelets *et al.*, 2015; Groenewoud *et al.*, 2007; Sutcliffe *et al.*, 2013; van Wetering *et al.*, 2010; van Gils *et al.*, 2009; Warmerdam *et al.*, 2010). Although the exact influence of the results of economic evaluations on final allocation decisions may be debated (Franken *et al*., 2016), the increasing use of economic evaluations in healthcare and its intended role in the decision making process emphasizes the need for clear, appropriate and consistently used methods. Assessing the costs and benefits of a health technology is not a straightforward exercise and many methodological debates are ongoing, ranging from the appropriate perspective to take (Brazier *et al.*, 2007; Heintz *et al*.; 2016; Versteegh & Brouwer, 2016), how to discount future costs and effects (Brouwer *et al*., 2005; Claxton *et al*. 2011; Heintz *et al*., 2016), how to include informal care and productivity costs (Heintz *et al.*, 2016; Hoefman *et al*., 2013; Krol & Brouwer, 2013; Krol & Brouwer, 2014) to how to estimate and express uncertainty around estimates (Griffin *et al*., 2011; Heintz *et al*, 2016).

1.2 Health benefits and the QALY

One area of ongoing research and debate is the measurement and valuation of health benefits (Drummond *et al*., 1997; Drummond *et al*., 2015; Gold *et al*., 1996,). In that context it is important to distinguish between three different types of economic evaluation: cost-benefit, cost-effectiveness and cost-utility analyses. These types of economic evaluation differ in how they measure and value health effects (Brouwer *et al.*, 2008; Drummond *et al*., 2015). In a cost-benefit evaluation, both costs and health effects are expressed in monetary terms. The advantage of such an approach is that costs and benefits can be compared directly, but it is difficult and contentious to express health in money terms. In cost-effectiveness analyses, costs are expressed in monetary terms and the benefits in so-called ‘natural units’, which, depending on the intervention, may take forms like life-years gained, percentage reduction of blood pressure or hip-fractures avoided. The advantage is that such outcomes resonate well in clinical practice, but their comparability across diseases is limited. In cost-utility analyses costs are also expressed in monetary terms, but the benefits are expressed in quality-adjusted life years (QALYs), a generic health outcome measure (Whitehead & Ali, 2010). The advantage of using QALYs is that they are comparable across diseases and reflect preferences (utilities) of people for different health states. Most countries that use economic evaluations within their decision process prefer cost-utility analyses, using QALYs as their outcome measure (Heintz *et al.*, 2016, Jakubiak-Lasocka & Jakubczyk, 2014). . This makes the QALY an important outcome measure in health care.

The QALY is a composite measure that combines length-of-life and quality-of-life into a single measure (Culyer *et al*., 2014; Whitehead & Ali, 2010). Quality-of-life is expressed on a scale that is anchored on two health states: perfect health, which is given the value 1, and dead, which is given the value 0. Most imperfect health states have a value in between 0 and 1, although very poor health states can have negative quality of life values. In order to calculate QALYs, the quality of life value of a health state is multiplied with the duration of that state (Whitehead & Ali, 2010). One year in perfect health equals 1 QALY (Culyer *et al*., 2014; Prieto & Sacristan, 2003). Two years in a state with a quality-of-life value of 0.5 also equals 1 QALYs (Drummond *et al*., 1997; Drummond *et al*., 2015; Prieto & Sacristan, 2003). If an intervention can bring someone who would live 5 more years in a health state with value 0.4 back to perfect health (without prolonging life duration), the gain is 0.6 QALY per year and (5*0.6=) 3 QALYs in total. By confronting such gains with the incremental costs of the involved intervention, an incremental cost-per-QALY ratio (ICER) can be calculated. Treatments with the lowest ICERs, i.e. which yield the most QALYs per euro invested and therefore would be the most efficient spending of healthcare budget, would have priority for funding, *ceteris paribus*.

Deriving QALY values

As the quality-of-life values play an important role in calculating the health benefits (QALYs) from a treatment, and feed into the decision making process, it is of course highly important that these values are valid and reliable. In the past, four methods have been used for obtaining quality-of-life values for health states: the Visual Analogue Scale (VAS); Standard Gamble (SG); Time Trade-Off (TTO), and Discrete Choice Experiment (DCE) (Whitehead & Ali, 2010; Brazier *et al*., 2012). We will briefly introduce them below.

With a VAS, a respondent in a health valuation exercise is asked to indicate on a single straight line, normally ranging from 0 to 100, where she would place a particular health state. This health state can be own current health, dead, or some hypothetical health state. The single straight line has verbal and/or numerical descriptors at each end, defining the endpoints in terms like ‘best possible health’ or ‘perfect health’ with value 100 and ‘dead’ or ‘worst possible health’ with value zero. The place marked on the scale is indicative of the relative value of the health state compared to the upper and lower end of the scale. An example of a frequently used VAS scale is the EQ-VAS (Euroqol, 2017). Its usefulness in deriving utility scores has sometimes been contested (Parkin & Devlin, 2006).

Example 1: Visual Analogue Scale

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and dead is marked 0.

We would like you to indicate on this scale how good or bad your own opinion you would rate the health state lower back. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

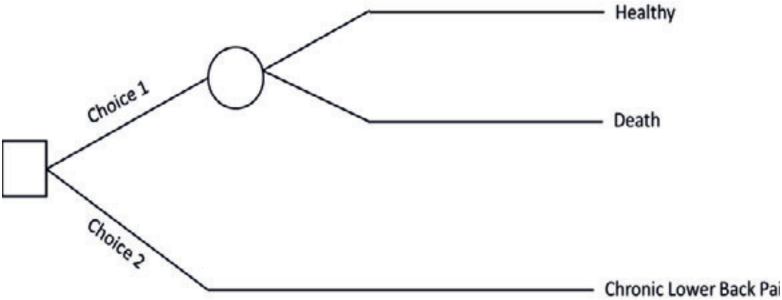
In a SG exercise, an individual is asked to choose between two alternatives prospects. One option is certain and involves living for a defined period of time in an imperfect health state. The other option is uncertain and involves a gamble with either a better or a worse outcome than the certain one. Typically, the gamble is between living in perfect health for the same period of time as the certain option or immediate death. The respondent is asked the highest risk of immediate death she is willing to take to become indifferent between the certain option or the gamble. Suppose a respondent has to choose between living 1 more year with low back pain or a risky operation which might restore perfect health during that year or results in immediate death. She is indifferent between both options when the chance of death is 30% and the chance of restoring perfect health is 70%. Given that perfect health is assigned the value of 1 and dead the value of 0, this choice implies that the value of the health state back pain can be calculated as: 0.7*1 + 0.3*0 = 0.7. Moving (for one year) from this point to perfect health would again imply a gain of 0.3 QALY. Living with back pain for 5 years implies a total amount of 5*0.7 = 3.5 QALYs.

Example 2: Standard Gamble

Quality of life values for lower back pain using SG

Please imagine having chronic lower back pain

You can either choose to live with chronic lower back pain or you can choose to gamble. In the gamble, your choices are: live in perfect health or instant death.



*simplified representation of a SG task

DCE as a technique to obtain quality-of-life values is receiving increasing attention. In a DCE, a respondent is asked to perform a number of pairwise choices between health states (e.g. Rowen *et al*., 2015; Norman *et al*., 2013). The respondent is asked to indicate which health state he/she prefers based on the characteristics describing the health states. The calculation of the quality-of-life values in a DCE is not as straightforward as in the other methods, but depends on sophisticated statistical analyses (Bansback *et al*.,2012; Brazier *et al*., 2012; Flynn, 2010; Rowen *et al*., 2015)

Example 3: Discrete Choice Experiment

Quality of life values for lower back pain using DCE*

Would you either live in option A, with chronic back lower pain for the described number of years and then die or live in option B, Perfect Health for the described number of years and then die.

Option A: Chronic Lower Back Pain	**Option B: Perfect Health**
Moderate problems walking about	No problems walking about
I have no problems with selfcare	I have no problems with selfcare
I have some problems with performing my usual activities	I have not problems with performing my usual activities
I have severe day time pain and no pain during the night	I have no pain or discomfort
I am not anxious or depressed	I am not anxious or depressed

In a TTO exercise, respondents are asked to imagine two health streams: a longer period in poorer health or a shorter period in better (typically perfect) health. The exercise subsequently varies the amount of time lived in perfect health in order to find the period in perfect health that would make the respondent indifferent between the two options.

Suppose a respondent is presented two prospects. One involves living 10 years with low back pain after which the person dies. The other involves living less than 10 years in perfect health after which the person dies. If the respondent indicates to be indifferent between 10 years with low back pain and 7 years in perfect health (with value 1), the quality of life value of low back pain equals 7/10=0.7.

Again, restoring perfect health (during one year) for a person with low back pain would imply a gain of 0.3 QALY. Living 5 years with the condition of low back pain means a total amount of 5*0.7 = 3.5 QALYs lived.

Example 4: TTO

Quality of life values for lower back pain using TTO*

You can chose between option A and option B. If you chose option A you will live another number of years (e.g. 10) with chronic lower back pain, after which you will die. If you opt for option B, you will live for less years, but in perfect health, after which you will die.

Option A: Chronic Lower Back Pain	Option B: Perfect Health
Moderate problems walking about	No problems walking about
I have no problems with selfcare	I have no problems with selfcare
I have some problems with performing my usual activities	I have not problems with performing my usual activities
I have severe day time pain and no pain during the night	I have no pain or discomfort
I am not anxious or depressed	I am not anxious or depressed

Comparison of the methods

Due to its impact in calculating the benefits in economic evaluations, the choice of method to elicit health state valuations is important. The method that generates the most valid and reliable quality of life values should be used. However, the choice is not trivial because all methods described above have their merits and flaws.

An important issue related to the VAS is that it does not require respondents to trade-off one health state for another. Hence it is not always considered a preference-based method. Its usefulness in economic evaluations has been contested (Whitehead & Ali, 2010) but also defended (Parkin & Devlin, 2006). The VAS is easy to understand and use for respondents.

The standard gamble is derived directly from expected utility theory. It is the only of the mentioned measures that involves the risk attitude of the respondent and has been reported to be a reliable method (Froberg & Kane, 1989; Sharma et al., 2002). However, it is susceptible to biases such as loss aversion (people are more sensitive to losses than to gains) and probability weighting (the tendency to overweight low and high chances as compared to intermediate chances'van Osch *et al*., 2004; Bleichrodt, 2002). Furthermore it is more cognitive demanding for respondents (Gold *et al.*,1996).

The DCE has been introduced fairly recently in the field of health state valuations, compared to the other methods. DCE is derived from random utility theory (Flynn, 2010; Stolk *et al*., 2010). Its application is increasing, also in the field of health state valuations, showing promising results (Bansback *et al*., 2012; Brazier *et al*., 2012; Robinson *et al*. 2017; Stolk *et al*., 2010). However, the reliability of health state values obtained with DCE and potential biases in the method require further exploration (Brazier *et al*., 2012; Robinson, *et al*., 2017; Stolk *et al*., 2010). DCEs are often considered to be easy to use by respondents (Norman *et al*., 2013).

The TTO can also be considered a preference based method, because the respondent is asked to trade-off two goods (length and quality of life). Like the SG, the TTO is susceptible to biases (e.g. loss aversion, utility curvature and scale compatibility). Some of these biases may influence responses in opposite directions, to some extent cancelling each other out (van Osch et al., 2004; Bleichrodt, 2002). Furthermore, the TTO task is quite easy to understand for respondents (Morimoto & Fukui, 2002; Torrance *et al*., 1972) and its results have shown to have good reliability (Froberg & Kane, 1989; Sharma *et al*., 2002).

While the field is evolving rapidly, at the time this research started, and to some extent still, the TTO is a much used method for obtaining quality-of-life values for health states, including many national valuation sets ('tariffs') (Arnesen & Trommald, 2005; Euroqol, 2017; Neumann *et al*., 1997). For this reason, we chose to focus on the TTO as a method for deriving health state valuations here. Below this method is further introduced.

1.3 Time Trade-Off

Torrance and Sackett (1972) developed the TTO as an alternative for the standard gamble. The TTO involves a trade-off between living more years in an imperfect health state or living less years in perfect health. Often the longer period is set to 10 years, hence the shorter period (X) should be somewhere between 0 and 10 years. Assuming no discounting, this can be formalized in the following way (van Nooten et al., 2009):

$$10 * V_{(h)} = X * V_{(ph)} + (10-X) * V_{(dead)} \quad (1)$$

In this equation $V_{(h)}$ is the value or quality-of-life weight of the imperfect health state, $V_{(ph)}$ denotes the value of perfect health, $V_{(dead)}$ represent the value for the health state denoted as dead, 10 is the number of years in the imperfect health state and X the number of years in the perfect health state.

If $V_{(dead)}$ equals zero, the equation can be reformulated into:

$$10 * V_{(h)} = X * V_{(ph))} \quad (2)$$

Moreover, normalizing $V_{(ph)}$ to 1, this leads to an expression of $V_{(h)}$:

$$V_{(h)} = X/10 \quad (3)$$

Eliciting the number of years in perfect health, X, that is considered equal to the time period in imperfect health can be done through a single question (matching), or through a more elaborate choice task (sometimes called "ping pong") in which the number of years is changed until the respondent is indifferent between the alternatives (Attema *et al*., 2013; Brazier & Ratcliffe, 2007).

Note that respondents may consider certain health states to be 'worse than dead' (WTD). In those cases the conventional TTO method is not appropriate to obtain valuations of the health state, as respondents cannot give up more years than were provided in the imperfect health state (*e.g.* in a 10-year TTO, the respondent can only give up a maximum of 10 years). A different TTO approach can be used for valuing WTD health states (Attema et al., 2013; Devlin et al., 2011). More recently, lead-time and lag-time TTO exercises were introduced to create one method capable of valuing all types of health states (Attema & Versteegh, 2013; Augustovski *et al*., 2013; Devlin *et al*., 2013), but these still require further research. Here, the focus will be on the conventional TTO exercise.

Factors influencing TTO scoring

Although the TTO is one of the most frequently used methods to obtain health state values, it remains important to ensure that the TTO is a valid and reliable instrument, and to study the factors that influence TTO responses. Otherwise the QALYs calculated using TTO results could be unreliable, which could ultimately lead to non-optimal health care decisions. Arnesen & Trommald (2005) performed a literature review of studies applying TTO and found large differences between TTO values for the same health states between studies, which can be attributed to methodological factors and respondent characteristics (Arnesen & Trommald, 2005).

Attema *et al.* grouped the items related to methodology that influence TTO responses into three headings: methodological, analytical and procedural (Attema *et al*., 2013) Methodological issues include value ranges, the time frame, the iteration process and the response scale and are related to how the trade-off between quantity and quality-of-life is conducted. Analytical issues include exclusion criteria, definition of anchor points, analysis of worse-than-dead values and time preference and are related to the analysis of the TTO data. Procedural issues include mode-of-administration, visual aids, sampling frame, context effects and indirect valuation, and are considered to be associated with the experiment itself, such as presentation of the TTO.

The other factor influencing TTO responses is respondent characteristics (Arnesen & Trommald, 2005). The literature review performed by Arnesen & Trommald (2005) found there was no discussion of the influence of respondent characteristics on TTO responses in two-thirds of the studies they reviewed (Arnesen & Trommald, 2005). This suggests that the influence of respondent characteristics is still understudied (Arnesen & Trommald, 2005). However, some attention has been given to the influence of respondent characteristics in the literature. For example, one of the first large TTO studies, with 3,395 respondents, found that valuations can be impacted by age, gender and marital status (Dolan *et al*., 1996). In the same study, the influence of other background variables (such as social class, and education) was found to be insignificant or of minor importance (Dolan *et al*., 1996). Other studies have also investigated the effect of age, gender, marital status and other characteristics. The results differ between studies both in terms of significance of variables and in terms of direction of their influence (Essink-Bot *et al*., 2007; Hsu *et al*., 2009; Kontodimopoulos & Niakas, 2006; Prosser *et al*., 2011; Rutten-van Mölken *et al*., 2009).

Some studies suggest that other demographic characteristics, such as having children (Devlin *et al.*, 2011; Essink-Bot *et al*., 2007; van der Pol & Shiell, 2007) or education and socio-economic class (Devlin *et al*., 2011; Prosser *et al*., 2011), play a role in TTO responses. Next to demographic characteristics it might also be of interest to investigate the influence of personal traits, “attitudinal” characteristics, on TTO responses. Research related to personal traits and “attitudinal” characteristics (*e.g.* beliefs about the future and death) is even scarcer. It seems, given the construct of the TTO exercise, that personal trait-related respondent characteristics like beliefs about future quality-of-life or expectations regarding length of life could also potentially influence TTO scoring (Essink-Bot *et al*., 2007; Handler *et al*., 1997; Mrus *et al*., 2006; van Nooten *et al*., 2009).

The main purpose of most of the studies that investigated the influence of respondent characteristics was, however, not to understand the influence of respondent characteristics on trade-offs, but to obtain health state valuations. The analyses to assess the impact of respondent characteristics on TTO responses were usually secondary objectives, and performed to test for heterogeneity in the sample (Essink-Bot *et al*., 2007; Izadpanah *et al*.*,* 2013; Kontodimopoulos & Niakas, 2006). The effect sizes were typically small, if reported at all (Dolan *et al*., 1996, Essink-Bot *et al*. 2007). Therefore, much remains unknown regarding the variety of respondent characteristics that influence TTO responses and their impact.

Typically, national tariffs for QALY values are obtained from health state valuations by the general public. This means that the normal source of valuations is not patients experiencing a specific health state, but a sample from the general public who value hypothetical health states (Versteegh & Brouwer, 2016). The sample selected from the general public should be a representative subset in order for the final tariffs to be representative. Hence, it is important to understand which respondent characteristics should be used to sample. Furthermore, to compare TTO responses from different studies it is important to provide the list of (influential) demographic and other variables.

1.4 Aim of This Thesis

This thesis aims to further investigate respondent characteristics influencing TTO responses and how they influence TTO responses. Therefore, the overall aim of this thesis is to contribute to a better understanding of which respondent characteristics and attitudinal factors influence TTO responses. This research will go beyond the frequently used demographic respondent characteristics such as age, gender and education. It will extend previous research including both previously researched variables as well as novel ones. Certain aspects of the TTO exercise which relate to a respondent's personal expectations, preferences, attitudinal (personal emotional-driven) traits, could influence TTO responses and will be investigated. For example, the TTO focuses on trading off life time, hence accelerating time of death, and in certain exercises reduced life-expectancy of the respondent is suggested anyhow, for example in a 10-year TTO. Therefore, opinions and attitudes that respondents have regarding death or euthanasia could play a role when completing a TTO exercise. Also, ideas individuals have about their future (*e.g.* expectations of length and quality-of-life in the future, certain life events that might still occur such as becoming a grandparent) that might be theoretically taken away from them in a TTO exercise could influence the scoring.

Furthermore, death and poorer health status of the respondent hypothetically implied in TTO exercises does not only influence the respondent themselves but also persons close to the respondent, such as children and partners. It has already been shown that marital status can be a significant variable influencing TTO responses. However, previous research focused on a married partnership status. In the Dutch society more and more couples decide not to marry, hence it would be interesting to understand if next to partnership, the type of partnership status influences TTO responses.

Therefore, this thesis addresses the following research questions:

1. Does subjective life expectancy (SLE) impact the willingness to trade-off (WTT) and the number of years traded-off for health state valuations?
2. How does the awareness of the reduced life span implied by a 10 year TTO affect health state valuations?
3. What is the influence of beliefs regarding future health and death, as well as desires to witness certain life events, on respondents' health state valuations?
4. Which responder characteristics influence TTO responses, with an emphasis on consideration of significant others, such as partners and children?
5. Do cultural differences in TTO responses exist?

Rather than looking at each respondent characteristic individually it is also interesting to understand if there are "groups" of respondents based on TTO response and if certain respondent characteristics can predict in which TTO scoring group a respondent would fall.

Thesis Outline

This thesis consist of 9 chapters. It starts with an introduction (Chapter 1). Chapter 2 provides an overview of the type of respondent characteristics that have been mentioned in the literature that could have an influence on TTO scoring. It aims to identify qualitatively how and when respondent characteristics might influence TTO valuations. This chapter uses a literature review to create this overview. The following chapters 3 to 8 present results based on empirical work.

Chapter 3 aims to answer question 1 and describes the results of two studies investigating whether SLE influences TTO responses. One study explores this in a lifetime TTO, the other in a 10-year TTO exercise. Question 2 is explored in chapter 4. Chapter 4 describes empirical evidence of how awareness of the reduced life span implied by 10-year TTO affects TTO responses. Chapter 5 focuses on question 3 and investigates the influence of beliefs about the future and death on TTO scoring. Chapter 6 aims to answer question 4 and describes empirical work on the influence of significant others in TTO exercises, using a 10-year TTO exercise.

Next to looking at the influence of significant others, such as partner or children, and beliefs about the future and death as reflected in question 5, Chapter 7 present results of cultural differences in TTO responses using the EQ-5D valuation sets from Japan, UK and Spain.

Chapter 8 is related to question 6 and takes a different approach to investigating the influence of respondent characteristics by identifying subgroups of respondent-based TTO responses from a 10-year TTO exercise using latent class analyses, rather than through econometric equations. After latent classes are identified, respondent characteristics associated with membership of the identified subgroups are explored.

The last chapter, Chapter 9, integrates the results of the previous chapters. It provides a discussion of the results and explores the implications and limitations of this thesis.

To note, this thesis consists of published or submitted research articles in scientific peer-review journals. As a result the chapters can be read independently and some of the chapters may overlap in content.

2.

What should we know about the person behind a TTO?

Based on: F.E. van Nooten, J. J. V. Busschbach,
M. van Agthoven, J. van Exel and W.B.F. Brouwer

What should we know about the person behind a TTO?

European Journal of Health Economics. 2018; 19(9):1207-1211
DOI 10.1007/s10198-018-0975-1.

2.1 What should we know about the person behind a TTO?

The Time Trade-Off (TTO) method is widely used to obtain quality of life scores for health states, but its validity is not uncontested, for instance because TTO values for similar health states can differ substantially between studies (Arnesen & Trommald, 2005). This could be caused by differences between studies in how the procedure is applied, but could also relate to differences in study samples.

In a TTO, respondents are asked to value health states by making trade-offs between length and quality of life within a limited remaining lifespan. This is a cognitively demanding task, which for some respondents may be more difficult to complete than for others. Moreover, despite the hypothetical nature of the exercise, having to imagine giving up life years from a finite amount of remaining life years to increase quality of life may invoke different emotions in respondents. This may be related to different expectations about their own remaining life expectancy or to different attitudes towards and experiences with this existentialistic question. These different emotional responses to the exercise may lead to answers that do not necessarily relate to the actual value of the health state presented to respondents, which TTO obviously hopes to elicit. We argue that in order to come to a better understanding of variations in responses to TTO questions, more systematic attention is required for the persons behind TTO responses.

TTO

A TTO question typically asks respondents to imagine being in a particular imperfect health state for a certain period of time, say ten years, after which they will die. Respondents are then asked to consider an alternative scenario in which they are in full health but live for a shorter period of time. Subsequently, respondents are requested to indicate how many years living in full health for them would be equivalent to living ten years in the imperfect health state. This point of indifference can be used to compute a value for the imperfect health state relative to full health. In order to do so, the scale on which health utilities are measured, is usually normalized by setting the value of the state ‘dead’ equal to 0 and that of ‘full health’ equal to 1. Then, if a respondent indicates for example that living for 10 years with moderate pain and severe problems with mobility (with value β) is equivalent to living 8.5 years in full health (with value 1), the value of the imperfect health state (β) for this respondent can be computed as follows. Ten years in β equals 8.5 years in full health, can be written as $10 * \beta = 8.5 * 1$ or $\beta = 8.5/10 = 0.85$. This can be formalised as follows:

$$T * v(X) = P*v() + (T-P)*v(D) \quad (1)$$

where v(FH) denotes the value of full health, v(X) the value of imperfect health state X and v(D) the value of being dead. A respondent is then asked to reveal the number of years P (with P<T) in full health (after which death follows), that this person considers equal to T years in the imperfect state X. Given that v(FH) is set to 1 and v(D) to 0, equation (1) reduces to

$$T*v(X) = P \quad (2)$$

Abstracting from issues such as discounting (Attema & Brouwer, 2010), for the sake of simplicity, the value of imperfect health state X can then be computed by dividing the years in full health (P) by the number of years in imperfect health (T), thus:

$$v(X) = \frac{P}{T} \quad (3)$$

The remaining life years in full health T presented to respondents varies between studies. Common time frames include 10 years or the subject's expected remaining life expectancy (Devlin *et al.*, 2011; Woloshin *et al.*, 2001;Z arate *et al*, 2008). Equation (3) suggests that this variation in subject's expected remaining life expectancy does not affect the value of the health state, as long as people trade-off life time proportionally for different periods of time (Torrance, 1976). Although this makes sense, it depends on several assumptions, including, as indicated, constant proportional trade-offs and for instance the absence of discounting. Since the introduction of TTO (Torrance, 1976) and its initial development, a considerable body of evidence has emerged showing that these assumptions do not always hold and, consequently, the way TTO is framed may influence the value given to health states(Arnesen & Trommald, 2005; Bleichrodt, 2002a; Doctor *et al.*, 2010; van Osch *et al.*, 2004) . Therefore, differences in TTO values may arise from lack of standardisation of the TTO procedure.

However, variation may also occur when respondent samples included in studies differ in characteristics that are relevant to how respondents react to the type of question or to the actual valuation of health states. An important distinction that needs to made in this context is the difference between valuing one's own health state or a hypothetical health state. It has been shown repeatedly that direct TTO values (i.e. people valuing their own health state) are considerably higher than indirect TTO values (i.e. people valuing hypothetical health states) (de Wit *et al.*, 2000). One reason for this is that own health state valuations, in contrast to those of hypothetical health states, may be influenced by coping (Stolk & van Nooten, 2005; Versteegh & Brouwer, 2016). Here, we focus on indirect TTO valuations, as these are most often used as input for utility questionnaires like the HUI, EQ-5D and SF-6D and commonly used as source for national tariffs.

Next, we first highlight some elements of what we already know from literature about influences of person-related variables such as demographic and attitudinal characteristics on TTO responses, without trying to be exhaustive. Then, we argue that more systematic attention to the persons behind TTO responses is required for a better understanding of variations in their responses to TTO questions.

What do we know?

The influence of several demographic characteristics and attitudinal variables of respondents have been regularly reported in TTO studies, such as age, gender, marital status, having children, health status, education level, socio-economic status, ethnicity, and religious beliefs. Below we highlight some of the findings regarding the influence of these variable on TTO values.

The evidence regarding the influence of age on TTO scores is mixed. Some studies observed no statistically significantly effects, although these typically used lower numbers of respondents or had less variance in age (Augestad *et al.*, 2013; Devlin *et al.*, 2011). Some large indirect TTO valuation studies investigated the relation between age and TTO. They found a non-linear relationship when using a fixed time frame of 10 years: generally valuations increase slowly up to the age of about 45, fall slowly between age 45 and 70, and then more sharply for older ages (Dolan *et al.*, 1996a; Dolan, 2000, Dolan & Roberts, 2002). In terms of effect size, the influence of age mostly was relatively small (Dolan *et al.*, 1996a).

For gender, most studies find no significant influence on TTO responses (Augestad *et al.*, 2013). In studies that report a significant gender effect, women tend to give up less life years and thus value health states higher than men (Bernert et al., 2009; Dolan *et al.*, 1996a; Dolan & Roberts, 2002).

Marital status, or living together with a partner (van Nooten *et al.*, 2015), and having children (van Nooten *et al.*, 2015), could influence TTO scores if these aspects would affect preferences for length over quality of life. Krol et al. (Krol *et al*., 2016) argued that two opposite effects may influence TTO scores in this context and provided evidence of their existence. On the one hand, the idea of leaving your loved ones behind can have a negative effect on the willingness to give up life-years. On the other hand, people may not wish to be a burden to their loved ones due to the illness and therefore can be willing to give up more life-years for a better quality of life in the years lived. In empirical work, this has mostly been investigated by looking at marital status and its association with TTO scores. The evidence here generally shows no relation with TTO scores (e.g. Bhatnagar *et al*., 2009), but sometimes a positive relation (Dolan & Roberts, 2002). Krol et al. (Krol *et al*., 2016) indicate that the two opposite influences may, on average, cancel out. The scarce evidence on the effect of having children on TTO valuations suggests a negative effect on the willingness to give up life-years and thus a positive influence on TTO scores, in particular among mothers (van der Pol *et al.*, 2007).

Own health status could be expected to influence TTO scores as well, as it may induce forms of coping and adaptation and shifts of reference points. Studies investigating the relation between (self-reported) health and TTO scores find both insignificant (Krol *et al.*, 2009) as well as significant effects (Ayalon & King-Kallimanis, 2010). Interestingly, Dolan (Dolan, 2011) showed that, in addition to the potential influence the current health state of a respondent can have on TTO scores, negative thoughts respondents have about their future health, can also influence TTO scores. Respondents gave up more years in a TTO exercise to reduce such negative thoughts (Dolan, 2011).

The evidence about the influence of socio-economic status on TTO responses is also mixed. For education level, studies report either no, positive or negative effects (Devlin *et al*., 2011; Rowen *et al.*, 2011. The same mixed results are observed for income (Devlin *et al*., 2011; Guest *et al*., 2013).

In addition, it is interesting to note that, without instructions, respondents generally do not take the financial consequences of ill health into account in TTO exercises, whereas mentioning loss of income because of illness explicitly only has a minor influence on TTO scores (Meltzer *et al*., 1999; Krol *et al*., 2009; Richardson *et al.*, 2009; Tilling *et al*., 2012).

Ethnicity generally does not appear to influence TTO scores (van Nooten *et al*., 2016), although some studies report a, typically small, effect (Izadpanah *et al*., 2013).

Religion and religious beliefs could influence TTO scores not only through an influence on preferences for length over quality of life, but also due to the nature of the TTO exercise. One might expect some religious convictions to be connected with both a reluctance in engaging in a trade-off 'actively' shortening remaining life span to increase quality of life and attitudinal characteristics like fear of death (van Nooten *et al*., 2016). The scarce evidence for religion suggests no statistically significant relation with TTO scores (Essink-Bot *et al*., 2007)

Related to the previous point, some studies explored the influence of attitudes towards life and death on TTO scores. This may be relevant since TTO questions often imply an early age of death and involve thinking about giving up life-years. Significant associations have been found, for example, for beliefs regarding life after death (Lamers *et al*., 2006a;; Rutten-van Mölken *et al*., 2009; van Nooten *et al*., 2016), fear of death (van Nooten *et al*., 2016) and attitudes towards euthanasia (Augestad *et al*., 2013; van Nooten *et al*., 2016. Respondents with a strong preference for staying alive 'at all costs' were reluctant to trade off any years (Kirsch *et al*., 2000).

Expectations regarding length and future quality of life were also shown to play a role. For example, when subjective life expectancy exceeded the life duration specified in the TTO question, respondents were less willing to trade-off years - and the higher subjective life expectancy, the lower the number of years traded-off (van Nooten & Brouwer, 2004; van Nooten *et al*., 2009). Expectations about future quality of life were found to be significant in one study (van Nooten & Brouwer, 2004).

These and some other person-related variables including cultural differences (Johnson *et al*., 2005; Knies *et al*., 2009) and numeracy (Woloshin *et al*., 2001), have been explored in the TTO literature, but typically the evidence remains scarce and most influences were studied in isolation, often as 'by-catch' in valuation studies. In other words, there is a lack of studies investigating the influence of person-related variables systematically.

Why would we want to know?

As we have shown above, the effect of person-related variables on TTO scores, or more precisely their association, tends to be small. Nevertheless, understanding these relations better, also in combination, is relevant for a number of reasons. First, knowledge of respondent characteristics that are influential for the outcomes of TTO studies may be helpful in sampling respondents. If TTO is used for generating nationally representative values for health states (often referred to as 'tariffs'), it may be important to sample respondents from the population according to such characteristics if and when relevant and influential. In addition, for comparability of results in time or across samples, it is useful to make the same selections of respondents and/or to correct for the same selection of potentially confounding variables. Moreover, better understanding of how different groups of respondents react to some of the defining features of TTO exercises, like the limited timespan and giving-up life years, may help improve the design and further standardisation of the TTO method and other methods deriving health state valuations. Finally, the influence of respondent characteristics may also differ depending on the severity of the valued health state, which is directly relevant to TTO exercises. Demographic and/or attitudinal variables may exhibit a different influence when dealing with mild health states as compared to severe health states (Krol *et al.*, 2009).

It is also relevant to investigate these respondent characteristics in combination. Let us consider age more closely, for example. Respondents of different ages may have different views about the importance of remaining length of life, the relevance of different dimensions of health for quality of life, or the relation between length and quality of life. Age may also be closely related to a number of other factors potentially influencing willingness to trade between length and quality of life. For instance, age may be associated with having (current or prior) experience with health problems, which may affect perceptions of how undesirable particular health states are. Age is also associated with role. Over time role functioning may shift from learning to being active in paid or unpaid activities, and include functioning as a partner, parent, or grandparent. And last but not least, age is related to health. If own health state influences TTO scores, then this effect, if not otherwise corrected for, may lead to an apparent association between age and TTO scores as well. Capturing all such elements under the umbrella of 'age' of course might not be adequate.

Moreover, while the influence of the separate variables on TTO scores may be small, the joint effect of several factors may be large. In that context, it is good to note that mostly, a large proportion of the observed variation in health state valuations or TTO scores remains unexplained, so far.

What's next?

We highlighted some of the current knowledge regarding the relation between TTO values of hypothetical health states and a variety of respondent characteristics. This evidence is largely a by-catch of valuation studies, fragmented and often mixed. We discussed several reasons why more systematic investigation of this relation is important for improving the design of TTO exercises and our understanding of the outcomes. Arnesen and Trommald (Arnesen & Trommald, 2005) already concluded that two-thirds of the studies in their review did not present or discuss the influence of even basic respondent characteristics such as age and gender. The focus in TTO research thus seems to be primarily on mean sample results, and much less on the heterogeneity in values or associations between values and sample characteristics.

Considering the importance of reliable and valid estimates of quality of life for research and policy in the health care sector, and the prominent role of TTO in generating such values, more attention for standardisation in design and application of TTO and understanding of what drives TTO answers is warranted. This is also true for other health state valuation techniques such as the standard gamble and discrete choice experiments. Systematic analysis of the influence of respondent characteristics on the interpretation of TTO questions, the willingness to trade, and the number of years traded is an essential part of this process of standardisation. Improved knowledge regarding which factors influence TTO scores, can also inform sampling procedures.

It is clear that, to date, we insufficiently understand what influences TTO scores. Therefore, we argue that there is more that we need to know about the person behind the TTO.

3.

The influence of subjective life expectancy on health state valuations using a 10 year TTO

Based on: F. E. van Nooten, X. Koolman and W. B. F. Brouwer

The influence of subjective life expectancy on health state valuations using a 10 year TTO

Health Economics 2009; 18:549-559
DOI 10.1002/hec.1385

Summary

Aim: To investigate if subjective life expectancy (SLE) impacts the willingness to trade-off (WTT) and the number of years traded-off in a 10-years' time trade-off (TTO) exercise to obtain health state valuations.

Methods: An Internet-based questionnaire was administered in a sample representative for the Dutch general public. Next to basic demographic characteristics and SLE, respondents were asked to perform three TTO exercises. The following EQ-5D health states were included 21211 (TTO1), 22221 (TTO2) and 33312 (TTO3). The WTT was studied using a probit regression model. The number of years traded-off was investigated using a generalized negative binomial regression model. The independent variables used in both models were age, gender, quality of life, education, the difference between age and expected age of death (SLE), and a variable indicating whether the SLE was less than 10 years (SLE < 10).

Results: Three hundred and thirty nine respondents completed the questionnaire. The mean utility scores were 0.96 (TTO1), 0.94 (TTO2) and 0.79 (TTO3). The probit model showed that SLE was the only variable with a significant influence on WTT. The gnbreg showed that the number of years traded-off was also significantly influenced by SLE. In addition, age and education significantly influenced the number of years traded-off.

Conclusion: The WTT years and the number of years traded-off were both influenced by SLE in 10-years TTO exercises. Reducing remaining life expectancy to 10 years in a TTO may thus increase loss aversion and, especially in respondents losing relatively many expected life years, diminish WTT and the amount of time traded off.

3.1 Introduction

The time trade-off (TTO) method is a popular utility measure in the context of economic evaluations and medical decision making. Not only is it frequently used to derive utility scores for health states in certain disease areas, but also important national tariffs for generic quality of life measures such as the EuroQol-5D instrument are based on utility elicitations using the TTO technique (e.g. Dolan, 1997; Lamers *et al.*, 2006). Moreover, while the method has a clear foundation in the formal quality adjusted life years (QALY) model (e.g. Bleichrodt and Gafni, 1996), it is relatively easy to apply. Basically, the TTO method derives the point of indifference between two alternative health streams (e.g. Torrance, 1976). Typically, in one stream the respondent lives for another 10 years in a relatively poor health state while in the other stream the respondent lives in perfect health but for less than 10 years. By eliciting the point of indifference between both the streams – generally by eliciting the number of years in perfect health (n_{ph}) that a respondent considers equivalent to 10 years in the imperfect health state – the utility score for the imperfect health state can easily be derived. Normally, this is simply calculated by dividing n_{ph} by 10. Thus, the TTO is a convenient hypothetical construct used to derive health state valuations with, which has particular features such as fixing the remaining life span (often to an unrealistic 10 years) and operating under certainty (while the real world is often uncertain, which makes risk attitudes of individuals important). This raises fundamental questions about the validity and usability of TTO scores in decision making that largely fall outside the scope of this chapter, but are important to acknowledge.

Unsurprisingly, it has been demonstrated that the TTO, in spite of its applicability and popularity, is not without methodological problems. The method has been shown to be influenced by several types of biases (e.g. Bleichrodt, 2002), which cause systematic distortion of results. There has been quite some attention for these biases in the literature (e.g. Bleichrodt, 2002; Spencer, 2003; van Osch *et al.*, 2004; Attema & Brouwer, 2008). There is also the question how individual traits, such as age and gender influence TTO utilities (e.g. Robinson *et al.*, 1997). By now, there is ample evidence that important respondent characteristics influencing the responses to TTO questions are age, gender and own current health (e.g. Dolan *et al.*, 1996a; Dolan, 2000; Dolan and Roberts, 2002). Other personal characteristics, such as beliefs regarding life after death, have sometimes also been suggested as influencing TTO responses (e.g. Lamers *et al.*, 2006), but the evidence is (much) scarcer there. A recent review of published studies moreover shows a wide and inexplicable variation in results of TTO exercises in different disease areas (Arnesen & Trommald, 2004, 2005). Arnesen & Trommald (2004) moreover point out that in many studies using TTO, there are a high number of non-traders, i.e. there are a high number of respondents that do not wish to trade-off length of life for a quality of life improvement. When these non-traders are, in general, not indifferent between the imperfect health state and perfect health, this means that the TTO does not adequately reflect their preferences. This may be the case, since some studies found that respondents may not willing to trade-off any life years due to religious motives (Buckingham *et al.*, 1996; Green *et al.*, 2000; van Osch *et al.*, 2004). The factors influencing non-trading may be considered as important as the influences on TTO scores in respondents that do trade-off length and quality of life, given the commonness of non-trading (Arnesen & Trommald, 2004). To date, much remains unclear as to what influences TTO scores. Given the widespread use of TTO, more knowledge seems to be required both on what influences the number of life years traded and what influences non-trading.

Recently, a more “implicit” characteristic of respondents was shown to influence TTO utilities, i.e. a respondents’ subjective life expectancy (SLE). Van Nooten & Brouwer (2004) report on an experimental study using a life-time TTO in which respondents were asked to imagine that they would live until age 80 in a suboptimal health state but could regain perfect health by sacrificing life years. In line with theoretical expectations (e.g. Dolan *et al.*, 1996b), respondents who had a higher expected age of death than the projected 80 years traded-off a significantly lower number of years to regain health than those who did not expect to reach the age of 80. The explanation for this is straightforward: the former group of respondents already was “cheated out” of some life years making them reluctant to trade-off additional years, while the latter group received “bonus-years” relative to their expectation, which may be relatively easily traded. While their experiment confirmed this reluctance in the quantity traded off, van Nooten & Brouwer (2004) did not investigate non-trading specifically. Moreover, they used a life-time TTO, whereas a 10-year time frame is more common in TTO exercises (e.g. Chapman *et al.*, 1998; Clarke *et al.*, 1997; Dolan, 1997; Lamers *et al.*, 2006; Witney *et al.*, 2006). Whether SLE also influences TTO exercises using a 10-year time frame has never been investigated to our knowledge. But given the commonness of the 10-year TTO, this is important, also because in such a situation many respondents will experience an immediate (but unequal) loss of life duration relative to their prior expectations, since most respondents would have expected to live longer than 10 years in the general population.

This chapter reports on a new empirical test to investigate the influence of

SLE on responses in a TTO exercise using a 10-year time frame. It looks both at the influence of SLE on the general “willingness to trade” (WTT) in order to explain non-trading and at the number of years traded-off. The structure of the chapter is as follows. Section 3.2 gives some background, Section 3.3 describes the methods, after which Section 3.4 presents the results, and Section 3.5 discusses our findings.

3.2 Background

The TTO method is based on the QALY model. In its most general form, this model assumes that preferences over lifetime utility can be represented by an additive utility function over life duration and health quality (e.g. Attema & Brouwer, 2007):

$$V = \sum_{0}^{T} \delta_t u(h_t) \tag{1}$$

where $u(h_t)$ represents a utility function mapping individual’s preferences over health states at each time point t while δ_t denotes the weight attached to the utility experienced at point t. This generalized QALY model (e.g. Bleichrodt and Gafni, 1996), however, is normally not used in practice. Typical TTO procedures will set δ_t in Equation (1) equal to 1 for each t, (unrealistically) implying that equal weight is attached to all life years (Attema & Brouwer, 2007). This results in the linear QALY model as noted formally in Equation (2):

$$V = \sum_{0}^{T} u(h_t) \tag{2}$$

The TTO method then typically asks respondents to indicate when they are indifferent between a certain period *(T)* in some imperfect health state *IH* followed by death *(D)*, and a shorter period *(k)* in full health *(FH)*, after which again death follows. The elicited indifference can then be represented by the following equation:

$$\sum_{0}^{T} u(\mathrm{IH}) = \sum_{0}^{k} u(\mathrm{FH}) + \sum_{k}^{T} u(D) \tag{3}$$

By normalizing the utility function over health, setting the utility of a period in full health equal to one, *u*(FH) , and a period in the state “death” equal to zero, *u(D)* = 0, it is easy to see that the utility attached to the imperfect health state *u*(IH) can be simply calculated by dividing the years in perfect health *(k)* by the number of years in imperfect health *(T)*, thus:

$$u(\mathrm{IH}) = \frac{k}{T} \qquad (4)$$

We assume here, as was shown in van Nooten & Brouwer (2004) for the case where *T* was set equal to 80 minus current age, that the elicited value of *k* depends on the expected age of death of respondents relative to the projected age of death at time *T*. Therefore, we assume that *k* depends (among other variables, such as the severity of the health state presented) on the SLE, which is defined as expected age of death minus current age. We hypothesize that if SLE >T this makes respondents more reluctant to trade-off years, resulting in more non-traders and fewer years traded by those who do trade. If on the other hand SLE < T, implying that the SLE is less than 10 years, so that people are granted additional life years in the hypothetical situation in comparison with their own expectation, people will be relatively willing to trade-off time.

3.3 Methods

In order to test this hypothesis in the context of a 10-year TTO, an Internet-based questionnaire was developed. The questionnaire was structured as follows. First, gender, age, educational level and current health status (using the EQ-VAS and EQ-5D) were elicited. Second, the questionnaire contained three different TTO exercises. The subjects were asked to imagine being in an imperfect state for 10 years after which death would follow. These health states were described using the EQ-5D dimensions (mobility, self-care, usual activities, pain/discomfort and anxiousness/depression) and levels. The health states presented to the respondents in the three TTOs were respectively some problems with mobility and usual activities, no problems on the other dimensions (21211 in EuroQoL terms), some problems on all dimensions except for pain/depression (22221 in EuroQol terms) and severe problems with mobility, self-care and usual activities, some depression/anxiety but no pain or discomfort (33312 in EuroQol terms). These health states were chosen to present respondents with a broad range of severity of health problems. The alternative of "perfect health" was defined as no problems on all five dimensions (11111) according to the EQ-5D instrument. Finally, in order to assess the SLE respondents were asked to indicate what age they expected to reach.

The TTO part of the questionnaire was set up as follows. The respondents were presented with two options (Appendix B3). The first option (A) was to live for 10 years in an imperfect health state after which death would follow. The second option (B) was living less than 10 years but in perfect health. The respondents could opt for either one of these options by ticking one of two boxes. The first box indicated that they opted for option B and then respondents had to indicate how many years in perfect health they would minimally require in order to prefer option B. This number should specify their point of indifference between the two scenarios. The second box indicated that they did not wish to sacrifice life years and opted for A. Reasons for opting for either A or B were not recorded. To allow subjects to become familiar with the TTO exercise, an example of the TTO exercise was given before the first TTO exercise was performed. The questionnaire was administered by a specialized company to a representative sample of the Dutch general public in terms of age and gender. Respondents who are enlisted with this company were invited to participate in the study.

In order to test our hypothesis we constructed variables from the elicited answers. First of all, k was set equal to the number of years respondents minimally required in perfect health in each of the health states. For ease of interpretation, in our analysis, instead of k we use 10-k as the parameter of interest, i.e. the number of years sacrificed. To evaluate the association between the quantity of years traded-off and the SLE of respondents a generalized negative binomial model (gnbreg) was used. The gnbreg model is appropriate because it allows flexible modelling of the overdispersion. In order to test whether SLE also influences the general "WTT" of respondents a probit model was used, in which the same variables were entered. Data of all questions and all respondents were included in both the models. We computed average marginal effects (AME) for discrete variables.

We estimated confidence intervals for these AME using a bootstrap procedure with 10,000 replications. The confidence intervals are non-parametric, bias corrected and corrected for clustering at the individual level. The explanatory variables used in both the models were age, gender, quality of life, education and the difference between current age and the expected age of death according to the respondents (SLE). Moreover, we included a variable indicating whether the expected age of death was smaller than *T* (i.e.

less than 10 years; SLE10), since in that case, people may be (asymmetrically) more willing to sacrifice life years.

With regard to the representativeness of our sample for the Dutch general public, our study was the representative for the general population in terms of age and gender, but was characterized by a small overrepresentation of respondents with a lower education. However, this finding does not seem to be of major importance for the central aim of this study, as the average SLE did not differ between the education groups.

3.4 Results

In total 339 respondents answered the questionnaire. Characteristics are shown in Table 3.1. The average age was 43 years, the average Visual Analog Scale (VAS) score was 75 and the average expected age of death was 80 years. Figure 3.1 shows the frequencies of the expected age of death across the sample. Figure 3.2 shows the associated SLE across different age categories.

In TTO1 (EQ-5D 21211) 78.2% of the respondents did not want to give up any life years, i.e. were non-traders. The average number of years given up across the entire group was 0.44 (SD 1.22), and 2.03 years (SD 1.92) among those that were not non-traders. In TTO2 (EQ-5D 22221) 71.1% of the respondents were non-traders and the average number of years given up was 0.55 (SD 1.18), and 1.91 years (SD 1.50) among those who were willing to trade. For TTO3 (EQ-5D 33312) these numbers were 23%, 2.08 years (SD 2.36) and 2.71 years (SD 2.36), respectively. In total, 15.9% of respondents were willing to sacrifice life expectancy to avoid all of the dysfunctional states they were presented with, and thus had no health state valuation of 1.00. Of the total group of respondents 21.9% were never willing to trade-off anything in any of the three TTO questions. This group was similar in age and gender to the group that was willing to trade-off years at any of the three TTO questions, however, the share of less educated respondents in this group was significantly higher (70.8% vs 56.2%, $p = 0.030$). The utility values associated with TTO1, TTO2 and TTO3 on the basis of these responses were 0.96 (SD 0.12), 0.94 (SD 0.12) and 0.79 (SD 0.24), respectively.

Table 3.1 Demographics

Variable (n = 339)	
Age	42.89 years (range 18 – 65, SD 13.28)
Gender	50.4% male
EQ-VAS (0 – 100)	75.22 (range 5 –100, SD 23.98)
Mvh_A1	0.7999 (range - 0.18 – 1, SD 0.26)
Education	Lower: 59.3%
	Middle: 32.2%
	Higher: 8.6%
Expected age of death	79.89 years (range 25 – 106, SD 11.14)
SLE	37.01 years (range 1 – 79, SD 16.11)
SLE < 10 years	3.2%

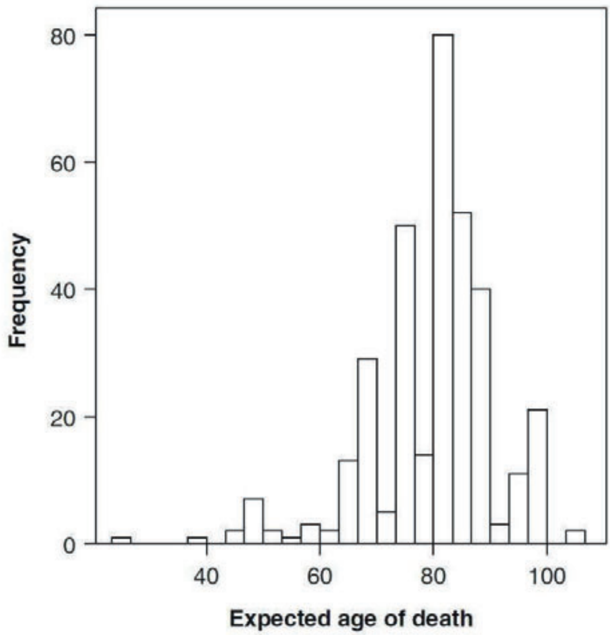


Figure 3.1 Histogram showing expected age of death

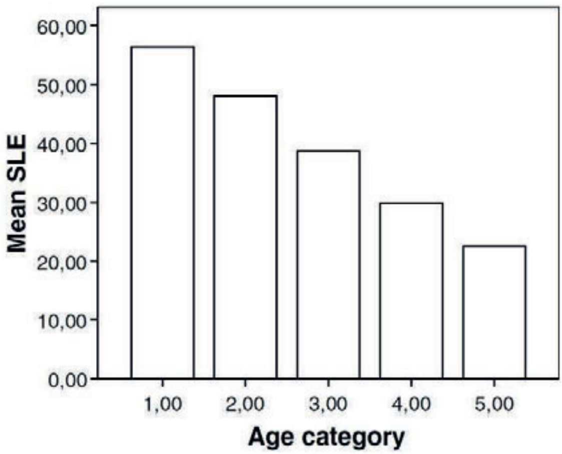


Figure 3.2 The mean SLE across different age categories (1 = 0 – 25, 2 = 26 – 35, 3 = 36 – 45, 4 = 46 – 55, 5 = 56 – 65)

Table 3.2 Results probit model (WTT)

	Average marginal effect	95% Confidence interval[a]	
Age	- 0.00162	- 0.00379	0.00052
Gender	0.01677	- 0.05159	0.08404
EQ-VAS	0.00029	- 0.00075	0.00139
Mvh_a1	0.06361	- 0.02934	0.16190
Edu 2	0.05718	- 0.01972	0.13053
Edu 3	0.11100	- 0.01369	0.23355
SLE	**- 0.00187**	**- 0.00366**	**- 0.00002**
SLE10	0.01209	- 0.00760	0.02940

Education 2 = middle; education 3 = high.

[a]Bias-corrected confidence interval.

In terms of explaining the responses, the probit model showed that the SLE was the only variable with a significant negative influence on the WTT (Table 3.2). In order to test whether SLE was not just a proxy for age, we ran the WTT regression without SLE, with age and age-squared and tested their significance. Both were not significant (p = 0.803 and 0.820).

The gnbreg showed that the number of years the respondents are willing to trade was also significantly negatively influenced by the SLE (see Table 3.3). For each additional year of SLE, the number of years that respondents are willing to give up decreased by approximately 0.01 years. In addition, as often found, age and education significantly influenced the number of years respondents are willing to trade-off. Although the SLE10 was not significant (probably due to small numbers of respondents with SLE < 10), its sign was positive, as expected. We also calculated the gnbreg model including dummy variables for TTO2 and TTO3. The effect of the dummy for TTO3 turned out to be significant next to the other variables that were significant in the first model. The dummy variables for the different TTOs were not correlated to any other independent variable.

3.5 Discussion

This chapter, to our knowledge, presents the first study to investigate the influence of SLE on the numbers of years respondents are willing to trade-off in a 10-year TTO as well as on their WTT any years. While the former issue has been addressed before in a life-time TTO (van Nooten & Brouwer, 2004) but not in a 10-year TTO, the latter issue has not been investigated before in any TTO context.

Table 3.3 Results gnbreg model (number of years traded-off)

	Average marginal effect	95% Confidence interval[a]	
Age	**- 0.01080**	**- 0.01974**	**- 0.00218**
Gender	0.03107	- 0.20439	0.30203
EQ-VAS	0.00278	- 0.00019	0.00607
Mvh_a1	- 0.12602	- 0.47932	0.21336
Edu 2	0.13060	- 0.14241	0.41780
Edu 3	**0.72061**	**0.09090**	**1.48955**
	Average marginal effect	95% Confidence interval[a]	
SLE	**- 0.00916**	**- 0.01693**	**- 0.00144**
SLE10	0.02054	- 0.00760	0.06026

Education 2 = middle; education 3 = high.

[a]Bias-corrected confidence interval.

Our hypothesis that the decision to trade-off life years and the number of years required in perfect health (k) to be indifferent between the two scenarios would depend on SLE was confirmed. The higher the SLE (remaining SLE, i.e. the difference between age and the SLE), the lower the number of years respondents are willing to trade-off. Therefore, it appears that the finding of van Nooten & Brouwer (2004) in a life-time TTO that the SLE influences TTO answers, also holds in more commonly used 10-year TTOs[1]. Furthermore, age and education influenced the number of years respondents are willing to trade-off, which is a common result. The WTT was only significantly influenced (negatively) by the SLE. In other words, those respondents for whom SLE exceeds the maximum number of years of life set by the TTO questions are less willing to trade-off any years of life, which is an important finding given the commonness of the 10-year TTO.

In a 10-year TTO, the influence of SLE on the WTT and the numbers of years traded is probably explained by the fact that respondents felt "cheated" out of life years anyhow. The more years already taken from them, in comparison with the reference point of own SLE, the fewer years people are willing to sacrifice additionally. In that sense, our findings can be linked to the well-known bias in TTO valuation of loss aversion, which refers to the fact that people are more sensitive to losses than to gains when viewed from a particular reference point (e.g. Bleichrodt, 2002, Kahneman & Tversky, 1979). By (unequally) reducing people's life duration relative to their reference point for life duration (i.e. remaining SLE), loss aversion may be (unequally) strengthened. Likewise, if people are given more years than they had expected (the small group of respondents with a SLE of less than 10 years), they would be more willing to give up years in a TTO. In this respect, Dolan *et al.* (1996) suggested that respondents who do not believe that they actually live for 10 more years, might willingly give up these "excess" life years, thereby depressing the apparent value attached to the health states. Although our results indeed indicate this influence, probably due to small numbers of respondents, it did not reach normal levels of statistical significance.

Our study, like that of van Nooten & Brouwer (2004), yielded relatively high utility scores for the health states presented, when for instance compared with MVH_A1 scores (Dolan, 1997). This may have to do with the TTO design we

[1] Note that we used slightly different health states as compared with van Nooten and Brouwer (2004), where EuroQol states 11211, 22222 and 11232 were used. We intentionally opted for a more severe first and third health state to avoid too many non-traders in the first exercise and to maintain a broad spread of health states over the possible health states. Moreover, it appears unlikely that the associations reported here would depend on the precise health states chosen.

used (Appendix B3).[2] The fact that our respondents were relatively reluctant to trade-off live years, also affects the effect-sizes of our regression results. The coefficient indicating the influence of SLE on numbers of years traded of was small, about 0.01 years. However, given the average number of years traded and the variation in SLE, this can still amount to a relevant influence.

In spite of the high utility scores, we are fairly confident that our central question can still be answered, since we did not set out to derive utility scores per se, but rather wished to investigate the influence of life expectancy on trading. Moreover, our results, which fit the variation across TTO studies well (Arnesen & Tromald, 2004), are in terms of background influences largely in line with the literature. For instance, the number of years traded off was significantly influenced by age and also educational level (Table 3.3). For instance Dolan *et al.* (1996b) indicate that experience of illness influences respondents' valuations of health states. Although the probit model did not result in a significant influence of own health on WTT, additional analysis revealed that the EQ-5D health state 11111 was more prevalent among those who were willing to trade-off compared with those who were not willing to trade-off at a 10% confidence level (34.7% vs 47.2% p = 0.063 by chi-square test). Furthermore, an analysis of healthy (EQ-5D health state 11111) vs non-healthy (other EQ-5D health states) respondents showed higher prevalence of healthy people among the higher numbers of years given up for all three health states evaluated.

Regarding the elicited expected ages of death, these appear on average reasonable. This indicates that it is feasible to directly obtain these expectations from the general public, without extensive instructions or guidance. We based our approach on that of Brouwer & van Exel (2005), who reported on SLE in the Dutch population, showing that subjective expectations regarding length of life exceed objective expectations by some 4 years. They report an average of 83.2 as expected age of death, while this is just below 80 in this study, which makes the average SLE in this study reasonably consistent with the average Dutch objective life expectancy.

In that context, among our respondents who were unwilling to trade-off anything, the percentage of respondents with lower education was significantly higher than in the group of people who were willing to trade-off, although the mean age and SLE did not differ. In a general population study

[2] Besides the design, the health states used may have had some influence also. The third state for instance, with severe problems in some dimensions but none on pain or discomfort may be difficult to envisage.

of 3395 respondents (Dolan *et al.*, 1996a), an overrepresentation of lower educated people among the non-traders was also reported. Dolan *et al.* (1996a) suggest that less educated respondents might be more likely to suffer from a so-called "status quo effect",[3] or they may simply have lower numeracy making it more difficult, for instance, for them to indicate the WTT a fraction of a year. In that context, it is interesting to note that for instance Woloshin *et al.* (2001) showed that limited numeracy may be an important barrier to meaningfully eliciting patients' values using the standard gamble and TTO techniques.

Although the influence of many factors on TTO derived utility values have been suggested in the literature, the influence of only a few of them (e.g. age, gender and current health status) has been repeatedly supported by evidence from empirical studies (Dolan *et al.*, 1996a). It seems that SLE has the potential of being a fourth such factor, given these and previous findings. Our analysis indicates the importance of including a question on the respondents' SLE in future TTO exercises and to give account for it in the analysis in the calculation of the utility scores from the TTO responses, both in a lifetime TTO as well as in TTO exercises with shorter durations.

Concluding, this chapter reports on the second empirical study on the influence of SLE on TTO and the first to study this influence in the context of a 10-year TTO. The results confirm that SLE has an influence on both the WTT and the numbers of years traded in a 10-year TTO. Although it has been suggested that any analysis of TTO data is likely to tell only a small part of the total story of what "lies behind the numbers" (Robinson *et al.*, 1997), since a considerable part of the variation of respondent valuations will remain unaccounted for (Dolan and Roberts, 2002); SLE may be one measurable variable that accounts for some additional variation. Furthermore this study has stressed the importance of WTT for TTO valuations. In studies based on TTO valuations it is often mentioned that because of unwillingness to trade-off anything a certain percentage of the respondents was removed from the analysis, without giving account for the underlying mechanism of this phenomenon or the preferences of respondents that are unwilling to trade. Given the widespread use of TTOs in the context of economic evaluations and medical decision making, more knowledge on why respondents act the way they do in TTO exercises is warranted.

[3] That is, these respondents focus relatively strongly on movements away from their current position.

Whether TTO responses are adequately sensitive, robust and representative of preferences to be trusted in policy making obviously is a matter for debate. The hypothetical nature of the TTO with its certain and fixed (and often unrealistically short) life span, its somewhat counterintuitive notion of living shorter but being healthy versus living longer but being unhealthy, might lead people to feel that the method itself is more problematic than any bias to which it may be susceptible. Given the popularity of the TTO, this remains an important issue to address in future research. Indeed, if we fail to understand what the TTO method exactly measures or how it is influenced by systematic biases and personal traits, we may need to start worrying about the life expectancy of the TTO method itself.

Acknowledgement

We are grateful to the attendees of the iHEA Conference 2007 in Copenhagen, Arthur Attema and two anonymous reviewers for their useful comments on earlier versions of this chapter. The usual disclaimer applies.

Appendix Chapter 3: A3

The influence of subjective expectations about length and quality of life on time trade-off answers

Based on: F.E. van Nooten and W.B.F. Brouwer

The influence of subjective expectations about length and quality of life on time trade-off answers

Health Economics 2004; 13:819 – 823
DOI 10.1002/hec.873

Summary

When answering TTO questions respondents sometimes have to imagine being in a certain health state during their remaining lifespan, often based on objective life tables. Respondents however may have subjective expectations about length and quality of life that differ from the objective ones. If respondents do not fully abstract from own expectations, TTO scores may be biased. In this note, we indicate how subjective expectations could influence TTO scores and present some empirical findings suggesting that they do. Our results indicate that subjective expectations may serve as unobserved reference points and as such influence TTO responses.

A3.1 Introduction

Economic evaluation has become an important aid for decision makers in deciding on reimbursement and implementation of various health care programmes. Often, such evaluations take the form of a cost-utility analysis, in which health effects are normally measured in terms of QALYs. Several methods exist which derive QALY health state valuations, of which time trade-off (TTO) is an important example. Normally in TTOs respondents are asked to sacrifice life years in order to improve their quality of life. The utility of being in state x can be derived from the number of years people are willing to sacrifice to return from x to good health, relative to the total remaining lifespan.

In the literature there has been attention for the fact that the answers people provide in TTOs do not necessarily correctly reflect the relative utility in the different health states. Bleichrodt (2002) recently indicates that there are at least three main reasons why disparities between TTO and standard gamble occur: probability weighting, loss aversion, and scale compatibility. Moreover, there is a large amount of literature on the use of TTO and possible problems in using TTO (Gafni & Torrance, 1984; Johannesson *et al.*, 1994; Martin *et al.*, 2000; Stiggelbout *et al.*, 1994, Stiggelbout *et al.*, 1995).

In this appendix, we discuss a possible additional disturbing influence in TTO – i.e. the effect subjective expectations about the age of death and quality of life may have on TTO answers. Individuals have their own beliefs about their expected age of death and future quality of life, but these expectations often do not coincide with the actuarial expectations (Mirowksy, 1999; Ross & Mirowsky, 2002). When remaining (average) actuarial life expectancy is used in TTO exercises to determine the relevant timeframe, the discrepancy between actuarial and subjective expectations may bias TTO responses. This holds if respondents do not fully abstract from own expectations about length and quality of life of the TTO exercise. The extent to which individuals are capable of this is unclear, and the resulting potential bias has not been investigated though Kattan *et al.* (2001) argue that the classical TTO artificially locks the patients with degenerative or life-threatening conditions into a specific health state. In this way the fear of deterioration, which is the primary fear of a patient with a deteriorating disease, is removed. Their study, in the specific context of using a patient population for health state valuations, suggests that the classical TTO is to be modified to include remaining subjective rather than actuarial life expectancy.

In this appendix, we will first indicate how subjective expectations may influence TTO answers. After that, we will present the results from a first empirical test (n = 105) of this influence of subjective expectations on TTO answers.

A3.2 Subjective expectations and TTO answers

It has been shown that people's expectations about longevity can differ from their actuarial life expectancy. Ross & Mirowsky (2002) reports that males expect to live about 3 years longer than the actuarial estimate and Afro-Americans expect to live about 6 years longer. This effect remains after a correction for socio-economic status and the signs and symptoms of good health. Moreover, Brouwer & van Exel (2005) report a considerable underestimation of expected future quality of life in the general public. Such discrepancies may lead to biases in TTO answers.

In TTO exercises, respondents are typically asked to imagine to be in a hypothetical health state for some specified period. However, it seems likely that respondents relate this hypothetical state to own expectations about length and quality of life. These subjective expectations may serve as an implicit reference point in the TTO exercise. For instance, when the actuarial estimate of age of death is used as to derive the relevant time frame, this may result in a longer or shorter time frame compared with the situation in which subjectively expected age of death had been used. If a respondent, expecting to reach the age of 85, receives an actuarial estimate of 80 years in a TTO exercise, this difference may be experienced as a loss of 5 years in relation to his reference point (85). This respondent may subsequently be more reluctant to trade-off "even more" years in the TTO exercise, as he already experiences a loss of 5 years. Alternatively, a respondent expecting to become 75, experiences a gain of 5 years and may trade these "additional" 5 years relatively easy (because they were a bonus to begin with). Similarly, the subjective expectations about future quality of life may also influence the responses given in a TTO exercise. Although individuals should normally assume a stable health status x, it may be, when they are asked to sacrifice life years in a TTO exercise, that the answers are influenced by the individuals' expectations about the quality of the sacrificed life years. Years that expectedly are lived in poor health may be more easily sacrificed than those years expectedly lived in (nearly) perfect health. Own beliefs may again serve as an implicit reference point, influencing TTO results if abstraction from these expectations is difficult.

Thus respondents' subjective expectations about the age of death and quality of life could influence TTO scores, if people are not capable of fully abstracting from these expectations. Of course, this is especially relevant when the trade-offs concern a timeframe that is set on the basis of some average or actuarial life expectancy. In many instances, TTO exercises are constructed with a fixed timeframe (often 10 years) irrespective of life expectancy or the age of the respondent. In such cases, the influence of own expectations may be less prominent, but still one can imagine that people may relate the exercise to prior beliefs they have on life expectancy and future quality of life. For instance, younger respondents may be more reluctant to give up years than older ones, as their remaining life expectancy has been more drastically shortened relative to their reference point, given the maximum of 10 years they can stay alive. Perhaps even more important, many respondents may be relatively unwilling to sacrifice life years in a 10 year time frame, leading to relatively high QALY scores.

Putting it to the test

In this section, we will briefly discuss the results of a first study to determine whether the responses in TTO exercises vary with the subjective expectations of respondents. A short questionnaire was drawn up for this purpose. In it, respondents were first asked some background questions (i.e. about sex, age and education) and their current health status using the EuroQol descriptive system and the EQ-VAS. Then the respondents were asked to perform three TTOs. For these TTOs, three health states out of 243 possible health states scenarios with the EQ-5D were chosen, with MVH_A1 scores of respectively 0.88 (health state 11 211), 0.58 (health state 22 222) and 0.16 (health state 11 232) to have some variation in the severity of health states and the corresponding valuation (Dolan, 1997). The TTOs were conducted as follows. First, the health state to be valued was presented, assuming that this health state would last until the age of 80 after which immediate death would follow. The age of 80 was chosen for men and women alike to keep the questionnaire simple. Second, the respondents were asked whether they would be willing to sacrifice life years in order to regain perfect health and if so, how many years they would be willing to sacrifice. After the TTO exercises, respondents were asked about their expectations on future quality of life, with the use of the EQ-5D descriptive system, to indicate their expected health profile at the ages of 60, 70, 80 and 90. Finally, respondents were asked what age they expected to reach.

The study sample was drawn from the general public. We did not attempt to get a representative sample as our study was purely explorative. Individuals in our convenience sample were handed the questionnaire for self-completion. In total 115 questionnaires were completed of which 105 were in principle useful for further analysis.

A3.3 Results

Respondents were more often female (58%), young (34.1 years old) and well educated than a representative sample would be. The vast majority of the sample population has a good quality of life.

Table A3.1 Expectations of quality and quantity of life for males and females

Female (n = 71)	Mean	SD	Male (n = 44)	Mean	SD
Age	33.92	13.02	Age	34.36	14.07
Quality of life now (EQ-VAS)	0.91	0.15	Quality of life now	0.94	0.11
Quality of life expectation at 60	0.87	0.18	Quality of life expectation at 60	0.88	0.17
Quality of life expectation at 70	0.72	0.16	Quality of life expectation at 70	0.75	0.15
Quality of life expectation at 80	0.52	0.31	Quality of life expectation at 80	0.46	0.32
Quality of life expectation at 90	0.17	0.41	Quality of life expectation at 90	0.14	0.44
Expected age of death	81.30	7.26	Expected age of death	81.71	6.49

Table A3.1 shows the subjective health expectations and current age and health of the respondents, for males and females separately. Table A3.2 displays the most frequently expected health profiles at the ages of 60, 70, 80 and 90, using the EQ-5D descriptive system.

Our sample was relatively unwilling to trade-off years for improved quality of life. Moreover, not all respondents distinguished between the severities of the three different health states in the TTO exercises. This may reflect that TTO exercises should preferably be introduced orally, which was not possible in this study. On average the TTO scores for the presented health states were respectively 0.99; 0.93 and 0.88, which is significantly higher than the corresponding MVH_A1 values. Though this makes the TTO results rather non-specific, it still is interesting to see whether or not the TTO answers were influenced by the subjective expectations of the respondents. We did not find a correlation between the answers of the first TTO question (health state 1) and expected age of death. However, we found significant (p < 0:01) correlations of 0.435 and 0.302, respectively, between expected age of death and the answers of the second and third TTO question.

Ordinary least squares regression analyses were used to assess the impact of subjectivity on the three health state valuations. Age and sex as well as education are added to the model as it has been shown before that they affect TTO valuations (Dolan et al., 1996). Table A3.3 shows the results.

These results indicate that the expected age of death is a statistically significant explanatory variable in the second and third regression. The sign shows that a higher expected age of death leads to a higher TTO score. Put differently, when people expected to reach a higher age, they were less willing to sacrifice life years in the context of a given maximum age of 80. This confirms our hypothesis. We also expected the expected quality of life to have an effect, especially those pertaining to the ages of 70 and 80, as that is the period of life in which these trade-off take place (i.e. respondents sacrifice years from 80 downwards). The expected quality of life at the age of 60 was only relevant when giving up many years (which our respondents were reluctant to do anyhow) and the expected quality of life at 90 is in principle not relevant as the indicated maximum age in the TTO exercises was 80. The quality of life expectation at the age of 70 is a significant explanatory variable, but only in the third regression. There it has the expected sign – when people believed the quality of life to be higher, they were more reluctant to trade-off such healthy years, leading to higher TTO values. (As an aside, that the adjusted R2s in Table A3.2 are so low is not in itself a cause for much concern since the object of this analysis is to assess the relative effect of quantity and quality of life expectations rather than to find the model(s), which include all the variance in valuations. Moreover, if age squared was used instead of age in the regression analysis no substantial changes occurred).

A final indication of the influence of subjectively expected age of death is denoted in Table A3.4. We hypothesized, that respondents with a higher expected age of death than the indicated age of 80 would be less willing to sacrifice additional life years, while those respondents with an expectancy equal to or lower than 80 would be more willing to do so. This hypothesis was confirmed, as all differences between the two groups in A2.3 were significant (Mann-Whitney test, P < 0.01).

Table A3.2 Most frequent chosen health states based on the EQ-5D at the ages of 60, 70, 80 and 90

	Health state	Number of different Health states
60 years of age	11111	16
70 years of age	11221	22
80 years of age	22221	30
90 years of age	22221	37

Table A3.3 Results of regression analyses (n=105, standard deviation in brackets)

	TTO health state 1		TTO health state 2		TTO health state 3	
Constant	0.948	(0.036)	0.529	(0.107)	0.574	(0.142)
Gender[a]	- 0.009	(0.005)*	- 0.019	(0.015)	- 0.029	(0.020)
Age	0.000	(0.000)*	- 0.000	(0.001)	- 0.001	(0.001)
Medium education[b]	- 0.005	(0.010)	- 0.032	(0.029)	- 0.056	(0.038)
Higher education[c]	- 0.004	(0.010)	- 0.054	(0.028)*	- 0.088	(0.038)**
Quality of life now (EQ-VAS)	0.038	(0.021)*	0.048	(0.061)	0.073	(0.082)
Expected qol at 60	- 0.044	(0.022)**	- 0.015	(0.065)	- 0.212	(0.086)**
Expected qol at 70	0.018	(0.026)	- 0.009	(0.078)	0.180	(0.104)*
Expected qol at 80	0.004	(0.012)	0.009	(0.035)	- 0.005	(0.046)
Expected qol at 90	0.001	(0.008)	0.006	(0.023)	0.005	(0.030)
Expected age of death	0.000	(0.000)	0.005	(0.001)***	0.005	(0.001)***
Adjusted R^2	0.034		0.183		0.188	

*Significant at $p < 0.1$, **significant at $p < 0.05$, ***significant at $p < 0.001$.

[a] 0=female, 1=male.

[b] Medium 0=no, 1=yes.

[c] Higher 0=no, 1=yes

Table A3.4 Life years sacrificed by respondents with a life expectancy above 80 and below 80 (Standard deviation in brackets)

Years of life Sacrificed	Life expectancy ≤80 (n = 52)		Life expectancy ≥80 (n = 53)	
Health state 1*	0.69	(1.53)	0.34	(1.57)
Health state 2*	6.10	(3.73)	2.77	(2.85)
Health state 3*	9.27	(4.93)	5.77	(3.95)

* *Difference significant, p < 0.01*

A3.5 Conclusion

We hypothesized that subjective expectations about length and quality of life might influence TTO answers especially (but not only) when working with timeframes based on average or actuarial life expectancy. People with a higher expected age of death than the proposed age of death may then be more reluctant to give up “additional” years, and people who believe the potentially traded years are of better quality are more reluctant to trade these as well. The latter relation-ship was less prominent in our results, which may have to do with scale compatibility, in the sense that people having to answer in terms of life-years will be more aware of and focused on duration than on quality of life (Bleichrodt, 2002). Our results therefore largely support our theoretical suspicions and indicate that subjective expectations indeed may play a role in TTO exercises.

It goes without saying, that given the limited sample-size and the fact that our respondents valued the different health states unexpectedly high relative to the MVH A1 values, more investigation is needed to confirm these findings. If confirmed, the influence of subjective expectations on health state valuations needs to be assessed further. This note hopes to have indicated that there is reason to believe that the normally unobserved variable of subjective reference points may influence TTO results.

Acknowledgement

We are grateful to Maiwenn Al, Elly Stolk and Job van Exel for their assistance during the study. We thank Han Bleichrodt and two anonymous reviewers for his useful comments on an earlier version of this paper. The usual disclaimer applies.

Appendix Chapter 3: B3

You can chose between option A and option B. If you chose option A you will live another 10 years in a stable but imperfect health state, after which you will die. If you opt for option B, you will live for less than 10 years, but in perfect health, after which you will die.

Option A	**Option B**
Live another 10 years in the following health state: • You have some problems with walking • You have no problems with washing or dressing yourself • You have some problems performing your daily activities • You experience no pain or discomfort • You are not anxious or depressed	Live … more years in perfect health

How many years in perfect health would option B have to entail minimally for you to opt for option B? Note that you can only opt for less than 10 years.

☐ I opt for option B if I minimally would live … years in perfect health, otherwise I would rather opt for A.
☐ I do not wish to sacrifice life years and opt for option A.

4.

Thirty down, only ten to go?!

Awareness and influence of a 10-year time frame in TTO

Based on F. E. van Nooten, X. Koolman, J. J. V. Busschbach and W. B. F. Brouwer

Thirty down, only ten to go?!
Awareness and influence of a 10-year time frame in TTO

Quality of Life Research 2014; 23:377 – 384
DOI 10.1007/s11136-013-0495-5

Summary

Background: Time trade-off (TTO) exercises typically present respondents with a limited time horizon, for example 10 years, thus implicitly considerably reducing remaining life expectancy for the average respondent. It is unclear how this affects health state valuations.

Aim: The aim of the study is to investigate how awareness of the reduced life span implied by a 10-year TTO affects health state valuations, using an experimental design.

Methods: Two Web-based questionnaires (Q1 and Q2) were administered in a sample representative of the Dutch population. Both questionnaires contained three 10-year TTO exercises valuing three distinct health states, specified using the EQ-5D. Q1 used a TTO instruction not explicitly emphasizing the fact that remaining life expectancy was reduced to 10 years, while in Q2 respondents were explicitly made aware of this fact by emphasizing their implied age of death. Respondents answering Q1 were asked retrospectively whether they had been aware of their reduced life span due to the 10-year TTO. Results In total, 656 respondents completed the questionnaires (Q1: 339 and Q2: 317). The average age of the respondents was 43 years and 51% of respondents were male. The average numbers of years traded off for the respondents of Q1 were for TTO1 0.443, TTO2 0.552, and TTO3 2.083 years. For the respondents of Q2, these averages were lower, i.e., TTO1 0.401 ($p = 0.085$ vs. Q1), TTO2: 0.546 ($p = 0.036$ vs. Q1), and TTO3: 1.467 years ($p = 0.000$ vs. Q1). Fifty-seven percent of respondents in Q1 confirmed that they were aware of the reduced life span. This spontaneous awareness had a limited and mixed influence on results. The generalized negative binomial regression analysis, explaining the time traded off showed that age, subjective life expectancy, and questionnaire Q2 (vs. Q1) were negatively associated with the years traded off, whereas education and worse health states in the TTO exercise had a significant positive impact on the years traded off. The probit model investigating the impact on the willingness to trade showed that age (-), education (+), subjective life expectancy (-), questionnaire Q2 versus Q1 (-), the interaction between Q2 and male gender (+), and worse health states in the TTO exercise (+) had a significant impact on the willingness to trade.

Conclusion: These findings emphasize the importance of expected and implied life expectancy in TTOs.

4.1 Introduction

The time trade-off method (TTO) is a popular method for health state valuations, to be used in economic evaluations of health technologies. Frequently used health state valuations, such as those related to the EQ-5D questionnaire, are indeed obtained via TTO exercises (Dolan, 1997; Lamers *et al*., 2006). A TTO exercise elicits the indifference between living a longer period in an inferior health state and living a shorter period in a superior health state. In a typical TTO, a respondent is asked to indicate her indifference between two health prospects: the first prospect is a fixed period of time in an imperfect health state H_1 while the second consists of a shorter period of time lived in perfect health. The respondent is asked to reveal the number of years (k) required in perfect health to become indifferent between both prospects. The period lived in H_1 can be the total life expectancy of a subject, but also a limited number of years, such as 10 years. For example, the MVH protocols from the EuroQol group use 10 years for valuing EQ-5D health states (Dolan, 1997). Several other studies have used a 10-year time frame as well (Szabo *et al*., 2010; Woloshin *et al*., 2001; Zarate *et al*., 2008).

By normalizing the value of perfect health to 1 and that of the state “dead’’ to 0, the results can be used to attach values to different health states (e.g., Torrance, 1976). For instance, the value of the imperfect health state H_1 is then commonly calculated by dividing k by 10 for a 10-year TTO (since 10 years in state H_1 is valued equally as *k* years in perfect health, which has a value of 1). So, if *k* would be 6, the value of H_1 would be 0.6.

However, although the TTO is a popular method of eliciting health state values or ‘‘utility scores’’, it is certainly not without methodological problems. It has been noted in the literature that several biases influence the values thus obtained by the TTO, like discounting (Attema & Brouwer, 2009) and scale compatibility (Bleichrodt, 2002). An important bias in the context of the current study is loss aversion (Bleichrodt, 2002), referring to the common observation that people give more weight to losses in decision making than to equally sized gains, i.e., prefer to avoid losses to acquiring gains (Kahneman & Tversky, 1979). Given some reference point (e.g., current health), a gain in health would receive less weight in decision making than a similarly sized loss would. The procedure followed to obtain health state valuations can also affect the TTO values (Arnesen & Trommald, 2005; Attema & Brouwer, 2008; Bleichrodt *et al*., 2003). Moreover, several characteristics of respondents, such as own health, age, gender, and education, influence TTO values, although, with the exception of own health, their influence is rather limited compared to elicitation protocol-related differences (see, e.g., Badia *et a*l., 1998; Dolan *et al*., 1997; Dolan, 2000; Souchek *et al*., 2005; de Wit *et al*., 2000).

An interesting feature of a TTO is that it instructs respondents to imagine that their remaining life span equals some fixed period T, for example 10 years, which is in general unlikely to coincide with their own expectation regarding remaining life expectancy, unless the respondent is older or for example has a life-threatening disease. It has been suggested that, consequently, TTO responses are influenced by this subjective remaining life expectancy of respondents (Heintz *et al*., 2013; van Nooten & Brouwer, 2004; van Nooten et al., 2009). This is probably because respondents in TTO exercises have difficulties in abstracting from their own situation. This has, for instance, been shown by van Nooten & Brouwer (2004) using a TTO with a time frame equalling rounded average life expectancy (i.e., 80) minus current age of the respondent. They found that when responders expected to live beyond 80, and thus were experiencing a reduced life span in the TTO relative to subjective expectations, they were less willing to trade-off life years for health improvements than those responders who did not expect to live until the age of 80. They concluded that subjective remaining life expectancy, in spite of the instructions, plays a role in TTO exercises as a *reference point*. Given that gains and losses are normally evaluated from a relevant reference point, falling below this expected life duration may lead to loss aversion, while moving beyond this reference point is evaluated as a gain (thus receiving less weight in decisions). Recently, van Nooten *et al.* (2009) confirmed the influence of expected age of death using a 10-year TTO, showing that the quantity of years experienced as ‘‘lost’’ relative to the subjective expected age of death was negatively correlated with the number of years traded off (k) in the 10-year TTO exercise. Furthermore, they found that the general willingness to trade, reflecting respondents’ willingness to trade-off any amount of time in order to regain perfect health, was also negatively influenced by respondents’ expected age of death relative to the implied age of death in the exercise. Heintz *et al.* (2013) recently investigated the influence of subjective life expectancy on time trade-offs by patients, in the context of diabetic retinopathy. They concluded that the influence of subjective life expectancy on TTO responses was substantial in the con-text of patient valuations.

These findings raise some important questions. A first puzzling question is how respondents exactly consider a time frame offered to them in a TTO exercise when this is shorter (or longer) than their subjective life expectancy. For instance, when a 10-year time frame is offered, do respondents then actually interpret it as limiting their life expectancy to current age plus 10 years or do they rather use it as a more abstract means of trading of years (abstracting from current calendar time or their own phase in life)? In the former case, a person aged 30 is thus con-fronted with a scenario in which he will not live beyond the age of 40. Thirty down, and only ten to go, therefore. In this case, a significant loss is experienced by the respondent and trading off ‘‘even more’’ years in the subsequent TTO exercise will add to that loss. In the latter case, the thirty-year old may be unaware regarding the fact that the 10-year time frame limits life expectancy or may not be inclined to project the 10-year time frame on his current situation. Rather, the respondent may for instance relate the 10-year period offered to the final 10 years of their own expected life span. Such years late in life are expected to be of poorer quality, making trading them less of a sacrifice - even when TTO instructions suggest otherwise (Brouwer & van Exel, 2005; van Nooten & Brouwer, 2004). Furthermore, the influence of time preference will then make trading easier (e.g., Attema & Brouwer, 2009). Moreover, such a respondent will not experience the 10-year period as a (severe) shortening of remaining life expectancy, and therefore, a smaller loss is projected, again expectedly increasing the willingness to trade-off life years in the TTO. These three aspects would be expected to result in more life years traded in people not projecting the 10-year time frame onto their own, current situation. Important in this context is that TTO exercises probably differ in the way the 10-year time frame is emphasized. For instance, the MVH protocol does this explicitly and even several times, but this need not be the case in other studies and protocols. Such differences could lead to differences in TTO scores and limit in the comparability of TTO scores.

The primary aim of this chapter is to investigate the effect on TTO scores of explicating the life span limitation implied by using a 10-year time frame in the TTO exercise. Moreover, we investigate whether respondents in TTOs are aware of the reduced life expectancy implied by a 10-year time frame, when this is not explicated. Our hypothesis is that respondents who are (made) aware of the reduced life span, thus projecting the 10-year time period on their own situation, will be less likely to trade-off years in the TTO and will trade-off fewer years than respondents who are not (made) aware of this reduction in life span.

4.2 Methods

Questionnaires

In order to test our hypotheses, we designed two Web-based questionnaires, Q1 and Q2, which were administered to representative samples of the Dutch general public in terms of age and gender, through a specialized Web survey company (Survey Sampling International). Respondents enlisted with this specialized Web survey company were invited to participate in the study. Respondents were assigned randomly to one of the two questionnaires. The results of Q1 were published previously (van Nooten *et al*., 2009) but not the results of Q2 nor the here central comparison between Q1 and Q2.

The two questionnaires were identical in all respects except for how they introduced the 10-year timeframe in the TTO to the respondents, highlighted below. In both questionnaires, the participants were instructed that they could choose between option A and option B in three distinct TTO exercises. In option A, they would live 10 more years in a stable but not perfect health state, which they would die. In option B, they would live less than 10 more years but in perfect health, after which they would die.

In questionnaire 1 (Q1) *after* completing of all three TTO exercises (not specifically alluding to the age of death), respondents were asked whether they had taken into account what their expected age of death would be in the TTO scenario presented and thus how old they would be in 10-year time. These questions were posed after the three TTO exercises in order to avoid contaminating TTO responses by focusing the attention of respondents on this issue. In questionnaire 2 (Q2), the respondents were explicitly made aware of the new age of death implied by the 10-year time horizon. This was done by having the respondent indicate before each of the 3 TTO exercises at what age they would die if they had only 10 more years to live. After that, they followed the exact same TTO procedure as in Q1.

The questionnaires first asked general background information from the respondents: age, gender, current health status (using VAS and the EQ-5D), and education. In both questionnaires, respondents had to value three health states using three TTO exercises. These health states were described using the EQ-5D dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and levels (no/some/severe problems). The three health states included in the questionnaires were 21211 (TTO1), 22221

(TTO2), and 33312 (TTO3)[1] in terms of the EQ-5D system. These imperfect health states were assumed to form a broad range across health states and were also used in van Nooten et al. (2009). The reference state ‘‘perfect health’’ was presented as 11111, i.e., no problems in any the five EuroQoL dimensions. Note that we did not in a protocol for states perceived as being worse than dead since we felt this would be too complex in the context of self-completed questionnaires.

At the time of the study, the standardized EuroQol Webbased TTO was not yet available. Hence, we used a similar design as a previous study using the same inclusion strategy (Krol *et al*., 2006). The survey provided respondents first with an example of how the TTO works to acquaint them with the exercise. An example of the exact TTO question provided to the participants is provided in Appendix A4. After presenting the TTO, respondents were asked whether they were willing to trade-off any years at all. Reasons for non-trading were not recorded. Next, in order to elicit the point of indifference between the two health streams, respondents were asked to indicate how many years they would minimally require in ‘‘perfect health’’ in order to be indifferent between the two streams. Finally, in order to assess the subjective life expectancy, respondents were asked how old they expected to become in ‘‘real life’’ (See Appendix B4 for questions).

4.3 Data Analyses

In order to test our hypotheses, we constructed variables from the elicited answers. First of all, k was set equal to the number of years respondents minimally required in perfect health in each of the health states. For ease of interpretation, in our analysis, instead of k, we use 10-k as explained variable, i.e., the number of years sacrificed out of 10 in order to regain perfect health.

We subsequently focused on the differences between the two questionnaires, investigating both the quantity of time traded in both questionnaires as well as the general ‘‘willingness to trade’’ (WTT). The WTT is a recode of the number of years traded off, operationalized as 0 when

1 These health states are described as follows: 21211: I have some problems walking about, I have no problems with self-care, I have some problems with performing my usual activities, I have no pain or discomfort, I am not anxious or depressed; 22221: I have some problems walking about, I have some problems with self-care, I have some problems with performing my usual activities, I have moderate pain or discomfort, I am not anxious or depressed; 33312: I am confined to bed, I am unable to wash or dress myself, I am unable to perform my usual activities, I have no pain or discomfort, I am moderately anxious or depressed.

respondents were not willing to trade-off any life years and 1 when respondents were willing to trade-off at least some lifetime in order to regain perfect health.

We moreover investigated the influence of type of questionnaire on willingness to trade and number of years traded off in two multivariate analyses, controlling for other influences. For these analyses, we combined the data from all TTOs and all respondents. For analysing the influence of questionnaire type on the quantity of years traded off, we used a generalized negative binomial regression model (GNB), including all responses, also those from respondents who did not trade any years. The GNB model is appropriate because of its flexibility in modelling overdispersion. For analysing the general WTT of respondents, we used a probit model. After testing different model specifications, the variables entered in the two final models were age, age2, gender, own VAS, education level, the difference between age and expected age of death [subjective life expectancy (SLE)], a dummy variable indicating the questionnaire type completed, dummies for the TTO question (TTO1, TTO2, and TTO3), and the interaction between questionnaire type and gender (to investigate whether men respond differently to Q2 than women do, e.g., van der Pol & Shiell, 2007). Using a dummy variable in both models, we furthermore investigated whether the number of years traded off and WTT differed between respondents in Q1 who ‘‘spontaneously’’ interpreted the TTO exercise as limiting their life expectancy to 10 years and those in Q1 who did not interpret the TTO exercise in that way. Significance levels were set at 0.05 and confidence interval at 95%.

We computed average marginal effects for the different variables and estimated confidence intervals for these average marginal effects using a bootstrap procedure with 10,000 replications. The confidence intervals are non-parametric, bias-corrected, and account for clustering at the individual level.

4.4 Results

Questionnaires Q1 and Q2 were answered by 339 respondents and 317 respondents, respectively. The average age of the total population was 43.3 years, 51% was male and the average SLE was 36.8 years. Background characteristics of the respondents for Q1 and Q2 are shown in Table 4.1. There were no significant differences in age, gender, education, and SLE between the respondents from Q1 and Q2. However, the respondents answering Q2 reported a slightly lower current quality of life, measured by the VAS (75 and 71 for Q1 and Q2, respectively, $p = 0.023$).

On average, pooling the data from the two questionnaires, the years traded off were 0.42 (standard deviation (SD) 1.26), 0.55 (SD 1.32), and 1.79 (SD 2.33) for TTO1, TTO2, and TTO 3, respectively. Table 4.2 presents the responses to the TTO exercises in Q1 and Q2. The number of years traded off by the respondents increased with the severity of the health state presented for both Q1 and Q2. The average numbers of years traded off for the respondents of Q1 were for TTO1 0.443 years, TTO2 0.552 years, and TTO3 2.083 years. For the respondents of Q2, these figures were for TTO1 0.401 years ($p = 0.085$ vs. Q1), TTO2 0.546 years ($p = 0.036$ vs. Q1), and TTO3 1.467 years ($p = 0.000$).

In terms of willingness to trade, in Q1, 78.2% of respondents was unwilling to trade-off any years in TTO1, while this percentage was 71.1% and 23.0% for TTO2 and TTO3, respectively. Overall, 21.2% were non-traders (i.e., never trading off lifetime to regain health). For Q2, these percentages were 83.9% (TTO1; $p = 0.073$ vs. Q1), 79.5% (TTO2; $p = 0.015$ vs. Q1), and 42.0% (TTO3; $p = 0.000$ vs. Q1). Overall, in Q2, 42% of respondents were unwilling to trade-off lifetime in any of the three TTO exercises. In general, Table 4.2 suggests an influence of questionnaire type on TTO responses.

Multivariate analysis for number of years traded off

The GNB model shows that education (positively), SLE (negatively), age (negatively), the type questionnaire (Q2 vs. Q1; negatively), TTO question 2 (versus TTO1; positively), and TTO question 3 (versus TTO1; positively) influenced the number of years traded off (Table 4.3). The variable age^2 was insignificant and therefore dropped from the model. The influence of respondent's own quality of life measured by VAS as well as gender did not reach statistical significance of $p < 0.05$ in this model. The fact that questionnaire type had a significant and relatively substantial influence on our results, also after controlling for other influences, confirmed our hypothesis.

Table 4.1 Sample characteristics

Questionnaire	**Q1**	**Q2**	**Q1 vs Q2** ***p-value****
N	339	317	
Variable			
Age (mean (range, SD))	42.89 years (18-65, 13.28)	43.74 years (18-65, 13.85)	0.423
Gender	50.4% male	52.1% male	0.696
EQ-VAS (0-100) (mean (range, SD))	75.22 (5-100, 23.98)	70.98 (1-100, 23.73)	**0.023**
EQ-5D scores based on national tariffs (mean (range, SD))	0.80 (-0.18-1, 0.26)	0.78 (-0.05-1, 0.27)	0.382
Education	Lower: 59.3 %	Lower: 63.1 %	0.337
	Middle: 32.2 %	Middle: 28.4 %	0.309
	Higher: 8.6 %	Higher: 8.5%	1.000
Expected age of death (mean (range, SD))	79.89 years (25-106, 11.14)	80.24 years (33-110, 10.76)	0.689
Subjective life expectancy (SLE) (mean (range, SD))	37.01 years (1-79, 16.11)	36.50 year (3-83, 15.51)	0.684

Bold value indicates significance at 0.05

**t* test

Table 4.2 Trade-offs in Q1 and Q2

Questionnaire	Q1	Q2	P**
N	339	317	
Percentage non-traders per TTO			
TTO1	78.2%	83.9%	0.073
TTO2	71.1%	79.5%	**0.015**
TTO3	23.0%	46.7%	**0.000**
Overall non-traders*	21.2%	42.0%	**0.000**
Years traded-off			
TTO1 (mean (SD))	0.443 (1.223)	0.401 (1.293)	0.085
TTO2 (mean (SD))	0.552 (1.182)	0.546 (1.455)	**0.036**
TTO3 (mean (SD))	2.083 (2.362)	1.467 (2.266)	**0.000**

Bold values indicate significance at 0.05

* Overall non-traders are defined as persons not willing to trade off any years in TTO1, TTO2 and TTO3

** Mann-Whitney test

Subgroup in Q1: respondents who were aware versus those who were unaware

In Q1, 145 respondents indicated to have been unaware of the fact that their remaining life span was shortened to 10 years only in the TTO exercise, whereas 194 indicated to have been aware of this fact. This means that 145 out of the total of 339 respondents, or 42.8%, did not interpret the 10-year TTO exercise to imply that they would not live beyond their current age plus 10 years.

Within Q1, the group of respondents who were unaware of the reduced life expectancy was expected to be relatively willing to trade-off life years compared to the group of respondents who were ‘‘spontaneously aware’’. Tested multi-varietal (results not shown), using the same probit and GNB regressions as described above, but with an additional dummy distinguishing spontaneous awareness and unawareness, indicated a significant effect of spontaneous awareness in the expected direction in terms of willingness to trade (probit), but not in the numbers of years traded (GNB). Thus, we cannot firmly accept or reject our hypothesis when it comes to spontaneous awareness of reduced life span.

Table 4.3 Results GNB model (dependent variable: years traded off)

	Average marginal effect	**Bias corrected 95% Confidence interval**	
Age	**- 0.013**	**- 0.019**	**- 0.007**
Gender	0.032	- 0.176	0.267
EQ-VAS	0.001	- 0.001	0.004
Education	**0.061**	**0.031**	**0.089**
SLE	**- 0.011**	**- 0.016**	**- 0.006**
Questionnaire	**- 0.437**	**- 0.755**	**- 0.113**
Questionnaire* gender	0.414	-0.075	0.981
TTO question 2	**0.230**	**0.050**	**0.445**
TTO question 3	**1.520**	**1.267**	**1.801**

Bold values indicate significance at 0.05
Questionnaire: Q1 = 1, *gender :1 = male

Table 4.4 Results probit model (dependent variable: willingness to trade)

	Average marginal effect	**Bias corrected 95% Confidence interval**	
Age	**- 0.001**	**- 0.003**	**- 0.001**
Gender	0.017	- 0.047	0.082
EQ-VAS	0.000	- 0.000	0.001
Education	**0.012**	**0.004**	**0.019**
SLE	**- 0.002**	**- 0.003**	**- 0.001**
Questionnaire	**- 0.171**	**- 0.235**	**- 0.105**
Questionnaire* gender	**0.102**	**0.006**	**0.200**
TTO question 2	**0.060**	**0.029**	**0.093**
TTO question 3	**0.468**	**0.427**	**0.507**

Bold values indicate significance at 0.05
Questionnaire: Q1 = 1, *gender :1 = male

4.5 Discussion

In this study, we investigated whether respondents performing TTO exercises were aware of the consequences of the 10-year time frame on life expectancy and the consequences when emphasizing this. We hypothesized that emphasizing the reduction in life expectancy relative to own expectations, increased loss aversion would result in fewer people trading and fewer years traded. Our results confirmed these hypotheses. This effect was most profound in poorer health states. Over 40% of the respondents in Q1, which did not allude to the projected age of death, indicated not to have been aware of the fact that the TTO exercise using a 10-year framework implies that their remaining life expectancy is shortened. Our results confirm that these subjects were somewhat less willing to trade than those who were aware of the reduction in their life span, but no significant difference was found in the number of years traded off. This indicates differences in interpretation of the TTO exercise between respondents when the instructions are not explicit, potentially leading to differences in the years traded off and willingness to trade. These problems may be especially manifest for poorer health states. Our results moreover confirmed the influence of SLE in TTOs reported previously (Heintz *et al*., 2013; van Nooten & Brouwer, 2004; van Nooten *et al*., 2009).

Our study has some limitations, which need noting. First, the TTO values obtained in our study turned out to be relatively high, also compared to similar studies using the same design and recruitment strategy (Krol *et al*., 2006, Tilling *et al*., 2012) as well as often used international tariffs (e.g., Dolan, 1997; Lamers, *et al*., 2006). The differences between the here reported TTO scores and those in national tariffs may have to do with differences in the TTO protocols used, like the absence of an interviewer in our protocol, Web-based, no administration of values below death, the absence of visual props, and no ‘‘ping-pong’’ protocol. However, our results are comparable to those reported in van Nooten & Brouwer (2004), who used a rather similar survey but one that was administered face to face and by post rather than using the Internet.

A second limitation is that while respondents were randomly assigned to either of the two questionnaires, the groups were not entirely comparable with regard to their current health in terms of indicated VAS scores. While this is relevant for direct comparisons between the two groups, as current health status has been shown to influence the number of years traded off in TTO exercises (Clarke *et al*., 1997; Mrus *et al*., 2006a, Mrus *et al*., 2006b; Stiggelbout *et al*., 1996). Our main results also hold in the multivariate analyses, where the difference in own health state was accounted for.

Despite these limitations, our findings have some interesting implications. First of all, our results imply that for more than 40% of respondents performing TTO exercises in which the age of death is not explicitly alluded to, it is unclear how they use or interpret the time period offered to them in the exercise. They expressed being unaware that a 10-year TTO time horizon often drastically reduced their remaining life span. Therefore, it may well be that these respondents use the time frame in some other way than projecting it to their own current situation. Alternatively, they may consider the 10 years simply as some abstract period of time, not directly relating it to their own current or future situation. This finding raises some questions regarding the interpersonal comparability of TTO responses, given that respondents may interpret the exercise differently, thus adding to the variability of responses. However, the impact of this variability on TTO scores may be limited, especially pronounced for more severe health states, and may also depend on the exact TTO protocol used. Unfortunately, this could not be investigated further in the current study.

Second, explicit emphasis of reduced life span appears to induce relevant differences in TTO scores (through reducing willingness to trade as well as the amount of time traded), whereas the influence of spontaneous awareness regarding this aspect appears limited. This is in line with earlier research examining the (effects of) inclusion of income in TTO exercises by respondents. Krol *et al.* (2006) and Krol *et a*l. (2009) concluded that respondents in TTO exercises did not consistently include income effects, but that spontaneous variability of this inclusion did not translate into different health state valuations, while explicit instructions to include or exclude income effects did have some effect on TTO valuations. It must be noted that framing effects, invoked by including or emphasizing particular elements of a question, in general are likely to affect the TTO values obtained (Robinson & Spencer (2006).

Third, in comparison with Q1, men respond differently to Q2 than women, especially in terms of general willingness to trade. This could be because men focused more on quality of life and women more on quantity of life when faced with a severely reduced life expectancy. Such response modes might be related to the position women have or envisage (e.g., as mothers). Indeed, a study with female respondents who recently became mothers to some extent confirmed this assumption, since these women assigned a higher value to life years relative to quality of life, although, contrary to our results, there appeared to be no reduction in the willingness to trade per se (van der Pol &

Shiell, 2007). This is an interesting and potentially important area for further research.

Concluding our results confirm the influence of SLE in TTO exercises and show the potential influence of (awareness of) the limited life expectancy in TTO exercises. Further research must be done to show the generalizability of our results in relation to other TTO designs, especially those used to derive national tariffs, using face-to-face interviews, visual props, and formal ‘‘ping-pong protocols’’. For the time being, our results suggest that when using a 10-year timeframe, variation exists in how respondents interpret the TTO exercise. Explicit emphasis of the reduced life span may reduce this variation. Whether this improves the validity of the resulting utility scores remains an open question.

Acknowledgement

We are grateful to the attendees of the ECHE Conference 2008 in Rome for useful comments on an earlier version of this paper.

Appendix Chapter 4: A4

TTO in Q1 and Q2

You can chose between option A and option B. In option A you will live 10 years in a stable but not perfect health state, after which you will die. In option B you will live less than 10 years but in perfect health, after which you will die.

Option A	Option B
Live another 10 years in the following health state: • some problems with walking • No problems with selfcare • Some problems with performing your daily activities • no pain or discomfort • Not anxious or depressed	Live … more years in perfect health

We would first like you to indicate which age you would reach if you had 10 years left to live (like in Option A): … years
[This question was only posed in Q2, not in Q1, which was otherwise identical]

You can choose for either option A or option B.
How many years in perfect health would have to be minimally provided in option B for you to choose option B? Please note: you can only indicate a number less than 10 years.

☐ I choose option B if I can live a minimum of … years in perfect health.
☐ I don't want to give up any time and will always choose option A.

Appendix Chapter 4: B4

Question in Q1 to ask about the 10 year awareness

Before, you had to make three choices between living in an imperfect health state for 10 years and living in perfect health for a shorter period. Every time you were asked to consider that after the maximum period of ten years you would die. Such instructions are often used in this type of research, but will probably not coincide with your own expectations. We would like to ask you some questions about this.

1. Did you consider what age you would achieve when you would have no more than 10 years to live, while answering the questions?

☐ Yes
☐ No

Question regarding the expected age of death

Often people have some expectation regarding what age they will reach themselves in life.

What age do you yourself expect to reach? ... years

5.

“Back to the future”: Influence of beliefs regarding the future on TTO answers

Based on: F. E. van Nooten, N.J.A. van Exel, D. Eriksson and W.B.F. Brouwer

“Back to the future”:
Influence of beliefs regarding the future on TTO answers

Health and Quality of Life Outcomes 2016; 14:4
DOI 10.1186/s12955-015-0402-6

Summary

Background: A common approach to obtain health state valuations is the time-trade-off (TTO) method. Much remains unknown regarding the influence of responder characteristics on TTO answers. The objective of this study is to increase understanding of the influence that beliefs regarding future health and death, as well as desires to witness certain life events, have on respondents' health state valuations.

Methods: An online survey was designed, including three TTO questions using a 10 year timeframe. Moreover, respondents completed demographic questions, the Health-Risk Attitude Scale (HRAS), the Expectations Regarding Aging (ERA) questionnaire, questions about beliefs regarding future health (i.e. life expectancy) and death (i.e. fear of death, belief in life after death and opinion about euthanasia), and about important life events taking place within the TTO timeframe. Regression analyses were performed in order to assess the influence of these different variables.

Results: One thousand sixty-seven respondents were included in the analyses. The following variables were significantly associated with years traded off: ERA mental health (decrease), ERA physical health (increase), HRAS (increase), support for euthanasia (increase), fear of death (decrease) and consideration of an important life event (decrease). The explained variance of the final model was low (0.08).

Conclusion: TTO responses may be influenced by considerations of future health, including life events and attitudes regarding health risks and death. Further investigation of TTO responses remains warranted.

5.1 Background

Several countries use cost-effectiveness analyses in the context of deciding on the reimbursement and funding of new medical technologies (e.g. United Kingdom, Sweden, the Netherlands). For the effectiveness component, some authorities prefer the use of quality adjusted life years (QALYs), as these are believed to allow for a universal comparison across diseases areas (Torrance, 1989). QALYs combine length and quality of life, with the latter normally expressed as a value between 0 (dead) and 1 (perfect health). Societal preferences obtained in the general public typically underlie these values (Greiner *et al*., 2005; Scalone *et al*., 2013).

Different valuation techniques are used to obtain health state valuations from the general public. The most widely used approach is the time-trade off (TTO) method. In a TTO exercise, a respondent is presented with two health streams. One of the health streams is a fixed lifespan (e.g. 10 years) lived in some imperfect health state “A”. The other stream entails a shorter life-span but lived in perfect health. Respondents are subsequently asked to indicate the minimum number of years lived in perfect health required to become indifferent between the two streams. If a respondent is not willing to live shorter in perfect health relative to the imperfect health state A, its value is assumed to be equal to that of perfect health (with value 1). When a respondent is willing to give up all remaining years, the value of health state A is considered to be equal to being dead (with value 0) (Torrance, 1986). If a respondent indicates to consider 6 years in perfect health (with value 1) equal to 10 years in state A, the value of health state A is assumed to be equal to 0.6. In this way, states of health impairment can be assigned a value between 0 and 1, with a higher score indicating a better health state. Other variants of the TTO exist for health states worse than death. In those cases the lead time TTO should be employed (Devlin *et al*., 2011, Attema *et al.*, 2013).

Despite the widespread use of TTO as valuation method for health state utilities in economic evaluations, relatively little is known about which characteristics of respondents are associated with responses to time trade-off exercises. Most of the previous research was focused on the typical demographics like age, gender, marital status and education (Dolan *et al*., 1996; Hsu *et al*., 2009; Kontodimopoulos & Niakas, 2006), but findings have been mixed both in terms of the direction of influences and their statistical significance. For example, several authors studied the relationship between age and TTO values as part of their analyses. Augestad *et al.* (2013), Best *et al.* (2010) and Hsu *et al.* (2012) found a positive relationship, while Ayalon &

King-Kallimanis (2010) Shimizu *et al.* (2008), and Zarate *et al.* (2008) found negative relationships. Similar findings were found for gender. Brown *et al.* (2002), Gupta *et al.* (2005) and Tamayama *et al.* (2009) found positive relationships with TTO scores and female gender, whereas Wells *et al.* (2004) and Rutten-van Mölken *et al.* (2009) found the opposite. These mixed findings may relate to numerous aspects, amongst others differences in studied populations (e.g. patient or general samples, cultural differences) and in methodology. Another explanation may be that unobserved variables, other than the standard demographic characteristics, influence TTO scores and confound some of the observed relationships.

Loss of future life years is a critical element of a TTO exercise, as length of life is traded off against quality of life. Therefore, attitudes towards future health and death for instance could play an important role. Some research has been performed in this direction by investigating the influence of subjective remaining life expectancy (SLE; calculated as the age the respondent assumes to reach minus current age). SLE turned out to have a negative relationship with the numbers of years traded-off in a TTO exercise. This means that a higher SLE is associated with a lower the number of years traded-off (Heintz et al., 2013; van Nooten & Brouwer, 2009; van Nooten *et al.*, 2009; van Nooten *et al.*, 2014)). Next to SLE, beliefs regarding life after death have been suggested to influence TTO values (Rutten-van Mölken *et al.* 2009).

Given the widespread use of TTO and its potential impact on reimbursement decisions, better understanding of such associations is important, not only to understand what drives TTO answers, but also for purposes of representative sampling. If these associations exist, comparing TTO scores from one study to another might not be valid if the sampled populations differ with respect to these variables.

The aim of this study is to obtain more insight in this underexplored relation of responses to TTO exercises with beliefs regarding future health and death.

5.2 Methods

A questionnaire was administered online by a survey company to a representative sample of the Dutch general public in the range 18 to 65 years, in terms of age, gender and level of education. Respondents who completed the survey in less than 15 min were considered to have devoted too little attention to the questions and, consequently, were excluded from the analyses. (This threshold for speeding through the questionnaire was based on the distribution of completion times in the pilot test.) The methods are described in more detail in van Nooten *et al.* (2015), who used the same dataset but focused on other parts of the same questionnaire.

The questionnaire first covered common demographic characteristics such as age, gender, marital status, nationality, education, having children, followed by questions regarding current health status, using EQ-5D and EQ-VAS (Euroqol). Next, respondents were asked to rank order six health states. Five of the six health states were described using the EQ-5D descriptive system. The five health states were perfect health, own current health status (as respondents reported previously in the EQ-5D), and three states of health impairment chosen to represent a broad range across health states (see Appendix A4 for an explanation of these health states). The sixth health state was labelled “dead”. After rank ordering these health states, respondents were asked to rate them on a visual analogue scale (VAS) ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). Finally, respondents were asked to perform three TTO exercises with a 10 year timeframe, for the three imperfect health states. The three states were presented in the order the respondent had ranked them (among own current health, perfect health and dead), from highest to lowest. The flow of the TTO questions is described in detail in Appendix B4.

After the TTO exercises, respondents were asked whether they had thought that the specified period of 10 years would start immediately. If answered affirmatively, respondents were asked if they had thought of a minimum period they wanted to stay alive, for example to witness a certain event or to reach a certain age, regardless of health). In case this question was answered affirmatively, respondents were asked to describe the event.

Respondents also answered a number of questions regarding beliefs about future health. First, respondents reported their subjective life expectancy (SLE) by providing a point estimate of expected lifetime.

Secondly, respondents completed the Health-Risk Attitude Scale (HRAS), which was developed to understand health related risk attitude (van Osch *et al*., 2014). The HRAS consists of 13 items, which are scored on a 7 points Likert scale (1 = totally disagree; 7 = totally agree). The 13 items are statements about how much risk respondents are willing to take with their health (for example: “When I look back at my past, I think that, in general, I did take risks with my health.” or, “Safety first, where my health is concerned”). The item scores are summed to obtain a total score ranging from 13 to 91, with a higher score indicating more risk seeking in the health domain (van Osch *et al*., 2014).

Thirdly, respondents completed the Expectations Regarding Aging (ERA) survey (Sarkisian *et al*., 2005). The ERA consists of three scales (i.e. expectations regarding physical health, expectations regarding mental health, and expectations regarding cognitive function, all related to aging) with four items each, making a total of 12 items, with 4 response options (1 – 4). The total score for each scale is calculated by summing the responses to each question, which is then rescaled to a range of 0 – 100, with higher scores indicating higher (that is, better) expectations regarding aging in the physical, mental health and cognitive function domains (Sarkisian *et al*., 2005).

Finally, respondents were asked several questions regarding beliefs about death, because the TTO involves shortening life duration in order to improve quality of life. Respondents were asked: “Do you believe in life after death?”, with the following answering possibilities (1) “no, I don’t believe in life after death”, (2) “yes, I believe heaven exists”, (3) “yes, I believe in reincarnation”, and (4) “yes, other (please explain)”. Then respondents were asked: “Are you afraid of death?” using a 0 – 100 visual analogue scale, with 0 representing no fear and 100 extreme fear. Finally, respondents were asked: “What is your attitude regarding euthanasia?” The following four answering possibilities were provided: (1) “I think that euthanasia should not be allowed under any circumstances”, (2) “I think that euthanasia should only be allowed under very strict circumstances (for example in case of unbearable suffering without any hope for the future)”, (3) “I think that euthanasia should be allowed after careful consideration and with professional support”, and (4) “I think that people should be free to opt for euthanasia”.

5.3 Data Analyses

Correlations coefficients were computed and were used to understand the relationship between the different variables included in this study, both the demographic variables (i.e. age, gender, partnership status, education, quality of life, number of children) and the beliefs regarding future health and death (i.e. SLE versus event, HRAS, ERA, fear of death, belief in life after death, attitude towards euthanasia).

Next, regression analyses were conducted. The dependent variable in the regression analyses was the number of years a respondent was willing to give up from the 10 year timeframe in order to regain full health, calculated by subtracting the TTO answer from 10. Remaining SLE was calculated by subtracting the actual age of the respondent from expected age of death. The variable for life after death was a dichotomous variable, in which the response options “yes, I believe heaven exists”, “yes, I believe in reincarnation” and “yes, other (please explain)” were classified as 1 and “no, I don’t believe in life after death” as 0. The variable called “event” was created in the following way: if respondents had answered affirmatively to both the question regarding whether they had thought of a minimum period they wanted to stay alive to witness a life event as well as the question whether they thought the 10 year period would start immediately, the variable event was defined as 1, otherwise as 0. Moreover, a dichotomous variable for euthanasia was created in which “I think that euthanasia is not allowed under any circumstances” was coded 1, whereas “I think that euthanasia is only allowed under very strict circumstances (for example in case of unbearable suffering without any hope for the future)”, “I think that euthanasia is allowed after careful consideration and with professional support”, and “I think that people should be free to opt for euthanasia” were coded 0.

5.4 Models

First, we estimated a base model, including the commonly investigated variables in the context of explaining TTO answers: age, gender, partnership status, having children, educational status, own health using the EQ-5D or VAS and SLE (van Nooten *et al.*, 2015).

Next, we expanded the base model with the variables related to beliefs regarding future health and death, and any important life events respondents wanted to witness within the 10 year timeframe of the TTO exercise. First, a principal-components factor analysis was conducted to explore the structure in the relationships between these additional variables (ERA mental, ERA cognitive and ERA physical scales, HRAS, belief in life after death, fear of death, views on euthanasia, and the wish to experience a specific future event). The factor analysis allowed for combining correlated variables into independent groups of variables, which were then added separately to the base regression model to determine which individual variables could further explain the TTO scores. Those variables that reached statistical significance ($p < 0.05$) in explaining years traded off in the TTO exercises were included in the final model.

The data were analysed using random-effects models to account for the repeated TTO measures. Confidence intervals were obtained via bootstrapping, as the data were not normally distributed. All analyses were conducted in Stata/IC version 12.1 for Windows (Stata-Corp LP, College Station, TX, US).

Table 5.1 Demographics of respondents

	All respondents (n=1,067)
Age (years) (mean, (SD), range)	43.2 (13.64) 18-65
Gender (male) (%)	50.2%
EQ-VAS (mean, (SD))	75.0 (16.59)
High education (%)	30.9%
Married (%)	49.0%
Living together (%)	15.3%
Children (yes) (%)	60.2%
SLE (years) (mean, (SD))	37.8 (17.21)
ERA Mental Health (mean, (SD), range)	65.4 (22.33) 0-100
ERA Physical Health (mean, (SD), range)	31.1 (17.45) 0-100
ERA Cognitive Health (mean, (SD), range)	40.3 (19.91) 0-100
HRAS (mean, (SD), range)	44.9 (9.63) 15-84
Fear of death (scale 0-100) (mean, (SD), range)	35.5 (30.14) 1-100
Support for euthanasia (allowed) (%)	95.1%
Belief in life after death (yes) %	55.8%
Considered Event (yes) %	15.8%

5.5 Results

Responder characteristics

From the original 1,223 respondents who completed the survey, 156 were excluded for speeding through the questionnaire, which left a total of 1,067 respondents for the analyses. Table 5.1 presents the demographics of this sample. The mean age of the total responder population was 43 years, half were male and the mean VAS score was 75. In responding to the TTO questions, 16% of respondents had considered an event they wanted to witness or an age they wanted to reach. In 62% of these cases, the event was related to children or grandchildren (e.g. birth, seeing them grow up to be independent, or attending their wedding). Five percent of the respondents were against euthanasia under all circumstances and 56% of the respondents believed in life after death. On a VAS scale from 0 to 100, the mean fear of death score was 36. The mean ERA score for the mental health subdomain was 60.4, for the physical health subdomain 31.1 and cognitive function 38.5. The average HRAS score was 44.9.

In general, the correlations between the demographic variables (i.e. age, gender, partnership status, education, quality of life, number of children), beliefs regarding future health and death (i.e. SLE, euthanasia, HRAS, ERA, fear of death, belief in life after death) and staying alive to witness a life event were weak, ranged between − 0.2 and 0.2.

Influence of future expectations

Table 5.2 first shows the results of the base model. Positive coefficients indicate an increase in the number of years traded, while negative coefficients indicate a decrease in the number of years traded. In this model, age (decrease), being male (increase), living together (increase), having children (decrease), quality of life (VAS) (increase) and subjective life expectancy (decrease) were statistically significantly associated with years traded-off. The association with education level and being married did not reach statistical significance.

The factor analysis resulted in three groups of variables to be added to the base model: 1) expectations regarding aging (the ERA mental, cognitive and physical scales), 2) beliefs regarding death (fear of death, belief in life after death), and 3) health related risk attitude (HRAS), and consideration of a future life event. Support for euthanasia did not load into any of the factors and showed limited variation (only 5% of the population was against euthanasia). It was added separately to the base case model.

These three groups of variables were first added to the base case model independently in order to assess their separate association with years traded off. For the variables in group 1 (expectations regarding aging), only the ERA mental and physical scales proved to be significantly associated with years traded off. For the variables in the other two groups (group 2: beliefs about death; group 3: health risk attitude) all variables were statistically significantly associated with years traded off when added separately to the base model. When added, support for euthanasia also proved to be significantly associated with years traded off. Therefore, the following variables were included in the final model: ERA mental and physical, fear of death, belief in life after death, HRAS and consideration of a future life event. When these variables were jointly added to the base case model, the following statistically significant results were observed: ERA mental health (increase), ERA physical health.

Table 5.2 Results including previous used variables (dependent variable: Years traded-off)

	Base Model			**Final Model**		
R^2	0.04			0.08		
	Coeffi-cients	Bias corrected 95% Confidence interval		Coeff-icients	Bias corrected 95% Confidence interval	
Age	**- 0.040**	**- 0.049**	**- 0.032**	**- 0.041**	**- 0.054**	**- 0.028**
Male	**0.275**	**0.117**	**0.426**	0.032	-0.184	0.259
EQ-VAS	**0.005**	**0.000**	**0.009**	0.006	- 0.001	0.013
Highest Education	0.045	- 0.12	0.218	0.094	- 0.132	0.311
Married	- 0.221	- 0.422	0.019	- 0.169	- 0.433	0.093
Living together	**0.369**	**0.140**	**0.595**	**0.327**	**0.030**	**0.635**
Children	**- 0.343**	**- 0.541**	**- 0.149**	**- 0.266**	**- 0.541**	**- 0.007**
SLE	**- 0.028**	**- 0.035**	**- 0.021**	**- 0.024**	**- 0.034**	**- 0.014**
ERA mental health				**- 0.009**	**- 0.015**	**- 0.004**
ERA physical health				**0.007**	**0.000**	**0.013**
HRAS				**0.017**	**0.007**	**0.029**
Fear of death				**-0.009**	**- 0.013**	**-0.006**
Support for euthanasia (allowed)				**0.228**	**0.096**	**0.360**
Belief in life after death (yes)				- 0.209	- 0.417	0.010
Considered event (yes)				**- 0.581**	**- 0.871**	**-0.284**

*Bold are statistically significant based on confidence intervals (increase), fear of death (decrease), HRAS (increase), consideration of a future life event (increase) and support for euthanasia (increase) (Table 5.2). Belief in life after death was not statistically significantly associated with years traded off (Table 5.2). Adding these additional variables to the model resulted in two variables from the base model to lose their significant association with years traded off (i.e. VAS and gender). Furthermore, compared to the base model, the variance explained by the final model doubled, although the absolute value of R^2 remained modest.

5.6 Discussion

While different responder characteristics can influence TTO responses (Dolan *et al*, 1996; Hsu *et al*., 2009, Kontodimopoulos & Niakas, 2006), previous research has focused mainly on demographics like age, gender and marriage (Dolan *et al*, 1996; Hsu *et al*., 2009, Kontodimopoulos & Niakas, 2006). This research has introduced a new category of variables, but has not been able to explain much more of the variance in TTO responses than previous studies. The objective of this study was to investigate the influence that beliefs about future health and death, and desires to witness certain life events, have on TTO responses. Since TTO exercises ask respondents to trade-off (future) quality of life and life duration, beliefs and desires regarding the future may well play a role in final responses. It already has been observed that expectations about length of life play a role in TTO responses (Heintz *et al*., 2013; van Nooten & Brouwer, 2004; van Nooten *et al*., 2009; van Nooten *et a*l., 2014; van Nooten *et al*., 2015). Given the importance of adequate health state valuations, it seemed worthwhile to explore this further. We found that both beliefs about future health and death and desires to witness a life event indeed had a significant, though modest, influence on years traded off in a TTO exercise. The effect sizes of these newly identified variables influencing TTO scores are small, however no smaller than previously identified variables (e.g. age and gender). The explained variance increased compared to the base case model showing that the newly identified variables provide more clarification of factors influencing TTO scores. This also indicate that there is probably not one variable influencing TTO scores, but that there are many pieces to this puzzle that need to be put together.

Before discussing the implications of this study, several limitations need to be noted. First, we used a web-based design in our study, which may have had consequences on, for instance, the involvement of respondents in the questionnaire and did not allow face-to-face explanations of questions. Based on a pilot test of the questionnaire we determined a minimum acceptable completion time of 15 min, in order to limit the effect of speeding through the questionnaire and low involvement. Second, there was no separate valuation exercise for health states ranked as being worse than dead, which may have influenced our results. Valuing worse than dead states using a distinct valuation exercise was considered cognitively demanding and alternative methods, which allow better than dead and worse than dead states to be valued in one exercise, such as the lead time TTO (Attema *et al*., 2013) may have the same problem and require further validation. Third, only a single iteration was performed in the TTO exercise, instead of a more common repeated choice ("ping-pong") exercise, which could have allowed for a more precise estimation of the responders' indifference point in the trade-off exercise, and to different results (Attema & Brouwer, 2013). Fourth, this study was performed in the Netherlands, where certain values, norms and beliefs about life and death may be different from other countries. Therefore, extrapolating these results to other countries requires caution. For example, the vast majority of the Dutch society is not against euthanasia (only 5% of the responders in this study were against euthanasia under all circumstances). This may not be the case in other countries. Performing a similar study in countries where people generally hold different beliefs about life after death, could lead to different results. Fifth, this study only included respondents up to the age of 65 years. The elderly, however, may exert preferences regarding for example euthanasia and beliefs about the future that are different compared to young adults due to age and lived experience. Hence a study including more respondents above the age of 65 could provide different results.

Notwithstanding these limitations, our results showed an influence of beliefs about future health and death on TTO answers. Given the limited knowledge so far in this area, these findings add to the existing literature. First of all, respondents who have higher expectations about future mental health, who aim to stay alive in order to witness a particular life event and those who are afraid of death, all traded off fewer years in the TTO exercises. Respondents who were not opposed to euthanasia were willing to give up more years, as well as those who were more risk seeking in the health domain. Although the explained variance of the final model remained low, the influence of support for euthanasia and the desire to witness a particular life event was remarkably high compared to other statistically significant variables in the model.

Research investigating the influence of beliefs about death on TTO scores is scarce. Rutten-van Molken *et al.*2006) found that beliefs about life after death were significantly associated with TTO scores. In our study this association was not statistically significant in the final model. This difference may relate to differences in studied populations, included variables and differences in TTO design. Furthermore this study also included fear of death and euthanasia, which although not the same as beliefs about life after death, could have mitigated the effect. More research in this area appears to be warranted.

Another interesting finding was that the ERA variable mental health was negatively associated with years traded off, suggesting that when responders have higher expectations regarding mental aging they are willing to give up fewer years. Although not in line with the instruction and intention of the TTO exercise, which specified a stable health state for the 10 year timeframe and should therefore render own expectations irrelevant, this implies that future years are more easily traded when one expects these will be spent in relatively poor mental health. How-ever, higher expectations regarding physical health were associated with more years sacrificed. This rather counter-intuitive result may be related to the skewed distribution of ERA physical health (see Table 5.1).

It should be noted that a 10-year time-horizon was applied for the health state valuation in this study. However it could be expected that in a TTO exercise with a longer time horizon, e.g. life time, the associations highlighted in this study may play an even greater role and this should be investigated.

5.7 Conclusion

This study showed that TTO scores are associated with considerations of future life events and beliefs about future health and death, next to the more common demographic variables observed in the literature like age and SLE. However, the overall impact on providing a better understanding of what drives responders to make choices in TTO scoring remained limited. Therefore, much remains unknown and more research in this area is warranted. Nonetheless, the results of this study suggest that beliefs about the future can be influential in TTO exercises and should therefore be considered in future research on predictors of TTO responses. These findings may also be relevant in sampling for representative societal valuations of health states.

6.

"Married with children" The influence of significant others in TTO exercises

Based on: F.E. van Nooten, N.J.A. van Exel, X. Koolman and W.B.F. Brouwer

*"Married with children"
The influence of significant others in TTO exercises*

*Health and Quality of life Outcomes 2015; 13:94
DOI 10.1186/s12955-015-0276-7*

Summary

Background: Which responder characteristics influence TTO scores remains underexplored. More research is needed in order to understand (differences in) TTO scores, but also in the context of generating representative health state valuations for some population. Previous studies have found age, gender, marital status and subjective life expectancy to influence the number of years traded off.

Objective: This study aimed to investigate which other responder characteristics influence TTO responses, with an emphasis on consideration of significant others, such as partners and children.

Methods and Design: We performed a web-based survey in a representative sample of the Dutch general public (aged 18–65). Data on demographics, health status and expectations about future length and quality of life were gathered. Respondents valued three distinct health states using TTO.

Results: A total of 1,067 respondents completed the questionnaire. Sixty percent of respondents had children and 49 % were married. The mean number of years traded off increased with severity of health states. Higher age and living together were positively associated with number of years traded off. Increases in subjective life expectancy, having children and being male (were negatively associated with the number of years traded-off.

Conclusion: Age, gender and subjective life expectancy, living together and having children were significantly associated with TTO responses. Consideration of significant others in TTO exercises thus may be important in understanding (differences in) TTO responses and when drawing representative samples from the general public.

6.1 Introduction

A commonly used method for deriving health state valuations is the time trade-off (TTO) method. These health state valuations are important in the context of constructing QALY scores for different health profiles, which can be used in cost-effectiveness analyses of health technologies. The TTO method is widely used, e.g. to derive national tariffs for often used measures like the EuroQol EQ-5D instrument (e.g. Dolan, 1997; Lamers *et al*., 2006).

In a TTO exercise, respondents are asked to indicate their indifference between two streams of health that differ in terms of length and quality of life. Commonly, one stream entails an imperfect health state that will last for a fixed period of time X (often 10 years). The second stream entails perfect health, but this is enjoyed for a shorter period of time than X. The TTO exercise then requires respondents to reveal their point of indifference between two streams, usually by varying the number of years in perfect health.

TTO exercises thus provide a relatively simple means of deriving health state valuations, allowing QALY scores for health states to be used in health economic evaluations. Nonetheless, the TTO method has been shown to be prone to several types of biases (Bleichrodt, 2002), its operationalization may influence results (Arnesen & Trommald, 2005), and the way in which the answers of respondents usually are interpreted has been argued to be too simplistic (e.g. not accounting for time preferences for health) (Attema & Brouwer, 2009). Moreover, it has been shown that certain respondent characteristics like age, gender and marital status can influence the number of years traded-off (i.e. health state valuations) (Dolan & Roberts, 2002). This emphasizes that it is important to know which characteristics of respondents influence TTO scores. Knowledge and understanding of the characteristics of respondents that influence TTO scores is important for several reasons: (i) when comparing health state valuations between respondents, populations or studies, (ii) for sampling purposes, when the aim is to generate representative health state valuations for a specific population, and (iii) to in-crease understanding of underlying considerations and mechanisms driving observed health state valuations.

Three clusters of responder characteristics influencing TTO scores may be distinguished: (i) the demographic characteristics age, gender, and marital status (Dolan & Roberts, 2002); (ii) health status (Devlin *et al*., 2011, Best *et al*., 2010) and health related characteristics (such as body mass index) (Hakim *et al*., 2002); and (iii) subjective reference points for future length and quality of life (Heintz *et al*., 2013; van Nooten & Brouwer, 2004; van Nooten *et al*., 2009).

It is conceivable that other responder characteristics are associated with the number of years traded in TTO exercises as well. TTO decisions may be especially affected by consequences on significant others, such as partners, children and family members. For instance, living shorter by trading off more life time may be considered less attractive when one has young (dependent) children, like observed by Van Der Pol & Shiell (2007). They found that recent mothers valued health states differently than the general population. Similarly, Devlin and colleagues found that females in a household with children provided significantly higher health state values (Devlin *et al*., 2011). Health state valuations may also be influenced by consideration of the consequences on partners and family members of living shorter or in poor health. For instance, Matza and colleagues (2014) observed that individuals with a caregiver role were less willing to trade off time to improve their health status than non-caregivers. Conversely, consideration of others could also lead to increased trading of years, in order not to be a burden to others as recipient of informal care (Krol *et al*., 2015).

While former studies have demonstrated the influence of marital status and having children on health state valuations using TTO, it remains unclear whether these are separate effects or that marital status perhaps combines the influence of having a partner and having children on TTO scores (Dolan and Roberts, 2002). This chapter reports the results of a study that aimed to investigate the association of a broad range of responder characteristics with TTO scores, with a special emphasis on the consideration of significant others and on disentangling the influence of having a partner and having children on TTO scores.

6.2 Methods

The questions for this study were part of a larger web-based questionnaire which was administered online by a professional survey company to a representative sample of the Dutch general public in terms of gender and age (in the range 18 – 65 years). A minimum completion time of 15 min was set, based on a pilot test of the survey. Respondents who completed the survey in less than 15 min were identified as speeders and excluded from the analyses. A total of 1223 respondents participated in the online survey; 156 (12.8%) respondents were disregarded because of speeding. For the remaining 1067 respondents, mean response time for the questionnaire was 27.8 min (SD 8.2; range 15–63).

At the time of the design of this web-based TTO exercise, neither a universal web-based TTO protocol existed nor a standardized EuroQol web-based TTO. Hence, one was created. The questionnaire first asked about general responder characteristics, such as age, gender, marital status, nationality, level of education, having children, number of children and age of the youngest child. Next, respondents were asked about their current health (using VAS and the EQ-5D) and whether they currently have or have had a chronic or serious condition. In addition, they were asked the following question: “If, due to some illness, you had to choose between a shorter life in good health and a longer life in poorer health, what would you choose at this moment?”

Then, respondents were asked to value six health states. Five of the six health states were described using the three level EQ-5D descriptive system. The five health states were own health (as previously indicated by the respondent), perfect health and three imperfect health states (EQ-5D profiles 21211, 22221 and 33312 - see Appendix A4). The latter were chosen to represent a broad range across health states and identical to those used in previous studies (van Nooten et al., 2009; van Nooten et al., 2014), also to facilitate comparisons. The sixth state was labelled as “dead”. To familiarize the respondents with the health states and tasks, they were first asked to rank these six health states and then to rate them using a visual analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). Finally, respondents were asked to perform TTO exercises for the three imperfect health states mentioned above. The three TTO exercises (Appendix B4) were presented to respondents in the order in which they had ranked them in the ranking exercise, with the highest ranked health state first.

In the TTO exercise respondents were first asked to choose between living 10 years in a specific imperfect health state followed by death (option A), or living 10 years in a perfect health state followed by death (option B). They could also indicate to be indifferent be-tween the two (Option C). If the dominated option A was chosen, respondents needed to confirm this preference or choose again. If option B was chosen, they were asked to choose between living 10 years in the imperfect health state, after which they would die (option BA), living 5 years in a perfect health state, after which they would die (option BB), or being indifferent between these options (option BC). If respondents chose option BA they were shown a slider ranging from 0 to 10 years and asked to indicate how many years in perfect health would be equivalent to living 10 years in the imperfect health state presented to them. If respondents chose option BB they were shown a slider ranging from 0 to 5 years and asked to indicate how many years in perfect health would be equivalent to living 10 years in the imperfect health state. The slider in options BA and BB allowed indicating years with one decimal level of precision. Finally, respondents were asked to confirm that they were indifferent between living X years in perfect health and 10 years in the imperfect health state, with X taking the slider value in case of options BA and BB or 5 in case of option BC. If respondents immediately chose option C they were asked to confirm that they were indifferent be-tween living 10 years in imperfect or perfect health state. In all cases, if respondents did not confirm their choice, they returned to the beginning of the question.

In this TTO exercise we did not include a protocol for states perceived as being worse than dead, since we felt this would be too complex in the context of self-completed on-line questionnaires (van Nooten et al., 2009). The ranking of the six states indicated that especially the worst health state (33312) was ranked as worse than dead by some of the respondents (around 9%). Since we did not have a separate valuation exercise for states considered to be worse than dead, we included all TTO responses obtained in the regular TTO exercise in the analyses. No responses were excluded therefore.

After the three TTO exercises, respondents were asked whether they had related the 10 year time frame to their own life while answering the questions, and whether they had assumed the 10 years period to start immediately. Moreover, we asked respondents whether they had considered specific moments in time that they wanted to reach (e.g. an anniversary or specific age) within that time frame. If so, they were asked to indicate which moments.

Finally, respondents answered some questions regarding their subjective life expectancy and their quality of life expectations at the ages of 60 (if aged 59 or less), 70, 80 and 90 years, using the EQ-5D as done before (Brouwer & van Exel, 2005; Penték *et al*., 2014; van Nooten & Brouwer, 2004).

A number of variables were constructed for the analyses. The number of years sacrificed out of the remaining 10 was calculated (10 minus the minimum number of years required in perfect health). Subjective life expectancy (SLE) was calculated by deducting the actual age from expected age of death. The utilities for quality of life expected at 60, 70, 80 and 90 years were calculated using the expected EQ-5D health profiles and the Dutch EQ-

5D tariffs (Lamers *et al*., 2006). Body mass index was calculated by dividing the weight responders provided by the square of their height (in meters).

Analysing the data, we first ran a regression model including explanatory variables that used in previous research, as discussed in the introduction: age, gender, marital status, educational status, own health and subjective life expectancy. Next, the contribution of three additional types of responder characteristics was investigated, i.e.: demographics (gender, age, highest education, marital status, children yes/no, number of children, age youngest child), health (VAS; chronic illness; serious ill-ness; weight) and expectations (subjective life expectancy and quality of life at the ages of 60, 70, 80 and 90). First, regressions were performed per TTO question per type of responder characteristic, giving in total three regressions. If a variable was statistically significant ($p < 0.05$) in at least one of the regressions it was included as explanatory variable in the final model. As the data was not normally distributed, bias corrected confidence intervals were obtained using a non-parametric bootstrap procedure using 10,000 replications.

6.3 Results

Responder characteristics

A total of 1223 respondents completed the questionnaire of which 156 (12.8%) were removed because of speeding through the questionnaire. The remaining 1067 respondents were included in the analyses. Table 6.1 shows that the mean age of the sample was 43 years, half of the respondents were male and mean VAS score was 75. Sixty percent of the total sample had children, varying be-tween 8% among those who were single and 86% among those who were married. Twenty eight percent of the respondents indicated to have (had) a serious condition and 36.6% indicated to have (had) a chronic condition. The average BMI was 26.4 and 19.1% indicated to be overweight.

In total, 56% of the respondents indicated to have a preference for a shorter life in perfect health over a longer life in imperfect health. Irrespective of their living situation (single or in a partnership), 63% of the respondents without children and 49% of the respondents with children would prefer quality of life over longevity. In addition, irrespective of having children, 50% of the responders who were married chose longevity over quality of life while this was 36% of the unmarried respondents who did have a partner.

In total, 16% of respondents indicated to wish to reach a specific moment in time, of which 68% had children.

Numbers of years traded off in the health states

The average number of years respondents wanted to trade off was 3.16 years for health state 1 (i.e. the highest ranked one), 3.80 years for health state 2 and 5.63 years for health state 3 (Table 6.2). Figure 6.1 shows the distribution of years traded off in each of the health states. Table 6.2 shows that respondents who indicated to have a preference for a shorter life in perfect health over a longer life in imperfect health, indeed traded off significantly more years than respondents who preferred to live longer in imperfect health ($p < 0.05$).

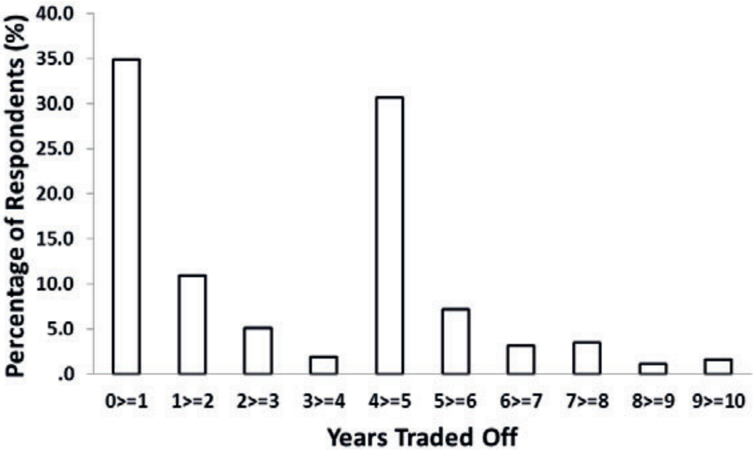


Figure 6.1 Percentage of respondents willing to trade of years for health state 1: 21211 (N=1,067)

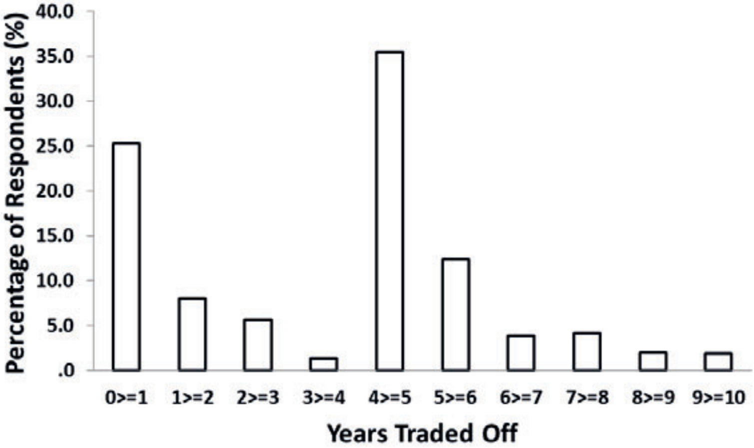


Figure 6.2 Percentage of respondents willing to trade of years for health state 2: 22221 (n=1,067)

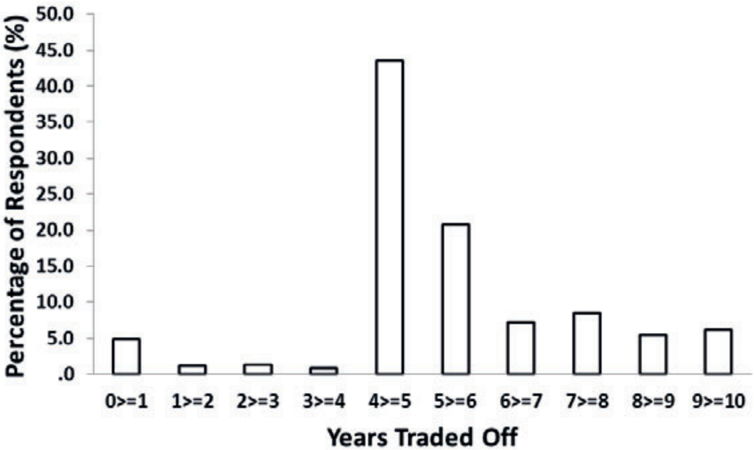


Figure 6.3 Percentage of respondents willing to trade of years for health state 3: 33312

Chapter 6

Table 6.1 Demographics of the sample (n=1067)

Age (mean, SD, range)		43.2 (13.64) 18-65
Gender (male) (%)		50.2%
Education (%)	Lower	15.4%
	Middle	53.7%
	Higher	30.9%
Marital status (%)	Married	49%
	Living together	15.3%
	Divorced	8.5%
	Widow(er)	2.2%
	Single	21.5%
	Don't want to reveal	3.5%
Children (yes) (%)		60.2%
Number of children (mean, SD, range)		2.1 (0.94) 1-11
Age of youngest child in years (mean, SD, range)		17.1 (11.4) 1-44
Dutch (%)		98.6%
Employed (%)		47.3%
Current quality of life EQ-5D VAS (mean, SD)		75.0 (16.59)
Current quality of life EQ-5D utility (mean, SD)		0.85 (0.23)
Do you have ever had a serious condition? (yes) (%)		28.2%
Do you have a chronic disease? (yes) (%)		36.6%
BMI (mean, SD)		26.4 (5.08)
Overweight? (yes) (%)		56.1%
Obese? (yes) (%)		19.1%
SLE (mean, SD)		37.8 (17.21)
Quality of life at years 60 (EQ-5D utility (mean, SD)) (N= 921)		0.77 (0.27)
Quality of life at years 70 (EQ-5D utility (mean, SD))		0.69 (0.30)
Quality of life at years 80 (EQ-5D utility (mean, SD))		0.51 (0.37)
Quality of life at years 90 (EQ-5D utility (mean, SD))		0.32 (0.42)

Table 6.2 Years traded off of respondents who prefer to live shorter in perfect health compared to respondents who prefer to live longer in imperfect health

	Mean (SD)	Shorter in perfect health (n=599)	Longer in imperfect health (n=468)	P*
Health State 1 (21211)	3.16 (2.58)	3.65 (2.48)	2.53 (2.56)	0.00
Health State 2 (22221)	3.80 (2.56)	4.43 (2.31)	2.99 (2.65)	0.00
Health State 3 (33312)	5.63 (2.01)	6.00 (1.79)	5.15 (2.17)	0.00

* independent samples T-test

Table 6.3 shows the results of the model that included variables also used in previous research (model 1). A negative sign means that respondents were willing to give up less years in the TTO exercise and a positive sign more years. In this model age (negative), marital status (negative) and subjective life expectancy (SLE) (negative) were statistically significantly associated with the number of years traded off. Education level and health measured by VAS were not statistically significantly associated with the number of years traded off.

The three regression analyses performed per TTO question per responder characteristic showed that only two variables from the long list of additional demographic, health, and expectations variables were statistically significantly associated with TTO answers: the dummy variable (yes/no) for living together (but not married) and the dummy variable (yes/no) for having children.

Table 6.3 shows the results of the model including the variable “living together” (model 2). The results show that age (negative), male gender (positive), married (negative), and SLE (negative) remained statistically significant, and that only the coefficient for being married changed meaningfully.

Table 6.3 shows the results of a model adding the variable “having children” to the model (model 3). Age, gender and SLE remained significantly associated with years traded off and their coefficients had the same sign as before. The variable VAS now also showed a significant (positive) association with years traded off, while having children was negatively associated with years traded off. Respondents with children traded off fewer years than those without, therefore. Interestingly, marital status lost significance while people living together significantly traded off more years than others. (A model including a single dummy variable “having a partner” - encompassing both “married” and “living together” - showed this new variable not to be statistically significantly associated with years while all other results remained similar to those presented in Table 6.3).

Table 6.3 Results (dependent variable: Years traded-off)

	Model 1			**Model 2**			**Model 3**		
	Average marginal effect	Bias corrected 95 % CI		Average marginal effect	Bias corrected 95 % CI		Average marginal effect	Bias corrected 95 % CI	
Age	− 0.044	− 0.057	− 0.032	− 0.044	− 0.053	− 0.036	− 0.040	- 0.049	− 0.032
Male	0.306	0.092	0.521	0.313	0.161	0.466	0.275	0.117	0.426
EQ-VAS	0.004	− 0.002	0.011	0.004	− 0.001	0.009	0.005	0.000	0.009
Married	− 0.467	− 0.694	− 0.240	− 0.377	− 0.552	− 0.197	− 0.221	− 0.422	0.019
Highest Education	0.066	− 0.166	0.298	0.078	− 0.092	0.247	0.045	− 0.120	0.218
SLE	− 0.028	− 0.038	− 0.018	− 0.028	− 0.035	− 0.021	− 0.028	− 0.035	− 0.021
Living together				0.312	− 0.091	0.531	0.369	0.140	0.595
Children							− 0.343	− 0.541	- 0.149

6.4 Discussion

Not many previous studies have studied the association between TTO responses and a wide variety of background characteristics of respondents. The aim of this study was to do so, with a special emphasis on the consideration of significant others. We investigated the association of these variables with the responses to a TTO exercise solved by a large and representative sample from the Dutch general public, valuing three distinct health states and using a 10 year time frame. The results provided some interesting insights. Expanding the regression model to include a broader range of variables showed having children and living together to be significantly associated with TTO scores. Before addressing our results in more detail, some limitations of our study are highlighted.

A first limitation was that we used an online survey for our study. Given the complex and uncommon task respondents need to fulfil in a TTO exercise, one may expect responses in an interview setting to more accurately describe health state preferences (although every method may have its own advantages and disadvantages). It needs noting that successful TTO studies have been performed online before (e.g. Tilling *et al*., 2012). Second, we did not include a separate valuation module for worse than dead health states, since we felt that the current TTO exercise was already cognitively demanding for members of the general public, as previously argued (van Nooten *et al*., 2009; van Nooten *et al*., 2014). This may have influenced our results and it remains interesting to see how the variables investigated here would relate to negative valuations of health states. Third, while our study included a broad array of variables, it needs noting that other potentially influential variables (e.g. more attitudinal questions) also remain understudied. This could be investigated in future research. Fourth, the health state valuations observed in our study were higher than the corresponding national EQ-5D tariffs (Lamers *et al*., 2006). This may well relate to our operationalization of the TTO method, which has been shown to be influential before (Bleichrodt, 2002). An important difference between the EuroQol protocol and this study is the number of iterations in the choice method used to obtain the indifference point. Here, we only used one iteration. A more intensive iteration procedure may help respondents to reach their indifference point more accurately and could result in more variation in the answers. It needs noting that the results from this study are closer to the corresponding national EQ-5D tariffs than previous studies which did not use choice based methods (van Nooten & Brouwer, 2004, van Nooten *et al*., 2014). More research in the relative advantages of different preference elicitation techniques remains warranted, also in the field of health (Attema & Brouwer, 2013).

Notwithstanding these limitations and areas for further research, this study has provided some interesting results. First, when running the regression model including the more commonly included variables, our results resemble those of earlier research. Notably, SLE and age had a significant influence on years traded off as observed before (van Nooten & Brouwer, 2004, van Nooten *et al*., 2014). This suggests that the influence of these characteristics is relatively stable across studies, supporting the generalizability of these and earlier findings.

Previous research classified marital status as being the third most important influence on TTO scores (Dolan & Roberts, 2002). However, marital status may be strongly related with both having a partner and with having children. This relationship became clear in our analysis. Only introducing the variable “living together” to the model next to “being married” did not affect the sign and significance of the variable “being married”, while “living together” did not reach significance. However, when also adding the variable “having children” to the model, the results changed. The effect and significance of “being married” was taken over by the variable “having children”, rendering the “being married” insignificant. This suggests the variable “being married” may proxy “having children” if the latter is not accounted for. Hence, the influence of marriage per se may be less strong than sometimes suggested.

Moreover, an intriguing result of the final analysis is the positive association between years traded off and “living together”. While being married commonly was associated with less years traded off (albeit insignificantly so in our final model), living together was associated with significantly more years traded off. The question why living together (but not being married) could lead to more years traded off cannot be answered with this study. However, a recent study by Krol and colleagues (2015) showed that respondents, when answering TTO questions, exhibit altruistic preferences. That is, they consider the consequences of their choices on significant others. However, two distinct considerations can be distinguished; one focusing on longevity (living longer for the others despite of poor health, for instance not to be missed), the other focusing on quality of life (giving up more years in order not to be a burden for loved ones). A possible explanation for our results could be that people living together may more often focus on quality of life, while being married (and especially having children) may lead to a focus on longevity. Such motivations behind response patterns are an important area for future research.

Respondents with children more often indicated they had a specific moment in time in mind that they would like to reach, while solving the TTO exercise. This moment in time was often related to children and grand-children (e.g. seeing them grow up, being at a wedding of children, living long enough for them to be old enough to be independent) among respondents with children. Of those without children, many of the reasons revolved around having a family.

Future quality of life expectations were not significantly associated with TTO responses in this study, although they have been shown to influence TTO responses using a time frame linked to full life expectancy (van Nooten & Brouwer, 2004). Quality of life at older ages may therefore be less relevant in a TTO using a 10 year time frame, also because the relevant ages normally are not reached within the 10 year time frame.

Concluding, this study has further explored the question which respondent characteristics influence TTO scores. Next to age, gender and subjective life expectancy, it seems that the variables living together and having children are also influential. The influence of these factors may have been attributed to the variable being married in previous research, but our results suggest the underlying mechanisms to be more diverse and complex. More research into these factors appears warranted in order to improve our understanding of TTO responses.

Appendix Chapter 6: A6

Health states used in the TTO exercise, in addition to 'own current health status' and 'dead'

11111 (perfect health):
I have **no problems** walking about

I have **no problems** with selfcare

I have **no problems** with performing my usual activities

I have **no** pain or discomfort

I am **not** anxious or depressed

21211:

I have **some problems** walking about

I have **no problems** with selfcare

I have **some problems** with performing my usual activities

I have **no** pain or discomfort

I am **not** anxious or depressed

22221:
I have some problems walking about

I have some problems with self care
I have some problems with performing my usual activities

I have moderate pain or discomfort

I am **not** anxious or depressed

33312:

I am **confined** to bed

I am **unable** to wash or dress myself

I am **unable** to perform my usual activities

I have **no pain** or discomfort

I am **moderately** anxious or depressed

Appendix Chapter 6: B6

The structure of the TTO question

In the TTO exercise, respondents received three questions for each health state they were asked to value. The health states were provided in the order they were ranked.

Introduction

Health is valued differently by people. Some people prefer to live shorter but completely healthy, while others prefer to live longer despite poorer health. For three of the five health states that you just rated you will be asked to choose between living another 10 years in that health state or living shorter but in perfect health. We start with the health you rated as the best, and end with the one you rated as the worst.

[*Question 1*]

Imagine you can choose between the following two options:

Box with the imperfect health state	*Box with the perfect health state*
I have **some problems** walking about I have **no problems** with self-care I have **some problems** with performing my usual activities I have **no** pain or discomfort I am **not** anxious or depressed	I have **no problems** walking about I have **no problems** with self-care I have **no problems** with performing my usual activities I have **no** pain or discomfort I am **not** anxious or depressed
OPTION A: You live another <u>10 years</u> in the health state above, without any change, and then you die.	**OPTION B**: You live another <u>10 years</u> in the health state above (perfect health), without any change, and then you die.

Which option do you prefer?

☐ OPTION A
☐ BOTH OPTIONS ARE EQUALLY GOOD

☐ OPTION B
[*If option A was chosen, respondent received the follow-up question below, else skip to question 2.*]
Are 10 years living in poorer health (Option A) better than 10 years living in perfect health (Option B)?

☐ Yes [*If yes, skip to question 2.]*
☐ No [*If no, back to question 1.]*

[*Question 2*]

Imagine you can choose between the following two options:

Box with the imperfect health state	*Box with the perfect health state*
I have **some problems** walking about	I have **no problems** walking about
I have **no problems** with self-care	I have **no problems** with self-care
I have **some problems** with performing my usual activities	I have **no problems** with performing my usual activities
I have **no** pain or discomfort	I have **no** pain or discomfort
I am **not** anxious or depressed	I am **not** anxious or depressed
OPTION A: You live another 10 years in the health state above, without any change, and then you die.	**OPTION B**: You live another 5 years in the health state above (perfect health), without any change, and then you die.

Which option do you prefer?

☐ OPTION A
☐ BOTH OPTIONS ARE EQUALLY GOOD
☐ OPTION B
[*If both options are equally good was chosen, respondent received the follow-up question below, else skip to question 3.*]Do you find living 5 years in perfect health (Option B) equally good as living 10 years in poorer health (Option A)?

☐ Yes [*If yes, skip to question 3.*]
☐ No [*If no, back to question 2.*

[*Question 3*]

Suppose you can choose between living 10 years in poorer health or less than ten years in perfect health. In both cases you die afterwards.

Box with the imperfect health state	*Box with the perfect health state*
I have **some problems** walking about I have **no problems** with self-care I have **some problems** with performing my usual activities I have **no** pain or discomfort I am **not** anxious or depressed	I have **no problems** walking about I have **no problems** with self-care I have **no problems** with performing my usual activities I have **no** pain or discomfort I am **not** anxious or depressed
OPTION A: You live another <u>10 years</u> in the health state above, without any change, and then you die.	**OPTION B**: You live less than <u>X years</u> in the health state above (perfect health), without any change, and then you die.

Note: X=10 if respondent chose option A in question 2; X=5 if respondent chose option B in question 2.

Please indicate on the scale below the <u>minimum</u> number of years that you would like to live in option B in order to make it equally good to option A. How many years in perfect health (option B) would make both options equally good?

[*Respondents were presented with a visual analogue scale ranging from 0 to X, with X=10 if respondent had chosen option A in question 2 and X=5 if respondent had chosen option B. The label of the scale was 'years'. Below the scale the following text was printed: "I find both options equally good* *when I'd live another … years in option B" (with the scale value displayed on the dots, at 1 decimal level)*].

[*After the respondent indicated a value, the following question was asked.*]

Do you really find living 10 years in poorer health (Option A) equally good as living Y years in perfect health (Option B)?

☐ Yes [*If yes, skip to next question.*]
☐ No [*If no, back to question 3.*]

7.

An exploration of differences between Japan and two European countries in the self-reporting and valuation of pain and discomfort on the EQ-5D

Based on: Y. Feng, M. Herdman, F.E. van Nooten, C. Cleeland, D. Parkin, S. Ikeda, A. Igarashi and N.J. Devlin

An exploration of differences between Japan and two European countries in the self-reporting and valuation of pain and discomfort on the EQ-5D

Quality of Life Research 2017; 26(8): 2067–2078.
DOI 10.1007/s11136-017-1541-5

Summary

Purpose: To investigate the systematic differences in the self-reporting and valuation of overall health and, in particular, pain/discomfort between three countries (England/UK, Japan, and Spain) on the EQ-5D.

Methods Existing datasets were used to explore differences in responses on the EQ-5D descriptive system between Japan (3L and 5L), the UK (3L), England (5L), and Spain (5L), particularly on the dimension of pain/discomfort. The role of different EQ dimensions in determining self-reported overall health scores for the EuroQol visual analogue scale (EQ-VAS) was investigated using ordinary least squares regression. Time trade-off (TTO) results from Japanese and UK respondents for the EQ-5D-3L as well as Japanese and English respondents for the EQ-5D-5L were compared using *t* tests.

Results: For the EQ-5D-3L, a higher percentage of respondents in Japan than in the UK reported ‘no pain/discomfort’ (81.6 vs 67.0%, respectively); for the EQ-5D-5L, the proportions were 79.2% in Spain, 73.2% in Japan, and 63–64% in England, after adjusting for age differences in samples. The ‘pain/discomfort’ dimension had the largest impact on respondents’ self-reported EQ-VAS only for EQ-5D-3L in Japan. Using the EQ-5D-3L, Japanese respondents were considerably less willing to trade off time to avoid pain/discomfort than the UK respondents; for example, moving from health state, 11121 (some problems with pain/discomfort) to 11131 (extreme pain/discomfort) represented a decrement of 0.65 on the observed TTO value in the UK compared with 0.15 in Japan. Using the EQ-5D-5L, Japanese respondents were also less willing to trade off time to avoid pain/discomfort than respondents in England; however, the difference in values was much smaller than that observed using EQ-5D-3L data.

Conclusions: This study provides evidence of between country differences in the self-reporting and valuation of health, including pain/discomfort, when using EQ-5D in general population samples. The results suggest a need for caution when comparing or aggregating EQ-5D self-reported data in multi-country studies.

7.1 Introduction

Patient-reported outcomes (PROs) are widely used as outcome measures in clinical trials of pain treatments. Indeed, given that pain can *only* be measured subjectively, studies of pain are entirely reliant on self-reporting (Cleeland *et al.*, 2011). The assessment of pain is incorporated in different ways in PROs, ranging from generic measures of health-related quality of life, such as EQ-5D (Kind *et al*., 2005) in which pain/discomfort is only one of five dimensions measured, to tools that are specific to the assessment of pain, such as the Brief Pain Inventory (Cleeland & Ryan, 1994). One important consideration when using such measures in multi-country studies is that factors such as race, ethnicity, language, and culture can potentially affect responses to PRO instruments (Alonso *et al.*, 1998; Scott *et al.*, 2006; Stewart *et al.*, 2000).

With regard to pain reporting, there is evidence that this can vary quite widely across countries; for example, the results of a survey published in 2014 showed that pain reporting and treatment rates were lower in China (6.2% and 28.3%, respectively) and Japan (4.4% and 26.3%, respectively) than in the other countries involved (≥14.3% and 35.8%, respectively (Goren *et al*., 2014). Substantial variations between countries in the rate of pain reporting have also been reported in primary care (Gureje *et al*., 1998) and among patients with cancer (Wang *et al*., 2010). Within Europe, studies have also shown variation in pain reporting; using the pain/discomfort dimension of the EQ-5D, rates of respondents declaring no pain/discomfort varied from 65% in France to 79.5% in Spain (König *et al*., 2010). At least some of the difference between countries may depend on cultural differences in pain response rather than on differences in objective levels of pain. For example, in experimental studies, Japanese subjects provided lower pain ratings for equivalent 'objective' levels of pain than European subjects (Komiyama *et al*., 2009), whereas other studies have shown that Euro-Americans consider seeking pain relief more acceptable than Japanese respondents (Hobara *et al.*, 2005). However, evidence from Uki *et al.* (1998) suggests that Japanese patients with cancer reported high levels of pain with inadequate pain management.

The EQ-5D is one of the most widely used generic preference-based measures of health status. It has been translated into numerous languages and is available in an increasing number of country-specific utility-based value sets. Nevertheless, little attention has been paid to whether the effects described earlier in relation to pain affect the self-rating and valuation of pain/discomfort on EQ-5D and whether such data can be compared and aggregated across countries. In one of the few such analyses performed, Tsuchiya et al. compared the results of the UK and Japanese EQ-5D-3L valuation data and noted that the two data sets were positively correlated (Tsuchiya *et al*., 2002). Furthermore, Japanese time trade-off (TTO) values were consistently higher than those from the UK, except for mild states.

The availability of self-reported and valuation EQ-5D-3L data from Japan and the UK together with self-reported and valuation data from England and Japan for the latest version of the instrument, the EQ-5D-5L, makes it possible to explore these questions in more depth. Self-reported data were also available for the EQ-5D-5L from Spain. The primary aim of this study was to investigate whether available EQ-5D data provide evidence of systematic differences in the way respondents in Japan self-report and value health using EQ-5D compared with respondents from England/UK and Spain, with a particular focus on the pain/discomfort dimension.

7.2 Methods

The EQ-5D

The EQ-5D consists of a descriptive system and a visual analog scale (EQ-VAS). Respondents rate their health on the EQ-5D descriptive system and assess their overall health on the EQ-VAS.

There are two versions of the instrument for use in the adult population, EQ-5D-3L and EQ-5D-5L. Both measure health using a descriptive system with five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). The original version of EQ-5D (EQ-5D-3L) uses three levels of severity (no, some, extreme problems or unable to) in each dimension. To increase the instrument's sensitivity, a new version of the instrument (EQ-5D-5L) was developed with five levels of severity (no, slight, moderate, severe, extreme problems or unable to) (Herdman *et al*., 2011). Self-ratings on the descriptive system are summarized as a five number 'code' where each number reflects the severity level on the individual dimensions. Each 'code' represents a unique health state, with the state 11111 representing full health. There are 243 EQ-5D-3L health states ($243=3^5$) and 3,125 EQ-5D-5L health states ($3{,}125=5^5$). By selecting one level of severity in each dimension on either version of the EQ-5D, respondents assign themselves one out of all possible health states as a description (known as a 'profile') of their own health.

The EQ-VAS consists of a vertical scale with anchor points of 0 (worst possible health) and 100 (best possible health). The respondent marks a point on the scale to show how they perceive their overall health.

In addition to providing these two types of self-reported data, the EQ-5D is frequently used in conjunction with health state valuation techniques such as TTO or discrete choice experiments to generate preference-based societal weights for each of the individual health states generated by the descriptive system (Oppe *et al*., 2014). It is recommended that weights be obtained for individual countries, as values assigned to health states might differ between countries for cultural or other reasons (Szende *et al.*, 2007).

Data

Valuation and self-reported health data obtained from the EQ-5D-3L valuation surveys in the UK and Japan as well as EQ-5D-5L valuation surveys in England and Japan were use (Dolan *et al*, 1997; Devlin *et al.*, 2016; Ikeda *et al*., 2015; Tsuchiya *et al.*, 2002). At the time the present study was carried out, self-reported health data on EQ-5D-5L were also available for Spain (Ramos-Goni *et al.*, 2014). In all cases, the samples included were intended to be representative of the general population of the country.

Socio-demographic and health-related characteristics available in each of the five data sets included age, gender, and respondents' experience of serious illness (in themselves, a family member, or others). The same background characteristics were recorded in each of the five data sets.

The Japanese EQ-5D-3L valuation study was a quasi-replication of the UK valuation study (Tsuchiya *et al.*, 2002). Each respondent valued the same set of the 17 health states, which were a subset of the 42 health states in the UK study (Dolan, 1997). The EQ-5D-5L valuation data used in the present analysis were collected during valuation studies carried out in Japan, England and Spain using the EuroQol Valuation Technology (EQ-VT) software, which was developed specifically for the EQ-5D-5L value set studies (Oppe *et al*., 2014). An identical methodology based on computer-assisted personal interviews (CAPI) was used in all countries. Each respondent was asked to provide TTO values for a block of 10 health states out of the 86 health states selected for direct valuation. All interviewers received training on administration of the CAPI and EQ-VT. Only the observed TTO values (rather than the value sets modelled from those data) were used in this study.

A summary of the characteristics of each survey is provided in Table 7.1.

Table 7.1 Characteristics of the five survey data sets

Survey places	EuroQol version	Self-rating of EQ-5D profile and EQ-VAS	Valuation data	No. of health states in the TTO	Background variables[a]	Year of survey	No. of respondents	Exclusion criteria for the valuation study	No. of respondents in the valuation data set[b]	Software[c]	Health state valuation method
Japan	EQ-5D-3L	Available	Available	1998	621	1998	621	Respondents completely missing TTO data; valued 1 or 2 states only; giving all states the same value; valued all states worse than dead	543	STATA	Conventional TTO
UK	EQ-5D-3L	Available	Available	42	Available	1993	3395	Missing in valuation data	2997	LIMDEP	Conventional TTO
Japan	EQ-5D-5L	Available	Available	86	Available	2014	1098	Three of the investigators (involving 72 respondents) did not follow the survey manual procedures	1026	STATA, SAS, and R	Lead time TTO
England	EQ-5D-5L	Available	Available	86	Available	2012–2013	996	Respondents who gave the same TTO value for all health states; giving health state 55555 a value no lower than the value for the mildest health state in their block	912	R and WinBugs	Lead time TTO and DCE
Spain	EQ-5D-5L	Available	Available	86	Available	2012	1000	N/A	N/A	N/A	N/A

[a] The background variables include age, gender, experience with serious illness by respondents themselves, their families, and others.
[b] Represents the number of respondents after applying the exclusion criteria. For the Japanese EQ-5D-3L data set, this study only has access to the sample that is included in the valuation study (N=543).
[c] The software used for the original analysis in the valuation studies.
d 'Available' and 'N/A' refer to the availability of data for this study.
N/A: not available; TTO: time trade-off; DCE: Discrete Choice Experiments.

Statistical analysis

There were five parts to the statistical analysis. First, respondents' socio-demographic and health-related characteristics were compared across the five data sets. Second, self-reported data on the descriptive systems of EQ-5D-3L and EQ-5D-5L were compared to investigate whether there were systematic differences between countries that could not be explained by differences in sample characteristics on age or gender, particularly in regard to reporting of pain/discomfort. Third, the impact of the five dimensions of the descriptive system on EQ-VAS scores was analyzed to determine, in particular, the contribution of the pain/discomfort dimension to VAS scoring. Fourth, valuation data were used to explore whether respondents in Japan and the UK for the EQ-5D-3L valuation studies and respondents in Japan and England for the EQ-5D-5L valuation studies have different stated preferences in terms of their willingness to trade off time in TTO tasks, particularly in relation to health states involving pain/discomfort. Fifth, the linked self-reported and valuation data sets for the EQ-5D-3L and EQ-5D-5L valuation studies in the UK/England and Japan were analyzed to investigate whether there was any relationship between respondents' self-reported pain/discomfort and the TTO values they assigned to the hypothetical health states.

Table 7.2 Socio-demographic and health-related characteristics of respondents in Japan and the UK for the EQ-5D-3L data; Japan, England and Spain for the EQ-5D-5L

	EQ-5D-3L			EQ-5D-5L			
Respondent characteristic	**Japan (n=543)**	**UK (n=3395)**	**P-value**	**Japan (n=1026)**	**England (n=996)**	**Spain (n=1000)**	**P-value**
Age, years, mean (SD)	48.1 (15.3)	47.9 (18.4)	0.753	44.9 (14.9)	51.2 (17.9)	43.8 (17.3)	0.000 (Spain, England) 0.000 (Japan, England) 0.119 (Japan, Spain)
Gender (%)							
Female	230 (42.4)	1926 (56.7)	0.000	511 (49.8)	591 (59.3)	525 (52.5)	0.000
Male	313 (57.6)	1469 (43.3)		515 (50.2)	405 (40.7)	475 (47.5)	
Experience of serious illness (self)[b] (%)							
Yes	80 (14.7)	1076 (31.7)	0.000	192 (18.7)	330 (33.1)	144 (14.4)	0.000
No[a]	463 (85.3)	2319 (68.3)		834 (81.3)	666 (66.9)	856 (85.6)	
Experience of serious illness (family)[c] (%)							
Yes	188 (34.6)	2156 (63.5)	0.000	377 (36.7)	692 (69.5)	633 (63.3)	0.000
No[a]	355 (65.4)	1239 (36.5)		649 (63.3)	304 (30.5)	367 (36.7)	

	EQ-5D-3L			EQ-5D-5L			
Respondent characteristic	Japan (n=543)	UK (n=3395)	P-value	Japan (n=1026)	England (n=996)	Spain (n=1000)	P-value
Experience of serious illness (other)[d] (%)							
Yes	178 (32.8)	547 (16.1)	0.000	130 (12.7)	416 (41.8)	347 (34.7)	0.000
No[a]	365 (67.2)	2848 (83.9)		896 (87.3)	580 (58.2)	653 (65.3)	

SD: Standard Deviation

[a] *This category includes missing values.*

[b] *Respondents who had experienced serious illness in themselves.*

[c] *Respondents who had experienced serious illness in a family member.*

[d] *Respondents who had taken care of others with a serious illness.*

The socio-demographic and health-related characteristics of the different samples were compared using *t*-tests for age, and chi-squared tests for gender and the proportions of respondents who reported having experienced serious illness in themselves, a family member, or others.

Respondents' self-reported EQ-5D data were analyzed by comparing the distribution of EQ-5D profile data by country and instrument version. The ceiling effect (measured by the proportion of respondents reporting the best possible health for EQ-5D), the number of EQ-5D profiles used, and the distribution of responses by dimension were also calculated and compared across countries and EQ-5D versions. Adjustments for age and gender were made when comparing distributions on the descriptive system across countries. Using England as an example, to adjust for age difference, five age ranges were used (≤ 30 years, 31–45 years, 46–60 years, 61–75 years, and >75 years). For each age range, the proportion of respondents in England who reported full health was calculated followed by the weighted average of the five proportions. The weight for each age band is the proportion of respondents in that age band in Spain or Japan. Adjustments for any differences in gender distributions were made in the same way.

The extent to which responses on the five dimensions of the descriptive system explained self-reported overall health on the EQ-VAS was examined using ordinary least squares (OLS) regression methods and the results were compared between countries for both the 3L and 5L. Data on the five dimensions of the EQ-5D were recorded as continuous variables (1 for level 1, 2 for level 2, 3 for level 3, 4 for level 4, and 5 for level 5). To show the model's goodness of fit, adjusted R-squared and results from residual analysis were reported after each regression.

To explore whether there were differences in TTO values in the valuation studies between respondents, in particular for pain/discomfort, EQ-5D-3L and EQ-5D-5L valuation data from the UK/England and Japan were analyzed. Respondents in the UK and Japan in the EQ-5D-3L studies yielded observed values for 42 health states in the UK and 17 health states in Japan. The EQ-5D-5L studies in Japan and England yielded observed values for 86 health states. TTO values in each version of the EQ-5D for those health states were compared between the two countries using *t* tests.

Among the 17 hypothetical health states in the EQ-5D-3L valuation data, there were three pairs of health states that *only* differed on the pain/discomfort dimension. Among the 86 hypothetical health states in the EQ-5D-5L valuation data, there were seven pairs of health states that *only* differed on the pain/discomfort dimension. For each version of the EQ-5D valuation studies, the difference in mean TTO values was compared between respondents by country (for those pairs that only differed in the pain/discomfort dimension) to gain insight into differences in how respondents in Japan and the UK/England value EQ-5D health states with respect to the pain/discomfort dimension.

Finally, respondents' self-reported and valuation data were linked using respondents' ID for the EQ-5D-3L and EQ-5D-5L data from Japan and the UK/England. Using both the self-reported and valuation data, the effect of self-reported pain/discomfort in explaining the TTO values was explored by using the OLS regression analyses to model the TTO values for the five dimensions of the EQ-5D profile and self-reported pain/discomfort. All analyses were performed using STATA/MP 13.

7.3 Results

There was no statistically significant difference in mean age between respondents in the UK and Japan in the EQ-5D-3L data (P=0.753), although the UK sample had a higher proportion of females and respondents who had experienced serious illness in themselves or in a family member, and a lower proportion of respondents who had taken care of others with a serious illness (P<0.05). There was no statistically significant difference in mean age between respondents in Japan and Spain in the EQ-5D-5L data sets (P=0.119), however respondents in England were older than respondents in Japan and Spain (P<0.05). The English sample also reported the highest proportion of females and respondents who had experienced serious illness themselves, in a family member, or who had taken care of others with a serious illness (P<0.05) (Table 7.2).

The most frequently self-reported profile using both versions of the EQ-5D was full health. The proportion of respondents reporting full health was highest in Japan for both versions of the EQ-5D and the differences between other countries were statistically significant (P<0.05). After adjusting for differences in age and gender in the EQ-5D-5L samples, the proportion of those reporting full health was still highest among respondents in Japan (66.5%), followed by Spain (54.9%) and England (53.8%). The reduction in the ceiling effect using the EQ-5D-5L compared with the EQ-5D-3L was similar in the UK/England (from 56.9% to 47.6% for the EQ-5D-3L and EQ-5D-5L, respectively) and Japan (from 77.2% to 66.5%).

Analysis of self-reported EQ-5D profile data showed that Japanese respondents employed a much smaller number of health profiles than respondents in other countries. Among the EQ-5D-3L data sets, three health states accounted for 90.1% of Japanese respondents compared with 12 health states in the UK (90.6%). The difference is even more marked for the EQ-5D-5L, where only four health states accounted for 91.4% of Japanese respondents compared with 80 health states in England (90.1%) and 16 health states in Spain (90.0%). Full EQ-5D profile distributions by country and EQ-5D version are available upon request from the authors.

Figure 7.1 shows the proportion of respondents reporting level one for the EQ-5D-3L by dimension and country. In the pain/discomfort dimension, 81.6% of respondents in Japan self-reported level one, compared with 67.0% of respondents in the UK.

Figure 7.1 The proportion of respondents in Japan and the UK who reported level one (no problems) by EQ-5D-3L dimension.

MO, mobility; SC, self-care; UA, usual activities; PD, pain/discomfort; AD, anxiety/depression.

Figure 7.2 shows that more respondents in Japan self-reported level one for the EQ-5D-5L than in England and Spain in four dimensions. In the pain/discomfort dimension, a higher proportion of respondents in Spain (79.2%) self-reported level one than in Japan (73.2%) and England (58.4%), although after adjusting for differences in age and gender between the samples, the proportion of respondents reporting level 1 on the pain/discomfort dimension in England was approximately 64%.

The results of modelling respondents' EQ-VAS scores as a function of their self-reported EQ-5D profiles are shown in Table 7.3. EQ-VAS scores decreased when the severity of problems increased in any of the five dimensions, and this finding was consistent across countries and EQ-5D versions. Using EQ-5D-3L, pain/discomfort was the most important dimension in explaining respondents' self-reported EQ-VAS scores in Japan (on average, a one level increase in the pain/discomfort dimension led to a decrease of 11.03 points on the EQ-VAS); whereas in the UK, it was the dimension of usual activities.

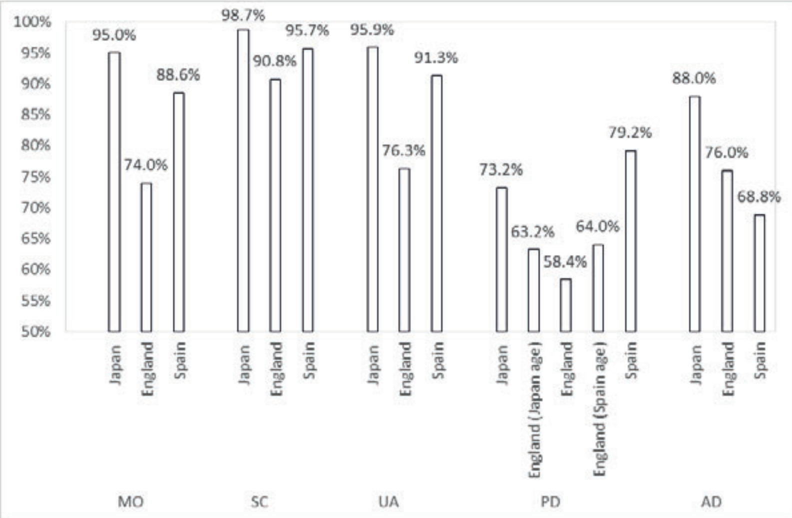


Figure 7.2 The proportion of respondents in Japan, England, and Spain reporting level one (no problems) by EQ-5D-5L dimension.
MO, mobility; SC, self-care; UA, usual activities; PD, pain/discomfort; AD, anxiety/depression. Data on the pain/discomfort dimension for England includes those adjusted by age distribution of Japan (Japan age) and Spain (Spain age).

The results of modelling respondents' EQ-VAS scores as a function of their self-reported EQ-5D profiles are shown in Table 7.3. EQ-VAS scores decreased when the severity of problems increased in any of the five dimensions, and this finding was consistent across countries and EQ-5D versions. Using EQ-5D-3L, pain/discomfort was the most important dimension in explaining respondents' self-reported EQ-VAS scores in Japan (on average, a one level increase in the pain/discomfort dimension led to a decrease of 11.03 points on the EQ-VAS); whereas in the UK, it was the dimension of usual activities. Using EQ-5D-5L data, anxiety/depression was the most important dimension in explaining differences in EQ-VAS scores in England and Spain; however, in Japan, the most important dimension was usual activities. The adjusted R-squared and results from residual analyses for each model are reported in Table 3. The results from residual analyses suggest no evidence of multicollinearity in the five models. However, there is evidence of residuals with non-normal distributions, heteroscedasticity, and non-linear functional form for some specifications.

Table 7.3 Modelling self-reported EQ-VAS scores by country and EQ-5D version

	Japan EQ-5D-3L		UK EQ-5D-3L		Japan EQ-5D-5L		England EQ-5D5L		Spain EQ-5D-5L	
	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value
Mobility	–0.88	0.785	–5.87	0.000	–2.87	0.029	–3.46	0.000	–0.35	0.719
Self-care	–5.42	0.394	–6.82	0.000	–4.41	0.128	–0.21	0.836	–3.29	0.065
Usual-activities	–9.50	0.018	–8.54	0.000	–9.61	0.000	–4.87	0.000	–5.86	0.000
Pain/ discomfort	–11.03	0.000	–6.80	0.000	–4.53	0.000	–3.28	0.000	–4.59	0.000
Anxiety/ depression	–7.22	0.003	–8.02	0.000	–5.83	0.000	–6.96	0.000	–7.87	0.000
Constant	115.80	0.000	125.80	0.000	114.94	0.000	105.99	0.000	109.82	0.000
Adjusted R-squared[a]	0.2027		0.4226		0.2285		0.5031		0.4026	
N	543		3372		1026		996		1000	

Note: (1) Shapiro–Wilk test rejected the null hypothesis of normally distributed residuals in all five regressions at the 5% significance level. (2) The Breusch–Pagan test accepted the null hypothesis that the variance of the residuals is homogeneous at the 5% significance level for the Japan EQ-5D-3L model, but rejected the null hypothesis in all other four models. (3) The Regression Equation Specification Error Test (RESET) rejected the null hypothesis of linear functional form for the specifications in Japan EQ-5D-5L model and Spain EQ-5D-5L model at the 5% significance level. (4) None of the five models indicated issues with multicollinearity (mean variance inflation factor <3).

[a] Adjusted R-squared is computed using the formula 1 - [(1 – Rsq) × (N – 1)/(N – k – 1)], where Rsq represents R-squared (the proportion of variance in the dependent variable that can be explained by the independent variables); N represents the number of observations and k represents the number of independent variables.

Table 7.4 reports the mean EQ-5D-3L TTO values for the 17 hypothetical health states valued by respondents in the UK and Japan. In Japan, none of the mean TTO values for the 17 health states were below zero (i.e. none of them were considered as being worse than dead). By contrast, six of the 17 health states in the UK had negative mean TTO values. TTO values for the five mildest health states (11112, 11121, 11211, 12111, 21111), were lower in Japan than in the UK, with the differences being statistically significant at the 5% significance level. However, Japanese TTO values for the remaining 12 more severe health states were all higher than UK values for those health states. The differences were statistically significant at the 5% significance level for 11 health states, but not for health state 22222.

Table 7.4 Comparing the mean TTO values for the 17 hypothetical EQ-5D-3L health states between respondents in Japan and the UK

Health state	Japan	UK	*t*-statistics
11112	0.79	0.83	–2.90*
11121	0.79	0.85	–5.13*
11211	0.82	0.87	–4.73*
12111	0.81	0.83	–1.90*
21111	0.78	0.88	–8.05*
11113	0.71	0.39	12.36*
11131	0.64	0.20	15.81*
11312	0.64	0.55	4.01*
11133	0.54	–0.05	20.37*
13311	0.60	0.35	9.74*
32211	0.33	0.15	5.90*
22222	0.51	0.50	0.41
23232	0.41	–0.09	17.54*
32223	0.22	–0.17	13.86*
32313	0.21	–0.15	12.83*
33323	0.09	–0.39	18.81*
33333	0.01	–0.54	29.15*

* *Significant at the 5% significance level.*

Table 7.5 reports the mean EQ-5D-5L TTO values for the 86 hypothetical health states valued by respondents in Japan and England. For respondents in both countries, only the worst state 55555 was assigned a mean TTO value below zero. Mean TTO values were higher in Japanese respondents for 63 of the 86 hypothetical health states, with 19 of those differences being statistically significant (P<0.05).

Table 7.5 Comparing the mean TTO values for the 86 hypothetical EQ-5D-5L health states between respondents in Japan and England

Health state	Japan	England	t-statistics	Health state	Japan	England	t-statistics
11112	0.91	0.85	2.41*	31524	0.46	0.45	0.18
11121	0.90	0.89	0.71	31525	0.39	0.43	– 0.58
11122	0.86	0.79	2.15*	32314	0.52	0.51	0.23
11211	0.91	0.89	0.83	32443	0.45	0.29	2.48*
11212	0.81	0.82	-0.23	33253	0.41	0.40	0.04
11221	0.85	0.84	0.13	34155	0.36	0.24	1.70
11235	0.60	0.53	1.24	34232	0.53	0.55	– 0.45
11414	0.60	0.41	3.23*	34244	0.34	0.26	1.23
11421	0.72	0.65	1.49	34515	0.30	0.32	–0.26
11425	0.54	0.53	0.09	35143	0.40	0.27	2.02*
12111	0.89	0.87	1.26	35245	0.30	0.18	1.89
12112	0.82	0.81	0.13	35311	0.60	0.51	1.60
12121	0.87	0.81	1.72	35332	0.38	0.59	–3.28*
12244	0.50	0.32	2.73*	42115	0.48	0.41	1.14
12334	0.61	0.44	2.55*	42321	0.59	0.54	0.89
12344	0.51	0.25	3.63*	43315	0.39	0.42	–0.37
12513	0.61	0.61	-0.01	43514	0.31	0.36	–0.70
12514	0.53	0.44	1.51	43542	0.33	0.23	1.50
12543	0.47	0.32	2.22*	43555	0.13	0.06	1.05
13122	0.75	0.81	–1.34	44125	0.38	0.32	0.78
13224	0.61	0.49	2.14*	44345	0.23	0.21	0.17
13313	0.63	0.69	–1.13	44553	0.17	0.09	1.06
14113	0.70	0.69	0.31	45133	0.49	0.36	1.85
14554	0.31	0.15	2.18*	45144	0.32	0.17	2.39*

Health state	Japan	England	t-statistics	Health state	Japan	England	t-statistics
15151	0.54	0.42	2.08*	45233	0.36	0.33	0.43
21111	0.90	0.89	0.37	45413	0.32	0.34	–0.32
21112	0.81	0.83	–0.39	51152	0.36	0.35	0.19
21315	0.60	0.54	1.02	51451	0.33	0.26	1.08
21334	0.54	0.50	0.50	52215	0.40	0.35	0.78
21345	0.42	0.43	–0.18	52335	0.32	0.33	–0.09
21444	0.40	0.15	3.93*	52431	0.43	0.54	–1.79
22434	0.45	0.53	–1.24	52455	0.15	0.07	1.07
23152	0.49	0.39	1.42	53221	0.58	0.58	–0.03
23242	0.52	0.44	1.30	53243	0.36	0.23	1.94
23514	0.54	0.40	2.29*	53244	0.26	0.12	2.00*
24342	0.46	0.36	1.48	53412	0.36	0.44	–1.23
24443	0.38	0.33	0.80	54153	0.28	0.27	0.22
24445	0.30	0.16	2.13*	54231	0.40	0.40	–0.05
24553	0.22	0.33	–1.49	54342	0.34	0.18	2.17*
25122	0.55	0.52	0.40	55225	0.21	0.17	0.52
25222	0.57	0.59	–0.54	55233	0.27	0.28	–0.06
25331	0.56	0.53	0.47	55424	0.17	0.25	–1.05
31514	0.45	0.39	0.98	55555	–0.02	–0.08	2.69*

* *Significant at the 5% significance level.*

No clear pattern was observed between the severity of health states and the presence of higher values from Japanese respondents; the 19 health states included mild states (e.g. 11112) and severe states (e.g. 55555). In only one case was a statistically significant higher value observed for English raters (state 35332).

Three pairs of health states in the EQ-5D-3L valuation data in the UK and Japan differed only on the pain/discomfort dimension. Analysis of those states showed that Japanese respondents traded off less time to avoid problems in the pain/discomfort dimension than UK respondents. The biggest between-country difference in mean TTO values was reported between health states 11121 and 11131, which would represent a decrease of 0.65 in TTO values in the UK compared with 0.15 in Japan.

Seven pairs of health states in the EQ-5D-5L in the English and Japanese valuation studies differed only in the pain/discomfort dimension. Specifically, four pairs differed between level 1 (no problem) and level 2 (mild problem), two pairs differed between level 1 (no problem) and level 4 (severe problem), while one pair differed between level 3 (moderate problem) and level 4 (severe problem). Comparing the mean TTO values between respondents in the two countries showed that Japanese respondents traded off either similar or less time to avoid problems in pain/discomfort than English respondents. The biggest difference in mean TTO values was reported between health states 12334 and 12344, which would represent a decrease of 0.19 in TTO values in England compared with 0.10 in Japan.

Regression analysis of the linked self-reported and valuation data sets showed that respondents' self-reported pain/discomfort was not significant in explaining the TTO values in Japan for EQ-5D-3L (P=0.395) and EQ-5D-5L (P=0.299), nor the UK for EQ-5D-3L (P=0.159). However, it has significantly positive effect in explaining the EQ-5D-5L TTO values in England (P<0.05).

7.4 Discussion

This is the first study to carry out an in-depth examination of the comparability of EQ-5D self-rated health status and valuation data from Japanese and European respondents, with a particular focus on pain/discomfort. A number of findings were clear from the empirical analyses in this study.

First, respondents in Japan tend to report better health in general than respondents in England/UK and Spain. Second, with respect to pain/discomfort, respondents in Japan reported problems less frequently than respondents in England/UK, but slightly more frequently than respondents in Spain. Third, Japanese respondents used a much smaller number of health states to describe their health than respondents in either of the other two countries, and Spanish respondents also used substantially fewer health states than respondents in England. Fourth, in the EQ-5D-3L valuation study, respondents in Japan were more willing to trade off time for the mildest health states, but less willing to trade off time for the severe health states compared with respondents in the UK. For nearly three-fourths of the EQ-5D-5L health states for which values were obtained, Japanese respondents' values were higher than those from English respondents. However, in contrast with EQ-5D-3L values, there was no clear pattern between this and the severity of the states. Fifth, in the EQ-5D-3L and EQ-5D-5L valuation studies, Japanese respondents were willing to trade off less time to avoid problems in the pain/discomfort dimension than respondents in England/UK. However, the differences in TTO values between respondents in Japan and England are much smaller than in the EQ-5D-3L valuation study.

It is not clear where these differences stem from, though similar findings have been reported previously. For example, in a comparison of EQ-5D results from 20 countries in a diabetes clinical trial, researchers found substantial variation in the reporting of functional health problems, but noted that the variation could not be explained by differences in demographic variables, clinical risk factors, or rates of complications (Salomon *et al.*, 2011). They suggested that the unexplained variability meant there were important problems of comparability across settings.

One possible cause of the differences found here is that the way terms used to describe health, e.g., the severity labels, varies across countries. For example, Luo et al. found that the interpretation and use of EQ-5D-5L response labels (e.g., 'slight', 'moderate', and 'severe') varied across Chinese, Malay, and English speakers in Singapore (Luo *et al.*, 2015a), whereas the English version gave similar outcomes in Chinese and non-Chinese English speakers in the same country (Luo *et al.*, 2015b), suggesting that there was no effect of culture on responses. Although a strict protocol is followed in producing other language versions of EQ-5D (Rabin *et al.*, 2014), it may not always be possible to find identical terms in all languages. There is also evidence suggesting that Japanese respondents might be less willing to report pain than those in Europe, possibly due to a tendency within the Japanese culture for pain to be repressed and controlled rather than shared or expressed (Shirado *et al.*, 1976).

Our findings on the reporting of pain/discomfort coincide with those of earlier multi-country studies that showed a tendency towards lower rates or intensity of self-reported pain in Japan than in other countries (Goren *et al.*, 2014; Gureje *et al.*, 1998; Tsang et al., 2008). Despite lower rates of self-reported pain/discomfort in Japan, we found that this was the most important dimension in explaining respondents' self-reported EQ-5D-3L VAS scores. There are two possible explanations for this discrepancy. First, while the EQ-VAS scores and self-reported EQ-5D profiles measure how good or bad respondents rate their currently experienced health status, the TTO valuation task evaluates health states that are hypothetical to the respondents. Second, the tasks involved EQ-VAS scores rating and TTO valuation, which are individually very different. It is possible for a respondent to rate a health state as poor on the EQ-VAS, but still not be willing to trade off any life years to avoid it (e.g., because of religious beliefs about the sanctity of life, being the primary caregiver to a small child, or having a very low personal discount rate). As the EQ-5D profile variables are treated as continuous, the importance of each EQ-5D dimension in explaining the EQ-VAS reflects the average effect in a dimension between two neighboring levels. An alternative approach is to treat the EQ-5D profile variables as dummies (i.e., one for each level and dimension). However, this would leave some categories with rather small sample sizes, particularly for severe levels (n < 5). Those results from the EQ-5D-3L data were not confirmed using EQ-5D-5L data. Given the design of the current study, it was not possible to determine whether the difference in findings was due to changes in methods or perceptions of the importance of pain/discomfort over time. It should also be noted that the lowest rates of pain/discomfort were observed in Spain. Other studies have also reported relatively low rates of self-reported problems on the EQ-5D descriptive system in the general population in Spain compared with other European countries although not on the EQ-VAS (Konig *et al.*, 2010). Similar findings have been reported for Spain using other instruments, such as the Brief Pain Inventory (Breivik *et al.*, 2006).

The comparison of valuation data also showed difference between countries. The Japanese EQ-5D-3L valuation data showed a tendency to compress towards the middle of the scale. A mid-range response style and lower levels of extreme response style have been reported in some studies in Japanese subjects (Harzing *et al.*, 2009; Stening *et al.*, 1994; Zax *et al.*, 1967), though it is not clear whether such an effect may also be present in valuation studies. Furthermore, we found that respondents in Japan were less willing to trade off time to avoid pain/discomfort on the EQ-5D-3L than respondents in the UK. It should be noted that the 17 health states valued by the Japanese respondents were a subset of the health states in the UK valuation study. The TTO values may be influenced by the mix of severity in the set of states presented. This may affect observed differences in TTO values.

This compression of values in Japan relative to UK values was no longer observed when analyzing results from the EQ-5D-5L. Only the worst state (i.e., 55555) was rated worse than death in both countries. Almost three-quarters of the EQ-5D-5L health states were given higher values by Japanese respondents compared with English respondents, indicating more reluctance to trade off time among Japanese respondents. However, unlike with the EQ-5D-3L data, these higher ratings were spread across all levels of severity. It is possible that these differences are due to changes in the methods that were used in the TTO valuation tasks between the two versions of EQ-5D, or changes in perceptions of health states, and/or relative importance assigned to different dimensions over time. However, it is not possible to answer it definitively here.

7.5 Limitations

Ideally, samples from Japan and European countries used to explore differences in the self-reporting and valuation of pain would have identical distributions for all factors that might influence results. However, it was not possible to control for all relevant variables, although we controlled for the effects of age and gender. It remains unclear whether differences in the rates of health problems between countries and other unobserved characteristics may have led to the differences we observed.

Furthermore, the EQ-5D-3L data used in this study were collected in the 1990s and may no longer be applicable to the present populations in the UK and Japan; however, we do not consider this to be a limitation of the present analysis as we were interested in comparing results between countries and not in exploring whether the data collected then would be relevant today. The fact that it was possible to compare findings from two different variants of the instrument at two different time points could in fact be considered a strength of the study because it gives an indication of the robustness of the results.

The possible misspecifications for modeling EQ-VAS scores by the EQ-5D dimensions should be noted. Although the assumption of normality does not hold in the five models and, as a consequence, will have an impact on the P values, the estimated coefficients themselves will still be consistent. Four models reported heteroscedasticity in the residuals and two models reported non-linear functional forms. These misspecifications might be explained by variables that are not included in modeling the EQ-VAS, but have an impact on the EQ-VAS, such as other health dimension(s) that are not covered by EQ-5D. If those variables are correlated with the EQ-5D dimensions, our estimated coefficients could be biased. Similar issues have been observed in previous studies (Dolan, 1997; Burström *et al.*, 2014; Sun *et al.*, 2015).

Finally, only data from two European countries were available for the present study and it is not clear whether results can be extrapolated to respondents in other Western countries.

7.6 Implications

Our findings have a number of implications. First, care should be taken when comparing and aggregating clinical data on pain between different countries, because respondents may use different criteria when responding, which could potentially lead to the same treatment being more or less effective in different countries. Second, the differences between respondents in Japan and European countries in self-reported and valuation behaviors could have a substantial effect on the results of cost–utility analyses. For instance, while applying the EQ-5D-3L instrument, the compression of values on the utility scale and the better baseline pain scores observed in Japan may result in relatively small improvements with treatment. Third, what constitutes a minimally important difference for EQ-5D index may be different between Japan and other countries. Fourth, if the findings related to pain/discomfort also applied to other pain measures used as inclusion criteria for clinical trials, then they might lead to questions about whether identical inclusion criteria for clinical trials are in fact being used across countries.

7.7 Conclusions

This study provides prima facie evidence of differences between Japan, UK/England, and Spain in the self-reporting and valuation of health, including pain/discomfort, when using EQ-5D in general population samples. The findings suggest the need for caution when comparing and/or aggregating EQ-5D data across the countries. Specifically designed studies, including the use of qualitative research and vignette techniques (Au & Lorgelly, 2014), would be helpful in exploring these issues further and confirming the findings.

Acknowledgements

We would like to thank the English, Spanish, and Japanese EQ-5D-5L value set study project teams for data sharing. Glaxo-Wellcome and Pharmacia Upjohn funded the data collection for the EQ-5D-3L Japanese value set study project. The authors are grateful for the advice from Aki Tsuchiya and Juan M. Ramos-Goñi. Funding for editorial support by

SuccinctChoice Medical Communications (Chicago, IL) was provided by Astellas Pharma Inc.

8.

A (latent) class of their own: Response patterns in trading-off quantity and quality of life in TTO exercises

Based on: F. van Nooten, W. B. F. Brouwer, N.J.A. van Exel, K. Houghton, M. van Agthoven and D. Stull

A (latent) class of their own:
Response patterns in trading-off quantity and quality of life in TTO exercises

Value in Health 2017;20(10):1403-1410
DOI 10.1016/j.jval.2017.06.008

Summary

Introduction: Conflicting results regarding associations of TTO valuations with respondent characteristics have been reported, mostly based on regression analyses. Alternative approaches, like Latent Class Analysis (LCA), may add to further understanding of variations in TTO responses.

Aim: To identify whether subgroups of respondents can be identified based on their responses to TTO exercises and to investigate which respondent characteristics are associated with membership of the identified subgroups.

Methods: Members of Dutch general public, aged 18-65 years, completed a web-based questionnaire concerning socio-demographic characteristics, three TTO exercises valuing health states described using the domains of the EQ-5D, and preference for quality versus quantity of life. LCA was used to identify patterns in the responses. Predictive variables were included in the final LCA model to identify the particular respondent characteristics that predict subgroup membership.

Results: The sample consisted of 1,067 respondents. Four latent classes were identified in the responses to TTO exercises. Two were 'high-traders', focusing on quality of life and trading-off a relatively high number of years. The other two were 'low-traders', focusing on length of life. Predictive analyses revealed significant differences between subgroups in terms of age, gender, subjective life expectancy, and preference for quantity over quality of life.

Conclusion*:* We showed that distinct classes of respondents can be discerned in TTO responses from the general public, distinguishing subgroups of low and high traders. More research in this area should confirm our findings and investigate their implications for health state valuation exercises.

8.1 Introduction

The time trade-off (TTO) method is commonly used for health state valuations (e.g. Greiner *et al*., 2006; Lamers *et al*., 2006; Scalone, *et al*., 2013). For instance, the frequently used quality of life measure EQ-5D has several readily available national health state valuations ('tariffs') for the health states described with the instrument, which were derived with the TTO method (e.g. Dolan, 1997; Lamers *et al*., 2006). A TTO exercise typically requires people to choose between two streams of health: one entailing a shorter life span which is spent in a better health state and the other entailing a longer life span spent in a poorer health state. Thus, the exercise requires participants to trade off length of life and quality of life. From this, relative health state valuations are derived, with perfect health anchored at the value of 1 and dead anchored at the value of 0 (Torrance, 1986). Despite being applied since the early 1970's, studies have shown that the answers to TTO exercises are still not fully understood, in particular the heterogeneity of responses and the associated respondent characteristics that may be driving this heterogeneity (Essink-Bot *et al*., 2007; Dolan & Roberts, 2002). Indeed, it is not uncommon to see utilities for a given health state with a wide range, meaning that some individuals are more (or less) willing than others to trade off years of life for that health state. This is true both for general public ratings and for disease-specific ratings (Lubeck *et al*., 2002; Mazur & Merz, 1995).

Several studies have shown that TTO responses may be influenced by basic respondent characteristics, such as age, gender and marital status (Dolan & Roberts, 2002). However, these findings vary between studies and sometimes even contradict (e.g. Augestad *et al*., 2013; Ayalon & King-Kallimanis, 2009; Best *et al*., 2010; Hsu et al., 2009; Kontodimopoulos & Niakas, 2006; Shimizu *et al.,* 2008). Beyond basic respondent characteristics, subjective life expectancy has also been shown to influence TTO responses, but the strength of its influence is limited (Heintz *et al*., 2013; Van Nooten & Brouwer, 2004; Van Nooten *et al*., 2009). Therefore, while there is often substantial variation in TTO responses between respondents, this variability is not yet well understood and further investigation is warranted. In that context, interestingly, Essink-Bot and colleagues suggested that individual response patterns (e.g. a 'general' tendency to give either high or low scores in valuation exercises) might be more influential than demographic or other respondent characteristics (Essink-Bot *et al*., 2007).

Previous research investigating heterogeneity in TTO responses has done so by identifying possible explanatory variables and testing for differences.

This approach requires an a priori understanding or hypothesis regarding what may be causing the variability. In addition, it also requires that collection of all potential explanatory variables such that they can be used analytically. If the heterogeneity is driven by many variables, uncovering subgroups using multiple comparisons may introduce multiplicity problems and can be prohibitively time-consuming (Stull & Houghton, 2013). In addition, if information on a key explanatory variable is not collected then true subgroups may go undiscovered. To more precisely uncover true subgroups that exist in heterogeneous TTO data, latent class analysis (LCA) is ideally suited. LCA is an analytic method that allows for the identification of subgroups of respondents by looking for common patterns of response within the heterogeneity of all responses in a study (Goodman, 2002; Collins & Lanza, 2010). In LCA, common responses on observed variables are assumed to be due to an unobserved latent variable representing the previously unknown class of response type (McCutcheon, 1987). By studying the patterns of variation among the observed variables, the latent variable of class membership may be identified, and possibly the characteristics that help explain latent class membership (McCutcheon, 1987). Indeed, LCA has previously been used in the examination of TTO responses. Meghani *et al*. (2009) applied the approach using data collected from men who had prostate cancer or were at risk of prostate cancer and identified three classes of respondents: low traders, high traders and non-traders. Differences between the three classes were found in terms of age, race/ethnicity, history of prostate cancer and the importance of sexual activity. Such findings indicate that applying LCA has the potential to contribute to further understanding of TTO valuations, which is highly relevant for economic evaluations and subsequent decision making.

Heterogeneity in TTO responses is problematic because where major differences between raters exist, consideration needs to be given to which group of preferences is the most important for the population subgroup of interest. This is particularly the case when policy makers must make decisions about providing health care resources to specific subpopulations. Two uses of TTO in health care decision making need to be distinguished here. In using patient preferences, understanding variation in TTO responses may be directly informative for treatment choices. Patients focusing on longevity for some reason may prefer other types of treatments than those focusing on quality of life. Here, we focus on variation in general public preferences for health states, the prominent source for health state valuations and national tariffs (albeit not undisputedly so – Versteegh & Brouwer, 2016). To serve as foundation for healthcare decision making from a societal perspective, health state valuations are commonly required to represent the values provided by the general public respondents, either based on hypothetical health states or their real experiences (Dolan *et al*., 2008). While commonly average general public valuations are used in decision making, understanding variation and subgroups in these valuations nonetheless remains important, for at least two reasons. First, understanding variation in general public preferences may be important for sampling reasons. If obtaining representative, average valuations is the aim, sampling may need to include existing subgroups in a balanced way. This can also increase the comparability of results between studies. Second, basing policy decisions for treatments aimed at specific subgroups of responders (e.g. elderly, women, low income groups, etc.) on average valuations, may misrepresent the relevant (ex ante) welfare impact of the treatments and even to 'wrong' decisions when average valuations are unrepresentative for the relevant subgroup. Large, systematic variations in preferences may thus beg the question whether using average valuations is optimal, or that valuations from a relevant subgroup (from the general public) may be better (see e.g. Versteegh & Brouwer, 2016 for analogous reasoning in the choice between patient and general public preferences). Empirical evidence regarding the extent to which TTO choices differ as a function of respondent characteristics will therefore provide further evidence of the implications of using the values of groups overall, versus the values of specific subgroups as well as indications for optimal sampling. Given the increasing use health state valuations based on TTO in health policy decisions, a better understanding of responses to TTO exercises, also in general public valuations, is important. Therefore, the present research used data derived from the general Netherlands public to investigate whether subgroups of TTO respondents can be identified and characterised in terms of their TTO responses.

8.2 Methods

A questionnaire was administered online by a professional sampling agency to a sample of the general public from the Netherlands, representative in terms of age (range: 18 to 65 years) and gender. An age limitation was imposed on the sample due to the questionnaire being designed to address multiple research questions (e.g., Rappange *et al*., 2015; Van Nooten *et al*., 2015; Wouters *et al*., 2015). A minimum time limit of 15 minutes for survey completion was imposed upon the data, based on a pilot-test of the questionnaire. Thus, respondents who completed the survey in less than 15 minutes were excluded from the final sample used for analysis.

Measures

Time Trade-Off

Respondents were presented with six health states and asked to rank them from best to worst. Following this, respondents were asked to individually rate each health state on a visual analogue scale (VAS), ranging from 0 ('worst imaginable health state') to 100 ('best imaginable health state'). Three of the six health states were imperfect health states specified using the dimensions of the EQ-5D (see Appendix A4). The three imperfect health states ranked based on MVH_A1 scores are were as followed: best: 21211, medium: 22221 and worst 33312. The fourth health state was the respondents' own current health status, also specified using the EQ-5D, as mentioned above. The fifth was "dead" and the sixth "perfect health" (see Appendix A4 for further details).

After ranking and rating the six health states, respondents solved three TTO exercises using a 10 year time horizon. These exercises used the three imperfect EQ-5D health states, which were presented to them in the order in which they had ranked them, from best to worst (see Appendix B4 for the exact TTO question provided to participants). Only the imperfect health states were used for the TTO exercise. Dead can be considered to have a quality of life of 0 and perfect health of 1. The protocol did not include a separate valuation exercise for states ranked or rated lower than dead.

Characteristics

Background information collected from the respondents consisted of age, gender, marital status, education, and children. Furthermore, the EQ-5D and VAS (EuroQol) was used to measure current health status. Finally, respondents were asked the following question: "If, as a result of disease, you had to choose between a shorter life in good health and a longer life in poorer health, what would you choose?"

8.3 Data Analyses

The main outcome was the number of years traded-off (calculated as 10 years minus the answer provided in the TTO exercise).

Latent Class Analysis

To test the hypothesis that there are patterns in the responses across the respondents, representing classes of respondents with varying willingness to trade-off years of life in the three health state valuations, latent class analysis (LCA) was performed in Mplus (version 7.11, Muthén & Muthén, Los Angeles, CA). The LCA models identify an unobserved categorical variable, which is measured by observed response variables. In LCA, which can be considered as a form of cluster analysis, an assumption is made that the associations between items can be explained by the existence of subgroups of respondents, which are not directly observed (i.e. latent). The assignment of individual respondents to a latent class is performed probabilistically. That is, each individual is assigned a probability of latent class membership based on posterior probabilities (Clark & Muthén, 2014; Vermunt & Magidson, 2002). Within each specific latent class, participants are assumed to have a similar response (i.e. the probability of choosing the particular response is the same for all individuals within a latent class); between latent classes, these responses differ. Thus, the objective is to categorize respondents into subgroups based on the observed item responses. In mathematical terms, LCA models with continuous outcomes have the general form

$$f(y_i) = \sum_{k=1}^{K} P(c = k) f(y_i | c = k)$$

where y_i is the vector of responses on the observed variables for individual *i* , and *c* is the categorical latent variable with *K* classes (*k* = 1, 2, …, *K*). A multivariate normal distribution is assumed for $f(y_i|c)$, with means that are specific to classes and the allowance that variances are also class-specific. Thus, for this LCA, the class-specific item parameters are item means and variances.

The statistical identification of clusters was tested by the following statistics to identify the model with the best fit for the data: Bayesian Information Criteria (BIC); Sample-size Adjusted BIC (SABIC); the Lo-Mendell-Rubin adjusted likelihood ratio test (LMR LRT); entropy; size of the smallest latent class; and class assignment probabilities. It is preferred to have smaller values of BIC and SABIC when selecting the number of latent classes (Tofighi & Enders, 2008). The LMR assesses the distribution of the LRT in evaluating the appropriate number of classes, with lower *P* values indicating the preferred model (Meghani *et al.*, 2009). The entropy gives an indication of the accuracy of latent-class assignment for each respondent (Leite & Cooper, 2010; Tofighi

& Enders, 2008). If the entropy is closer to 1, this indicates greater precision of latent-class assignment (Leite & Cooper, 2010; Lubke & Muthén, 2007). The size of the smallest latent class extracted is relevant because very small classes may be "by chance" findings. This could lead to a false indication of the number of latent classes within the heterogeneous data (Stulz *et al*., 2010). While there is no explicit criterion for defining a small class size, the authors believe conclusions should not be based on subgroups comprising less than 5% of the sample. Finally, visual inspection of the latent classes is required, as this provides insight into the reasonableness of the number of latent classes (Stull & Houghton, 2013).[5]

Comparisons of Emergent Subgroups

Once the best-fitting latent-class model was determined, latent class predictor variables were built into the model to identify the respondent characteristics that predict latent class membership. This method is known as the 'Modal ML' (Vermunt, 2010), or the '3-step method' (Asparouhov & Muthén, 2014). The LCA makes up the first step, and during the estimation a nominal variable (*N*) for the most likely latent class is created based on the latent class posterior distribution. That is, *N* is set to be the class *c* for which the probability $P(C = c|U)$ is the largest, where *C* is the latent class variable and *U* is the latent class indicator. Because LCA assigns individuals to latent classes probabilistically, *N* has a classification uncertainty value, which is computed as follows:

$$\rho_{c1,c2} = P(C = c_2|N = c_1) = \frac{1}{N_{c1}} \sum_{N_i = c_1} P(C_i = c_2|U_i)$$

where N_i is the most likely class variable for the *i*-th observation, C_i is the true subgroup variable for the *i*-th observation, and U_i represents the subgroup indicator variables for the *i*-th observation. The probability $P(C_i = c_2|U_i)$ is computed from the estimated LCA model, and is saved by Mplus automatically (when specifying save=CPROB). Given this, we can then compute the probability

$$q_{c2,c1} = P(N = c_1|C = c_2) = \frac{P_{c1,c2} N_{c1}}{\sum_c P_{c,c2} N_c}$$

[1] For transparency and review purposes, the data and the analysis code have been made available to the Editor of the journal. Researchers who are interested in the code can contact Katherine Houghton at khoughton@rti.org.

where N_c is the number of observations classified in latent class c by the most-likely class variable N. Therefore, N can be treated as an imperfect measurement of C, with the measurement error $q_{c2,c1}$. The third step involves the use of the most likely latent class variable as a subgroup indicator variable, with the uncertainty rates set at the probabilities $q_{c2,c1}$. During this step the latent class predictor variables are included, which allows the measurement of the relationships between the subgroup variable and the included respondent characteristics, while directly building in the subgroup classification uncertainty (Asparouhov & Muthén, 2014).

Variables used for this analysis were age, gender, education, marital status, having children, quality of life (based on EQ-5D VAS), subjective life expectancy (SLE) and the stated preference for quality versus quantity of life. Except for the latter variable, these variables were chosen as they have been shown to be associated with responses to TTO exercises in previous research (e.g. Augestad *et al*., 2013; Ayalon & King-Kallimanis, 2009; Best *et a*l., 2010; Hsu et al., 2009; Kontodimopoulos & Niakas, 2006; Shimizu *et al*., 2008; van Nooten *et al*., 2009). To our knowledge, the variable preference for quality versus quantity of life has not been used. This question was added to the survey to check whether a respondent's preference for quality or quantity of life could be meaningfully obtained directly and would reflect their TTO responses. All post-hoc comparisons were conducted in Stata version 13 (StataCorp, College Station, TX).

8.4 Results

Demographic characteristics

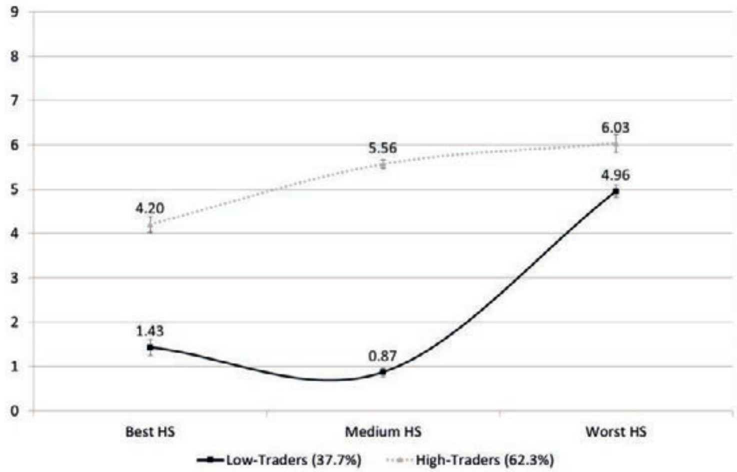
A total of 1,223 respondents completed the questionnaire. A total of 156 respondents (12.8%) were removed due to completing the survey in less than 15 minutes. Thus, the final sample for analysis consisted of 1,067 respondents. As presented in Table 8.1, the mean age of respondents was 43, half were male, 60% reported having children, and the mean VAS score for health was 75.

A two-latent class solution split respondents into high-traders (62% of the sample) and low-traders (38% of the sample). Figure 8.1 presents the mean number of years traded-off for each of the three health states within each subgroup, with 95% confidence intervals. The number of years traded-off differed significantly between health states within each of the two subgroups (difference between low and high traders for the best health state: 2.77 years;

medium health state: 4.69 years; and worst health state: 1.07 years) (Figure 8.1)

Table 8.1 Demographics of respondents

	All respondents (n=1,067)
Age: mean (SD); range	43.2 (13.6); 18-65
Gender: % male	50.2%
EQ-VAS: mean (SD)	75.0 (16.6)
SLE: mean (SD)	37.8 (17.21)
Education level: %	Lower: 15.4% Middle:53.7% Higher: 30.9%
Children: % yes	60.2%
Marital status: %	Married: 49.0% Living together: 15.3% Divorced: 8.5% Widow(er): 2.2% Single: 21.5% NA: 3.5%
Prefer quantity of life over quality of life (%)	43.0%
Health State 1: 21211 (mean (SD))	3.16 (2.58)
Health State 2 : 22221 (mean (SD))	3.80 (2.56)
Health State 3 : 33312 (mean (SD))	5.63 (2.01)



Note: error bars represent 95% confidence interval

Figure 8.1 Number of Years Traded-Off Per Health State: 2-Latent Class Solution

After review of the decision points based on empirical and visual examinations, a four-latent class solution was identified as the model providing the best fit to the data. While the majority of the fit statistics suggested a five-latent class model as the best solution, the size of the smallest latent class (n=40; 3.7% of the sample) was considered too small. Thus, a four-latent class solution was chosen, since it still had an acceptable smallest latent class size (n = 103, 9.7% of total sample), smaller BIC and SABIC values than the 3-latent class solution, a significant Lo-Mendell-Rubin–adjusted likelihood ratio test value, and an entropy value closer to 1. Model fit statistics are presented in Table 8.2.

Table 8.2 Latent Class Analysis of Years Traded-Off: Model Fit Information

Decision Point	2-Latent Class Model	3-Latent Class Model	4-Latent Class Model	5-Latent Class Model
BIC	13,969.454	13,534.163	13,299.967	12,999.359
SABIC	13,937.693	13,489.696	13,242.796	12,929.483
LMR LRT *P*-value	671.744 <0.0001	447.150 <0.0001	253.015 0.012	317.127 <0.0001
Entropy	0.932	0.949	0.967	0.986
Smallest latent class size	n = 402 (37.7%)	n = 86 (8.1%)	n = 103 (9.7%)	n = 40 (3.7%)

BIC: Bayesian information criteria; LMR LRT; Lo-Mendell-Rubin Likelihood Ratio Test; SABI: sample-size adjusted Bayesian information criteria.

As presented in Figure 8.2, the four subgroups of respondents identified by LCA differed in their willingness to trade-off years across the three health states. Interestingly, this four-latent class solution appears to be the result of further parsing the results from the two-latent class solution (presented in Figure 8.1). That is, a group of 'high-traders' (9.7% of the data) and a group of 'medium-high-traders' (50.8%) were identified. Nearly 97% of the 'high-traders' group from the 2-latent class solution were placed into either the 'high-traders' or 'medium-high-traders' groups within the 4-latent class solution.

The four-latent class solution further identified a group of 'low-traders' (27.0%) and 'medium-low-traders' (12.6%). 100% of the 'low-traders' group from the two-latent class solution was placed into either the 'low-traders' or 'medium-low-traders' groups within the 4-latent class solution.

The two low-trader groups differentiated most between the worst and the medium health state, while the two high trader groups most between the best and the medium health state (Figure 8.2). The number of years traded-off differed significantly between health states within each subgroup (Figure 8.2) Interestingly, in the low traders group the number of years traded off in the medium health state was lower (0.33 years) compared to the best health (1.32 years), and in the high traders groups the number of years was higher (8.15 years) in the medium health state compared to the worst health state (7.42 years).

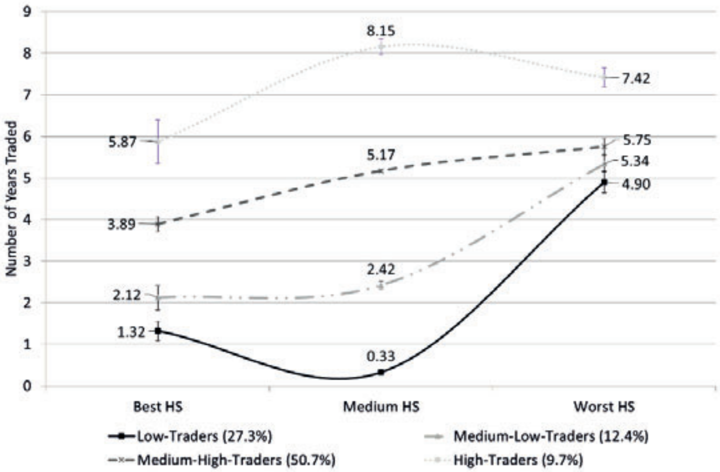


Figure 8.2 Number of Years Traded-Off Per Health State: 4-Latent Class Solution

Comparison of Emergent Subgroups

Descriptive results of comparisons of proportions and means of the predictive variables are presented in Table 8.3. Results of the LCA including predictor variables among the two- and four-latent classes are presented in Table 8.4 and Table 8.5, respectively. The results of the predictive analyses among the two-latent class solution showed that those most willing to trade-off life years ('high-traders') were younger, males, less likely to have children, lower SLE and prefer quality over quantity (Table 8.4). The results of the predictive analyses among the four-latent class solution (Table 8.5) showed that individuals in the low and medium-low trader groups were significantly older, had a higher SLE, and preferred quantity over quality of life when compared to the high trader group. Individuals in the medium-high trader group were significantly older than those in the high trader group. No further significant differences were found between the latent classes.

Table 8.3 Descriptive characteristics among the emergent subgroups.

Characteristics	2-Latent Class Solution		4-Latent Class Solution			
	Low-Traders (n=402)	High-Traders (n=665)	Low-Traders (n=291)	Medium-Low-Traders (n=132)	Medium-High-Traders (n=541)	High-Traders (n=103)
Age: mean (SD)	46.0 (12.8)	41.5 (13.9)	47.3 (12.0)	42.9 (14.1)	41.9 (13.8)	38.9 (13.7)
Gender: % male	45.5%	53.1%	41.9%	56.1%	52.9%	52.4%
EQ-VAS: mean (SD)	74.6 (16.1)	75.3 (16.9)	73.9 (16.4)	76.3 (15.6)	75.8 (17.1)	72.8 (15.6)
SLE: mean (SD)	36.4 (15.8)	38.6 (18.0)	35.2 (15.1)	39.4 (17.4)	38.6 (17.8)	38.5 (18.9)
Education: % highest	30.1%	31.3%	28.2%	37.1%	31.8%	25.2%
Children: % yes	70.7%	53.8%	73.9%	62.1%	54.5%	48.5%
Marital status: % married	57.5%	44.0%	59.5%	51.5%	45.5%	35.0%
Living together:	11.0%	18.0%	10.7%	12.1%	16.6%	25.2%
Prefer quantity of life over quality of life: %	61.7%	33.1%	65.6%	52.3%	32.2%	33.0%

Table 8.4 Comparison of Characteristics among the Emergent Subgroups: Results of a logistic regression analyses among the two-latent class solution

	Regression coefficient	**Standard error**	**P value**
Age	0.029	0.008	**0.001**
Gender	-0.341	0.148	**0.021**
VAS	0.005	0.004	0.220
SLE	0.014	0.006	0.026
Education	0.073	0.160	0.648
Children	0.412	0.184	0.025
Married	0.083	0.182	0.647
Living together	0.354	0.239	0.139
Prefer quantity of life over quality of life	1.127	0.144	**<0.001**

Note. Latent class 2 (high traders) is the reference class.
SLE: subjective life expectancy; VAS: visual analogue scale.

Table 8.5 Comparison of Characteristics Among the Emergent Subgroups: Results of a logistic regression analyses among the four-latent class solution

	Regression coefficient	Standard error	P value
Low traders on:			
Age	0.065	0.014	<0.001
Gender	-0.545	0.254	0.032
EQ-VAS	-0.003	0.007	0.716
SLE	0.030	0.011	0.005
Children	0.419	0.308	0.174
Education	0.332	0.295	0.261
Married	0.308	0.328	0.348
Living together	-0.559	0.354	0.114
Prefer quantity of life over quality of life	1.138	0.259	<0.001
Medium-low traders on:			
Age	0.039	0.015	0.010
Gender	0.050	0.287	0.863
EQ-VAS	0.004	0.008	0.621
SLE	0.029	0.011	0.012
Education	0.617	0.317	0.052
Children	0.268	0.334	0.423
Married	0.187	0.360	0.603
Living together	-0.650	0.397	0.101
Prefer quantity of life over quality of life	0.620	0.290	0.032
Medium high traders on:			
Age	0.032	0.012	0.010
Gender	-0.070	0.233	0.766
EQ-VAS	0.006	0.007	0.365
SLE	0.018	0.010	0.066
Children	-0.056	0.278	0.841
Education	0.336	0.270	0.213
Married	0.227	0.307	0.459

	Regression coefficient	Standard error	P value
Living together	-0.350	0.299	0.243
Prefer quantity of life over quality of life	-0.167	0.244	0.493

Note. High traders is the reference class.
SLE: subjective life expectancy; VAS: visual analogue scale

8.5 Discussion

Previous studies have identified a number of respondent characteristics that are associated with variations in TTO valuations (e.g. Arnesen & Trommald, 2005; Dolan & Roberts, 2002). Most of these studies investigated these relations using regression analyses. The major drawback from all of these studies jointly is the inconsistency of the results (Essink-Bot *et al*., 2007), indicating that an alternative approach to investigating on the variation in TTO responses may be required. Essink-Bot and colleagues suggested that individual response patterns might be more influential determinants of health state valuations than respondent characteristics (Essink-Bot *et al*., 2007). So far only one study, in (potential) prostate cancer patients, investigated whether subgroups of respondents could be identified based on their TTO responses based on hypothetical health states using LCA (Meghani *et al*., 2009). The advantage of using LCA over traditional analytic approaches is that LCA, unlike traditional analytical approaches, does not specify subgroups a priori (e.g., based on demographic characteristics). Latent variable approaches such as those applied in the present analysis explicitly model the heterogeneity within the data, without a priori subgroup specification (Stull & Houghton, 2013), and allow the subgroups (latent classes) to emerge from the heterogeneity. In the present study, LCA found subgroups of differential TTO responses in this sample of the Dutch general public. Predictive analyses found significant differences between these subgroups in their characteristics. The LCA analyses indicated that four subgroups of respondents existed in these TTO data: two high-trader groups, one of which was more willing to trade-off years ('high-traders') and one slightly less willing to trade-off years ('medium-high-traders'); and two low-trader groups, one of which was least willing to trade-off years ('low-traders'), and one which was slightly more willing to trade-off ('medium-low-traders'). Interestingly, these results were quite similar to those reported in the research performed by Meghani and colleagues who identified three groups in their sample: high traders, low traders and non-traders (Meghani *et al*., 2009). In our study, results showed significant differences between subgroups in terms of age, gender, subjective life expectancy, and stated preference for quantity versus quality of life. Compared to the highest traders, lowest traders were older, female, had a higher subjective life expectancy, and preferred quantity over quality of life. Of note, in this 4-latent class solution, the more disparate the trading classes are in terms of willingness to trade, the greater the difference in their characteristics. That is, individuals who are low traders versus high traders have more differences in their characteristics than those who are medium-high traders versus high traders.

Interestingly, this research shows that the question on preference for quantity versus quality of life is associated with responses to the TTO questions and resulting class membership fairly well. These results indicate that respondents were able to classify themselves up front in the low or high trading group, based on their preference for quality or quantity of life (i.e., preference for quantity was associated with low trading, preference for quality with high trading).

In an 'ideal' or perhaps 'simple' world one might expect to find one homogenous group of respondents that trades from low to high based on the severity of the health streams if presented from best to worst, as done in this exercise. Our LCA clearly shows that reality is more complex. In our study, for example, the high-traders traded-off more years on the medium health state compared to the worst health state (8.1 vs 7.3 years, respectively). Something similar was seen in the low-traders group, where respondents traded-off less years for the medium health state compared to the best health state (0.3 vs 1.3 years, respectively). In a study by Lamers *et al*. (2006), it was found that at least one state was valued inconsistently by the majority of respondents (Lamers *et al*., 2006). In spite of ranking and rating exercises, respondents did not always value health states between the 'best' and 'worst' health state lower than the former and higher than the latter. Such violations of transitivity or preference reversals may relate to methods used, but may also signal that respondents do not always trade-off according to easily understandable or linear patterns.

It needs noting that the LCA analyses were performed on the years traded-off per health state, regardless of how the health states were ranked. Most respondents ranked the health states as a priori expected (n=897), while only a minority did not. When re-analysing the LCA results using the number of years traded-off for the health states based on the ranking of the health states (e.g. if the medium health state was ranked as best by a responder, these results were used to specify "best" health state), the results did not change. Thus, the ranking of the health states did not influence the trade-off or the latent class assignment.

Some possible limitations need to be noted. First, we used a Web-based approach for performing the TTO exercises. It has been suggested that a face-to-face approach is better suited for performing TTO studies (Shah *et al*., 2013). However, we considered online data collection more efficient and justified for the current purpose (revealing underlying patterns under the responses), and the obtained responses on average showed plausible patterns. Second, we did not include a separate valuation exercise for health states worse than dead. It was assumed that this might be too complex for the respondents to perform (Van Nooten *et al*., 2009). Third, we used a sample from the Dutch general public. Given cultural differences between countries (for example related to aspects like religion and physician assisted dying), different patterns may be found when similar studies are performed in other countries. Fourth, one of the difficulties in conducting LCA is determining which latent-class solution constitute an optimal representation of the data. Unfortunately, there is no single decision criterion that one can use to assert an optimal fit. Rather, a number of decision criteria must be used. An important step in this process involves viewing the individual-level data within each latent class. If the individual-level data closely tracks the mean of the latent-class to which the model has assigned individuals, this adds to confidence in the findings. Moreover, comparisons of characteristics among the latent-classes can be helpful as model validation: if the identified latent-classes are significantly differentiated on the basis of relevant characteristics, this adds confidence to the chosen solution (Stull & Houghton, 2013). Here, the observed association between years traded-off in each of the classes and respondents' stated preference for quantity versus quality of life adds confidence to our findings.

It should also be noted that this research included the three-level EQ-5D, as the five-level EQ-5D was not yet available at the time of this study. Nowadays, it is gaining momentum in use, hence the results identified in this research should also be verified using the five-level EQ-5D.

8.6 Conclusion

Notwithstanding these limitations, in our view, the findings of this study contribute to a further understanding of TTO valuations beyond the information that traditional regression analyses provide. In particular, this study has shown that there is a great deal of heterogeneity in TTO valuations, but that patterns exist within this heterogeneity. Distinct classes of respondents were discerned in the general public, distinguishing subgroups of low and high traders. Characteristics that can help explain the likelihood of belonging to a subgroup were age, gender, SLE and stated preference for quantity versus quality of life.

In terms of implications of our findings, we first emphasize that our results require confirmation in other studies. Secondly, additional research is required to understand their exact nature, in particular whether the observation of high and low traders in our sample is an expression of a ‘true’ preference. Finally, it would be interesting to see whether alterations to the TTO procedure or instructions could lead to less distinct response patterns between high and low traders and potentially lead to better estimation of health preferences. If our results are confirmed, an important implication is that the use of average values resulting from TTO exercises may be problematic, as they do not reflect the preferences of any sub-group in society. Moreover, it may be that the results of TTO exercises are related to the composition of the sample in terms of low and high traders. If so, this may have consequences for sampling of TTO studies.

Finally, it seems worthwhile to use techniques like LCA to further explore and improve our understanding of the variation in responses to TTO exercises.

9.

Discussion

9.1 Introduction

Economic evaluations can help to inform decisions regarding which health technologies should be publicly financed. Several types of economic evaluations exist wherein the benefits (or effectiveness) of a medical intervention are assessed in relation to its costs, relative to a relevant comparator such as ‘usual care’. The most prominent types are cost-benefit, cost-effectiveness and cost-utility analyses, which differ in the way health benefits are expressed and valued. Cost-utility analysis currently is the preferred type in most countries that use economic evaluations in their decision-making process. In cost-utility analyses, health benefits are expressed in terms of ‘utilities’, calculated as “quality adjusted life years” (QALYs). The QALY measure combines length and quality-of-life in the form of a single numeric measure, in which perfect health is assigned the normalized value 1 and the state ‘dead’ is assigned the value 0 (Culyer, 2014). One year in perfect health then equals one QALY. An important advantage of using a generic measure such as the QALY is that it allows comparisons of the benefits and hence efficiency of different interventions across different diseases, within a health care system.

A challenge in operationalizing the QALY is assigning quality-of-life values or weights (between 0 and 1) to the imperfect health states between perfect health and dead. Several methods have been put forward for this purpose: the Visual Analogue Scale (VAS); Standard Gamble (SG); Time Trade-Off (TTO), and Discrete Choice Experiments (DCE) (e.g., Whitehead & Ali, 2010; Brazier *et al.*, 2012). This thesis focused on the TTO method, because it was frequently used, also for obtaining national valuation sets (‘tariffs’) (Arnesen & Trommald, 2005; Euroqol, 2017; Neumann *et al.*, 1997). In a TTO exercise, a respondent is typically presented with two health profiles; one in which she will live in less than perfect health for X years and the other in which she will live in perfect health but for less than X years. The respondent is instructed to imagine that after the specified number of years she will die. The chosen time horizon X can be varied, but usually is some fixed number, like 10-years, or remaining lifetime (resulting in a ‘lifetime TTO’). Respondents are then asked to state how many years (less than X) in perfect health they consider to be equal to X years in the imperfect health state. The value of the imperfect health state then can be calculated as the ratio between the number of years in perfect health and X.

Research has shown that when utilizing the TTO method, there may be considerable variation in quality of life values for the same health states, both between respondents and between studies (Arnesen & Trommald, 2005). Such variation in quality of life values can influence the outcomes of an economic evaluation and, through that, the decision whether or not to fund a health technology. Hence, it is important to understand this variation. Attema et al. (2013a) distinguished three categories of factors that can influence TTO health state valuations: methodological, procedural and analytical factors. But, TTO health state valuations can also be affected by sample selection and differences therein between studies (Arnesen & Trommald, 2005). Therefore, it is important to understand which respondent characteristics influence TTO responses, especially when the aim is to obtain valuations that are representative for some group (like in the case of national tariffs). More in general, obtaining more understanding of the characteristics of respondents that influence TTO responses is important for several reasons:

1. for comparing health state valuations across studies and populations;
2. for sampling purposes,
3. to enhance knowledge regarding the mechanisms underlying (heterogeneity in) observed health state valuations.

Looking at the literature, it needs to be concluded that we currently lack a thorough and systematic understanding of the influence of respondent characteristics on TTO responses (Chapter 2). The available evidence suggests that several respondent characteristics can influence TTO responses. Numerous studies have reported that respondents' demographic characteristics, such as age, gender, marital status or socio-economic status, can impact TTO responses, but the findings differ across studies. Intrinsic attitudinal aspects (e.g. beliefs about life and death, or expectations about the future), which may also be relevant given the nature of TTO exercises trading off length of life, have been investigated less rigorously (Augestad *et al*., 2013; Essink-Bot *et al*., 2007). Therefore, more research into this topic seems necessary. Meanwhile, it could be useful to already consider these variables when comparing TTO-based quality-of-life weights obtained in different studies, samples, and sources.

Given these caveats in our understanding of TTO responses, this thesis aimed to investigate and provide more insight in the influence of respondent characteristics, both socio-demographic and attitudinal, on TTO responses. In doing so, the following research questions were addressed and discussed within this thesis:

4. Does subjective life-expectancy (SLE) impact the willingness to trade (WTT) and the number of years traded off for health state valuations?
5. How does the awareness of the reduced life-expectation, implied by a 10-year TTO, affect health state valuations?
6. What is the influence of beliefs regarding future health and death, as well as desires to witness certain life events, on respondents' health state valuations?
7. Which responder characteristics influence TTO responses, with an emphasis on consideration of significant others, such as partners and children?
8. Do cultural differences in TTO responses exist?
9. Can classes of respondents be identified based on their response patterns in TTO exercises and which respondent characteristics predict membership of the identified subgroups?

The influence of respondent characteristics

The question related to the influence of SLE, the willingness to trade (WTT) and the number of years traded off (question 1) is central in Chapters 3 and 4. This aspect has also been included in the research performed in Chapters 5, 6 and 8. In all these chapters surveys with respondents from the general public were used to collect the information. To obtain SLE, respondents were asked to provide their expected age of death. Respondent current age was subtracted from this expected age of death to calculate remaining SLE. The WTT variable was created by recoding the number of years traded off. Respondents who were not willing to trade off any life years were compared to those who were willing to trade off at least some lifetime. The results from regression analyses in Chapter 3 showed that SLE was significantly negatively associated with WTT and the number of years traded off, although the influence was small.

When SLE was higher than the life expectancy projected in the TTO exercise, respondents were less willing to give up years (Chapter 3). This might be explained by the fact that respondents in that case implicitly had the feeling that life years were already “taken away” from them given the instructions of the TTO exercise. This may have caused them to be less willing to give up ‘additional’ time, compared to other respondents who felt the projected lifetime was longer than expected. Those might be more willing to trade with the 'additional life years' they were projected to receive in the TTO exercise. Similar results to those presented here were obtained in research performed by Heintz et al. (2013) who performed a similar study among patients with different severities of diabetic retinopathy. Heintz and colleagues (2013) also showed that SLE had a negative association with years traded off in TTO exercises. These findings suggest that respondents, notwithstanding the TTO instructions, appear to consider their own SLE when performing a TTO exercise. Hence, it could be useful to have information about the SLE of respondents when comparing or explaining results of TTO studies, as it sheds more light on the aspects people consider in the context of a TTO exercise.

Research question 2 (How does awareness of reduced life expectation, implied by the 10-year TTO, affect the health state valuations?) was addressed in Chapter 4. In a TTO a respondent is instructed to imagine that at the end of the pre-specified time horizon he/she will die. The way in which this message is conveyed in TTO exercises can differ. The Euroqol group developed a standardized way of performing TTO exercises, in which this aspect of death after the specified time horizon is emphasized several times to the respondents (Oppe *et al*, 2014). The study presented in Chapter 4 illustrated that if this point is not made explicit to respondents, they might not assume they will die at the end of the pre-specified time horizon. We found that only 57% of respondents who were not made aware of their shortened life expectancy explicitly in the study presented in Chapter 4, confirmed that they were actually aware that their life expectancy was shortened in the TTO exercise. Additionally, respondents who explicitly were made aware of their shortened life expectancy, were less willing to trade-off years compared with those who were not. This is somewhat similar to the findings from Chapter 3. Therefore, we found that despite the hypothetical nature of the exercise, explicit awareness of reduced life expectancy can impact TTO responses. Optimal ways of instructing respondents could be further investigated.

Related to SLE and the way the implied (reduced) life expectancy is made explicit in TTO exercises, is the influence of beliefs regarding death (research question 3). Previous research provides some indication that the influence of beliefs and attitudes regarding death and euthanasia could be relevant in TTO exercises (Augestad *et al.*, 2013). This appears plausible since a TTO exercise often involves 'actively' reducing the number of years lived through trading off length of life for quality improvements. Augestad and colleagues (2013) concluded that the way in which death is incorporated in TTO exercises may influence TTO responses. In Chapter 5 respondents were asked about issues like fear of death, support for euthanasia and beliefs regarding life after death. The analysis conducted in Chapter 5 showed that beliefs regarding death, specifically fear of death and support for euthanasia, were statistically significantly associated with TTO responses. Greater fear of death was associated with reduced willingness to trade off years. Support for euthanasia was associated with higher willingness to trade off years. Although these results were statistically significant, the effect sizes of these associations were small. Still, a main finding was that beliefs and attitudes regarding future health and death may be associated with TTO responses.

Closely related to attitudes regarding death are a respondent's attitudes and thoughts about the future. These include attitudes towards and expectations about quality-of-life in the future, future life events (e.g. becoming a grandparent) and expectations regarding aging. Arguably, these could potentially influence TTO responses. For example, when trading off years, respondents could consider future life events that they would like to witness occurring within the timeframe of the TTO exercise. To investigate this, in the study presented in Chapter 5, respondents were asked whether they had thought about specific future events while doing the TTO exercise. The results showed that some respondents did think about major future life events when answering the TTO questions. These included events such as marriage of children, graduation of children or grandchildren as well as the birth of grandchildren. Furthermore, considering such anticipated major future life events was associated with fewer years traded off.

Respondents may also have certain perceptions about their future selves, with regards to how they will age. Such perceptions might also affect TTO responses. Similar to SLE, the existence of such an influence would imply that people do not fully adhere to the instructions of the TTO exercise (which explicitly specifies their future health), but let own expectations influence their TTO responses. In this case, the perceptions respondents have about their future health might make it more or less attractive to reach older ages. If such own perceptions about health are taken into account in a TTO exercise, respondents with poor expectations regarding health at older ages might be more inclined to trade-off future years. This was explored in Chapter 5 (in the context of a 10 year TTO), as well as in the Appendix of Chapter 3 in a lifetime TTO; Appendix A3). Respondents were asked to rate their own future quality-of-life expectations using the EQ-5D 3L dimensions at the ages of 60, 70, 80 and 90. The results in Chapter 5 showed that in a 10-year TTO, quality-of-life expectations were not significantly associated with TTO responses. (In a lifetime TTO exercise, shown in the Appendix of Chapter 3, only the quality-of-life respondents expected to have at the age of 60 was significantly associated with TTO responses.)

To further explore the influence of perceptions of future health, the expectations regarding aging (ERA) instrument was used in Chapter 5. The ERA instrument measures expectations regarding aging in the physical health, mental health and cognitive function domains (Sarkisian *et al.,* 2005). The results from Chapter 5 showed that future expectations about physical and mental health, as measured by the ERA, were associated with TTO responses, in a 10-year TTO exercise. Interestingly, the ERA mental health domain was negatively associated with TTO responses (meaning less years traded off when expectations are better) whereas the ERA physical health domain was positively associated with TTO responses (meaning more years traded off when expectations are better) (Chapter 5), though in both cases with small effect sizes. Additionally, the health risk attitude scale (HRAS) was administered to assess whether health risk attitudes were also related to TTO responses. We observed small, significantly positive results, indicating that those who are more willing to take risks with respect to their health traded off marginally more years.

Overall, it seems that beliefs regarding future health and death, as well as desires to witness certain life events, could affect respondents' health state valuations.

Research question 4 focused on the influence of consideration of significant others, such as partners and children, on TTO responses. Previous research suggests that marital status may be associated with TTO scores. Dolan & Robinson (2002) for instance found that health state valuations were most influenced by age, gender and marital status. Usually, when using marital status as a variable, it is implicitly assumed that only married partnership is of interest. However, increasingly young couples decide not to get married (at first) - at least in Dutch society, but do live together in a sustained relationship. The difference between "being married" and "living together" was explored in Chapter 6. The study in Chapter 6 showed that "being married" and "living together" were associated with TTO responses in opposite ways. Being married led respondents to trade-off less years, whereas living together unmarried led respondents to trade-off more years. Note that, again, the effect sizes of these characteristics were small.

Chapter 6 also investigated the relationship between TTO responses and having children, finding that having children is associated negatively with TTO responses (that is, fewer years traded off). Previous research also found that respondents with children provided higher quality of life values (Devlin *et al*., 2011; Van Der Pol & Shiell, 2007). The study in Chapter 6 suggested that respondents do take significant others (both partners and children) into consideration in TTO exercises. Interestingly, the results were different when the variables were analyzed in isolation or jointly. When all three variables, “having children”, “being married” and “living together,” were included in the regression model, “being married” was no longer significant, whereas “living together” and “having children” remained significant. This could imply that being married and having children are confounding factors. Again, the identified associations were small. It does seem the case, however, that TTO valuations were influenced by significant others such as partners and children. Interestingly, the type of partnership (married or living together) may have different influences. This may relate to underlying ideas of how illness would affect significant others; i.e. becoming a burden for loved ones due to illness, hence trading off more years, or not wishing to be missed, hence trading off less years. Krol et al. (2016) demonstrated these two opposite effects to exist and to indeed affect TTO responses.

Research question 5 focused on exploring the influence of cultural differences on TTO responses. Several studies already identified the association of cultural differences with quality of life values obtained from TTO exercises (Badia *et al*., 2001; Bernert *et al*., 2009; Johnson *et al*., 2005; Knies *et al*., 2009; König *et al.* 2010; Norman *et al*., 2009). The existing research focused on the EQ-5D 3L version. Hence, it would be interesting and important to assess whether cultural differences can also be observed in the context of the EQ-5D 5L version. In Chapter 7, the influence of cultural differences was explored by comparing the TTO valuation sets from the UK and Japan, based on the EQ-5D 3L as well as the EQ-5D 5L. The results of the analyses of EQ-5D 3L showed that Japanese respondents seemed to trade off years more willingly for the mildest health states and less willingly for severe states compared with respondents in the UK. No clear patterns could be established comparing the UK and Japanese TTO responses for the EQ-5D 5L. As the methods used to obtain health state valuations for the EQ-5D 3L as well as EQ-5D 5L in both countries were highly comparable, these results suggest cultural differences play a role in how respondents chose to trade off, at least for the EQ-5D 3L. Moreover, the method in which the health states are shown to respondents (3L vs 5L EQ-5D versions) also appeared to matter, given the differences in results between the 5L and 3L EQ-5D versions. These results highlight the complexity of the influence of cultural differences on TTO health state valuations and encourage future research.

Finally, research question 6 focused on whether classes of respondents can be identified based on their response patterns in TTO exercises, and which respondent characteristics potentially predict membership of the identified subgroups. Essink-Bot et al. (2007) already stated: "*Individual response patterns might be more important determinants of TTO or VAS valuations of health states than age or other respondent characteristics measured.*" The approach used in Chapter 8 of this thesis to investigate individual TTO response patterns and to create classes of these response patterns was latent class analyses (LCA). This latent class analysis used the TTO responses (years traded off) and identified response patterns therein to generate classes (subgroups) of respondents who differed in the number of years they were willing to trade-off. Four classes were identified, which could be merged into two main groups: 'high-traders', who focused on quality-of-life and 'low-traders', who focused on length-of-life. Once these classes were identified, respondent characteristics that predicted membership of the identified classes were investigated. Analyses revealed significant differences between the two main subgroups in terms of age, gender, SLE, as well as in their indicated preference for quantity *versus* quality of life. Respondents in the low trading group were more likely to be older females with a longer SLE and a preference for quantity over quality of life compared to individuals in the high-trader group. The fact that respondents' preferences for either quantity or quality of life matched with being a high or a low trader is interesting, as this suggests that TTO responders are able to indicate up front whether they are high or low traders. In summary, two distinct classes of respondents could be identified based on their patterns of TTO valuations: low and high traders. Membership in these classes can be predicted by age, gender, SLE, as well as in their indicated preference for quantity *versus* quality-of-life.

9.2 Summary

The main research question of this thesis was: Do respondent characteristics, including socio-demographic and attitudinal characteristics, influence TTO responses? In summary, this thesis provided supportive evidence that tsuch respondent characteristics are associated with TTO responses. Respondent characteristics that were found to influence TTO responses in this thesis included SLE, expectations about the future and attitudes towards death, as well as having/considering significant others (partners and/or children). Cultural differences can also play a role in certain cases. It is also became clear that the associations between TTO responses and these variables, when significant, typically is weak (in terms of 'effect size'). Given the potential multitude of small influences, may still make it worthwhile investigating these issues further, also for the purpose of sampling. Moreover, while this thesis has added to the literature by investigating these matters, many questions remain unanswered and more studies, also involving more standardized ways of performing TTO's and using larger samples, remain needed in order to confirm these results and further develop them.

Limitations

Whilst this thesis provided new insights into which respondent characteristics could be associated with TTO responses and confirmed some old ones, it is subject to several noteworthy limitations, which are highlighted in this section.

First, this thesis should be considered as containing exploratory rather than confirmatory analyses. It provided new insights and knowledge regarding the influence of personal characteristics on TTO responses, but further steps should be taken to confirm the results with studies adequately powered to further test old and new hypotheses related to the influence of respondent characteristics on TTO responses. This thesis provides initial evidence in several areas as well as interesting avenues for future research.

One of the limitations of this thesis is that almost all studies were performed in the Netherlands (Chapter 3, 4, 5, 6 and 8), which limits the generalizability and applicability of results to other countries and contexts. After all, as shown in Chapter 7, cultural differences can (sometimes) influence TTO responses. The results presented in this thesis may also be related to cultural factors. For example, the presented influence of expressed support for euthanasia on TTO responses may not only be different in other countries due to other levels of support for euthanasia, but also take different forms in different contexts. Chong *et al*. (2009) for instance found that several socio-cultural factors can predict the support for euthanasia. Hence, if a respondent sample is included in a TTO exercise with a different socio-cultural mix the results could be different compared to the results in this thesis.

Other important limitations are related to the methodology used in this thesis. Commonly, TTO exercises use a choice based method, sometimes called 'the ping-pong method', in which low and high response values for the perfect health state are presented to the respondent who can then chose to accept or not accept the presented trade-off, until the point of indifference is reached (Oppe *et al.*, 2014; Gudex, 1994). The ping-pong method leads to more consistency in respondents' preferences compared with direct matching

(Attema & Brouwer, 2013; Lenert *et al*., 1998). In the present research, only one study used TTO responses obtained via a full ping-pong method, using the EQ-5D procedure for TTO exercises (Chapter 7). Chapters 5, 6 and 8 were based on a TTO exercise in which there was only a one-time ping-pong procedure, which arguably is less adequate than a full ping-pong procedure. The other studies used a direct matching question to determine the point of indifference. The fact that no (full) ping-pong method was used in most studies could have led to higher TTO quality of life values (e.g. less years traded off) compared to other TTO studies (e.g. UK EQ-5D utility weights (Dolan *et al.*, 1997)) and could affect the validity of the results in this thesis.

Another methodological limitation is the medium through which the TTO exercise was performed. Only in Appendix Chapter A3 and Chapter 7 of this thesis face-to-face interviews were used, while the other studies were performed *via* the internet (Chapter 3, 4, 5, 6 and 8). It has been shown that internet-based TTO responses are lower in accuracy than face-to-face interviews (Versteegh *et al.,* 2013; Norman *et al*., 2010), which may affect the validity of the results obtained in this thesis. A distinction should be made here in terms of the accuracy of the TTO results (as reflecting 'true preferences') and the question of which factors influence the results, which was the main question in this thesis.

The results from this thesis may be impacted by all these limitations, which limits their generalizability and warrants future research to confirm our findings. It is important to note that the TTO quality of life values obtained in this thesis do follow the same order as in the UK or Dutch EQ-5D value sets (Dolan *et al.*, 1997; Lamers *et al*., 2006). Furthermore, when comparing the influence of SLE in Chapters 3, and 4, in which a matching method was used to that in Chapters 5, 6 and 8, in which a one-time ping-pong was used, the results consistently confirmed that SLE was significantly associated with TTO responses. The consistency of the finding of the influence of SLE on TTO responses, not only in the here presented studies but also for instance in the study of Heintz et al. (2013), in a patient population, is encouraging.

Another limitation to be considered is that in two of the here presented studies (which formed the basis of chapters 3, 4, 5, 6 and 8), respondents above the age of 65 were excluded, despite the growing number of elderly in many populations. Respondents above the age of 65 might have different health preferences in general, as well as different opinions about or attitudes towards for example SLE, euthanasia and beliefs about the future, compared to younger adults. Not including respondents over 65 therefore could have influenced our findings. Hence, it is important to ensure that elderly respondents are included in future research investigating personal characteristics.

The TTO exercises used in several studies in this thesis did not allow for valuing health states as worse than dead (WTD) (Chapters 3, 4, 5, 6 and 8). Hence, respondents who felt the health state they were asked to rate was worse than dead, if present, were not able to provide an appropriate answer. This could have influenced the results. The influence of personal characteristics on WTD states and their valuation is an interesting issue for future research. There are techniques that can be used to allow respondents to rate a health state that they consider to be worse than dead, like the lead time TTO (LT-TTO) or the "composite" TTO which includes the lead time TTO (Attema *et al.*, 2013b, Devlin *et al.*, 2011, Devlin *et al*., 2013, Oppe *et al*., 2014). In the LT-TTO respondents are provided with a 'lead time' in full health preceding each health state under valuation. Respondents are allowed to trade-off their lead time to avoid health states they consider to be worse than death (Devlin *et al*., 2011). Given that we did not allow for WTD states to be valued, the selection of health states under valuation was done in such a way to only include better than dead states (based on their EQ-5D utility weights). Nonetheless, some respondents may have considered some health states to be worse than dead.

Finally, we focused on TTO in this thesis. Meanwhile, alternative health state valuation techniques, especially Discrete Choice Experiments, are becoming more prominent. We cannot generalize our findings to other health state valuation techniques.

Areas for future research & policy implications

Although this thesis has advanced our knowledge regarding which personal characteristics can influence TTO responses, it was not able to provide a conclusive answer to the question "which respondent characteristics influence TTO responses". Much remains unknown both in terms of which factors influence TTO responses and the extent to which they do, also in relation to different TTO procedures. Further research into which respondent characteristics influence TTO responses, as well as responses to other health state valuation techniques remains needed therefore. This thesis advanced knowledge in this area, also by highlighting new influences, especially SLE. Future research can build on the findings presented here. It is recommended that researchers include attitudinal characteristics next to socio-demographic characteristics when attempting to further explain TTO responses. These attitudinal characteristics could concern subjective life expectancy, beliefs regarding life and death issues, as well as a respondent's preference for quality or quantity of life. However, other areas which were not explored in this thesis should also receive attention, including personality type.

When further research on this topic is performed, this could lead to understanding which respondent characteristics should systematically be included in health state valuation studies. The prioritization could for example focus on those respondent characteristics with the greatest influence on responses to ensure that a manageable list of respondent characteristics is collected so as to minimize respondent burden and study costs. Such a list may also be useful for sampling purposes.

It needs to be noted that, although in multivariate analyses of TTO responses adding more background variables resulted in a clear increase of the explained variance of the models, it still remained low. This means that much of the variance in TTO responses remained unexplained and that unobserved factors might be influential in explaining TTO responses. Respondent characteristics in isolation appear to exert only a small influence on TTO responses, but jointly they better explained the observed TTO responses. Further research testing characteristics jointly rather than in isolation may thus add more fruitfully to our understanding of what influences TTO responses.

Given the fairly modest influence personal characteristics appear to have on TTO responses, one may wonder whether further insight into this matter is relevant at all. In this context, it should first of all be noted that factors as the ones revealed in this thesis also help to understand how people interpret TTO exercises and to what extent they adhere to the instructions provided. This can help to further improve methodology. Moreover, the results may be useful for reasons of sampling. Even though differences in samples and methods may alter results only slightly, it is important to note that the results of cost-effectiveness analyses are often quite sensitive to the quality of life weights used in calculating QALYs (Benedict & Muszbek, 2015; Schackman *et al*., 2004). Small changes in the quality of life weights used in the cost-effectiveness analyses therefore may have a noticeable impact on results.

Interestingly, some countries, such as Sweden, prefer the use of health state valuations by patients in economic evaluations. Versteegh & Brouwer (2016) recently argued that patient valuations could be used in economic evaluations next to general public valuations. If TTO-based quality of life values are used for this, it would be interesting and relevant to see how patient characteristics are associated with TTO responses. Heintz *et al*. (2013)

highlighted that SLE plays a role in patient valuations as well, but other (disease specific) factors remain underexplored. Future research could focus on this, also in the context of other health state valuation techniques.

The results presented in this thesis emphasized that comparing TTO responses for the same health state from different TTO sources should be done with caution. Methodological differences between studies could lead to differences in TTO responses (Attema *et al.*, 2013), but differences in respondent characteristics in the samples may add to this. Therefore, reporting respondent characteristics (especially those considered the most important) is relevant to allow sound comparisons and to help understand differences in valuations. Elements in the methods that might trigger attitudinal responses, such as the way death is introduced in the exercise, deserve further investigation as well.

Most studies that investigated the influence of respondent characteristics in some way or another looked at mean effects. In this thesis we also employed latent class analysis. This suggested that respondents can be classified into high and low traders. Future research should use different techniques, including LCA, to investigate this further in different groups and cultures.

As indicated above, while this thesis focused on the influence of respondent characteristics on TTO responses, other health state valuation techniques are becoming more prominent, especially DCE (Bansback *et al.*, 2012, Robinson *et al.*, 2016, Stolk *et al.*,, 2010). The influence of personal characteristics on responses to DCE questions and other valuation techniques requires further investigation. It would be interesting to understand if similar respondent characteristics influence health state valuations obtained with other techniques. One might hypothesize that some influences might be dissimilar, for instance because of the fact that in a TTO exercise people need to explicitly trade-off length of life, which could increase the influence of aspects like SLE and fear of death. This may be different in other valuation techniques. More research in this area is encouraged.

9.3 Conclusion

In summary, this thesis provided further evidence that a range of socio-demographic and attitudinal respondent characteristics can be associated with TTO responses. While some of the influential variables were already known from previous research, this thesis also highlighted new variables which can influence TTO responses. In particular the influence of subjective life expectancy, found in several studies (also outside this thesis), needs to be mentioned in that context. The several studies included in this thesis highlighted that diverse aspects were associated with TTO scores, including, next to SLE, expectations about the future and attitudes towards death, having a preference for length or quality of life as well as having/considering significant others such as partners or children. While the associations between TTO responses and these variables were typically small in terms of ‘effect size’, knowledge on these factors may still be helpful, for instance in explaining TTO responses, understanding how people respond to TTO exercises, optimizing TTO methods as well as sampling of respondents for health state valuation exercises.

Hence, the investigation of the influence of person characteristics on health state valuations is certainly not finished. However, this thesis has shed more light on the relevance and influence on health state valuations of characteristics of the person behind the TTO.

References

REFERENCES

1. Alonso J, Black C, Norregaard JC, et al. Cross-cultural differences in the reporting of global functional capacity: an example in cataract patients. Med Care. 1998;36(6):868-878.

2. Arnesen T, Trommald M. Roughly right or precisely wrong? Systematic review of quality-of-life weights elicited with the time trade-off method. J Health Serv Res Policy. 2004;9(1):43-50.

3. Arnesen T, Trommald M. Are QALYs based on time trade-off comparable?--A systematic review of TTO methodologies. Health Econ. 2005;14(1):39-53.

4. Asparouhov, T. & Muthén, B. (2014) Auxiliary variables in mixture modeling: Three-step approaches using Mplus. Structural Equation Modeling: A Multidisciplinary Journal, 21:3, 329-341.

5. Attema AE, Brouwer WB. Can we fix it? Yes we can! But what? A new test of procedural invariance in TTO-measurement. Health Econ. 2008 Jul;17(7):877-85.

6. Attema AE, Brouwer WBF. The correction of TTO-scores for utility curvature using a risk-free utility elicitation method. J Health Econ. 2009;28(1):234-243.

7. Attema AE, Brouwer WBF. The value of correcting values: influence and importance of correcting TTO scores for time preference. Value Health. 2010;13(8):879-884.

8. Attema AE, Brouwer WBF. In search of a preferred preference elicitation method: A test of the internal consistency of choice and matching tasks. J Econ Psychol. 2013;39:126-140.

9. Attema AE, Edelaar-Peeters Y, Versteegh MM, Stolk EA. Time trade-off: One methodology, different methods. Eur J Heal Econ. 2013;14(SUPPL. 1).

10. Attema AE, Versteegh MM, Oppe M, Brouwer WB, Stolk EA. Lead time TTO: leading to better health state valuations? Health Econ. 2013;22(4):376–92.

11. Attema AE, Versteegh MM. Would you rather be ill now, or later? Health Econ. 2013;22(12):1496-1506.

12. Au N, Lorgelly PK. Anchoring vignettes for health comparisons: an analysis of response consistency. Qual Life Res. 2014;23(6):1721-1731.

13. Augestad LA, Rand-Hendriksen K, Stavem K, Kristiansen IS. Time trade-off and attitudes toward euthanasia: implications of using “death” as an anchor in health state valuation. Qual Life Res. 2013;22(4):705-714.

14. Augustovski F, Rey-Ares L, Irazola V, Oppe M, Devlin NJ. Lead versus lag-time trade-off variants: does it make any difference? Eur J Health Econ. 2013;14 Suppl 1:S25-31.

15. Ayalon L, King-Kallimanis BL. Trading years for perfect health: results from the health and retirement study. J Aging Health. 2010;22(8):1184-1197.

16. Badia X, Monserrat S, Roset M, Herdman M. Feasibility, validity and test-retest reliability of scaling methods for health states: the visual analogue scale and the time trade-off. Qual Life Res. 1998;8(4):303-310.

17. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. J Health Econ. 2012;31(1):306-318.

18. Benedict A, Muszbek M. Life after clinical trials - difficulties in assessing value of drugs in advanced oncology. European Journal of Cancer. 2015: 15(suppl 3): S185

19. Bernert S, Fernández A, Haro JM, et al. Comparison of different valuation methods for population health status measured by the EQ-5D in three European Countries. Value Heal. 2009;12(5):750-758.

20. Best JH, Garrison LP, Hollingworth W, Ramsey SD, Veenstra DL. Preference values associated with stage III colon cancer and adjuvant chemotherapy. Qual Life Res. 2010;19(3):391-400.

21. Bhatnagar V, Frosch DL, Tally SR, Hamori CJ, Lenert L, Kaplan RM. Evaluation of an internet-based disease trajectory decision tool for prostate cancer screening. Value Health. 2009;12(1):101-108.

22. Blankers M, Nabitz U, Smit F, Koeter MWJ, Schippers GM. Economic evaluation of internet-based interventions for harmful alcohol use alongside a pragmatic randomized controlled trial. J Med Internet Res. 2012;14(5):e134.

23. Bleichrodt H, Gafni A. Time preference, the discounted utility model and health. J Health Econ. 1996;15(1):49-66.

24. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ. 2002;11(5):447-456.

25. Bleichrodt H, Pinto JL, Abellan-Perpiñan JM. A consistency test of the time trade-off. J Health Econ. 2003;22(6):1037-1052.

26. Brazier J, Rowen D, Yang Y, Tsuchiya A. Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. Eur J Heal Econ. 2012;13(5):575-587.

27. Brazier J, Ratcliffe J, Tsuchiya, A SJ. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford: Oxford University Press; 2007.

28. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006;10(4):287-333.

29. Brouwer WBF, Niessen LW, Postma MJ, Rutten FFH. Need for differential discounting of costs and health effects in cost effectiveness analyses. BMJ. 2005;331(7514):446-448.

30. Brouwer WBF, van Exel NJA. Expectations regarding length and health related quality of life: some empirical findings. Soc Sci Med. 2005;61(5):1083-1094.

31. Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. J Health Econ. 2008;27(2):325-338.

32. Brown MM, Brown GC, Sharma S, Landy J, Bakal J. Quality of life with visual acuity loss from diabetic retinopathy and age-related macular degeneration. Arch Ophthalmol. 2002;120(4):481-484.

33. Buckingham JK, Birdsall J, Douglas JG. Comparing three versions of the time tradeoff: time for a change? Med Decis Making. 1996;16(4):335-347.

34. CADTH. Guidelines for the Economic Evaluation of Health Technologies, 3rd Edition, 2006.; 2006. Ottowa: Canadian Agency for Drugs and Technologies in Health. https://www.cadth.ca/about-cadth/how-we-do-it/methods-and-guidelines/guidelines-for-the-economic-evaluation-of-health-technologies-canada.

35. Carroll C, Hummel S, Leaviss J, et al. Systematic review, network meta-analysis and exploratory cost-effectiveness model of randomized trials of minimally invasive techniques versus surgery for varicose veins. Br J Surg. 2014;101(9):1040-1052.

36. CBS. We koersen af op een langer leven in goede gezondheid. https://www.cbs.nl/nl-nl/nieuws/2014/26/we-koersen-af-op-een-langer-leven-in-goede-gezondheid. Diemen: Centraal bureau voor de statistiek. Published 2014. Accessed December 18, 2016.

37. CBS. Zorguitgaven; kerncijfers. Zorguitgaven; kerncijfers; 19 mei 2016. Heerlen: Centraal bureau voor de statistiek. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=83037ned. Published 2016. Accessed December 16, 2016b.

38. CBS. Levensverwachting; geslacht, leeftijd (per jaar en periode van vijf jaren). Heerlen: Centraal bureau voor de statistiek. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37360ned&D1=3&D2=a&D3=0&D4=a&HDR=T,G3&STB=G2,G1&VW=T. Published 2016. Accessed December 16, 2016a.

39. Chapman GB, Elstein AS, Kuzel TM, et al. Prostate cancer patients' utilities for health states: how it looks depends on where you stand. Med Decis Making. 1998;18(3):278-286.

40. Clark SL, Muthén B. Relating Latent Class Analysis Results to Variables not Included in the Analysis. http://www.statmodel.com/download/relatinglca.pdf. Accessed October 5, 2014.

41. Clarke AE, Goldstein MK, Michelson D, Garber AM, Lenert LA. The effect of assessment method and respondent population on utilities elicited for Gaucher disease. Qual Life Res. 1997;6(2):169-184.

42. Clarke CE, Patel S, Ives N, et al. Physiotherapy and Occupational Therapy vs No Therapy in Mild to Moderate Parkinson Disease: A Randomized Clinical Trial. JAMA Neurol. 2016;73(3):291-299. 52.

43. Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. Health Econ. 2011;20(1):2-15.

44. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore. 1994;23(2):129-138.

45. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. Clin Cancer Res. 2011;17(21):6646-6650.

46. Collins LM LS. Latent Class and Latent Transition Analysis with Applications in the Social Behavioral, and Health Sciences. (Wiley, ed.). Hoboken; 2010.

47. Culyer AJ. Encyclopedia of Health Economics. First. Netherlands: Elsevier B.V.; 2014.

48. De Wit GA, Busschbach JJ, De Charro FT. Sensitivity and perspective in the valuation of health status: whose values count? Health Econ. 2000;9(2):109-126.

49. Devlin NJ, Tsuchiya A, Buckingham K, Tilling C. A uniform time trade off method for states better and worse than dead: feasibility study of the "lead time" approach. Health Econ. 2011;20(3):348-361.

50. Devlin NJ, Buckingham K, Shah K, et al. A comparison of alternative variants of the lead and lag time TTO. Health Econ. 2013;22(5):517-532.

51. Devlin NJ, Shah K, Feng Y, Mulhern B, van Hout B. Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England. Oxford; 2016.

52. Doctor JN, Bleichrodt H, Lin HJ. Health utility bias: a systematic review and meta-analytic evaluation. Med Decis Making. 2010;30(1):58-67.

53. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. J Health Econ. 1996;15(2):209-231.

54. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ. 1996;5(2):141-154.

55. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-1108.

56. Dolan P. Effect of age on health state valuations. J Health Serv Res Policy. 2000;5(1):17-21.

57. Dolan P, Roberts J. To what extent can we explain time trade-off values from other information about respondents? Soc Sci Med. 2002;54(6):919-929.

58. Dolan P. Developing methods that really do value the “Q” in the QALY. Health Econ Policy Law. 2008;3(Pt 1):69-77.

59. Dolan P. Thinking about it: thoughts about health and valuing QALYs. Health Econ. 2011;20(12):1407-1416.

60. Douven R, Ligthart M, Mannaerts HIW. Een Scenario Voor de Zorguitgaven 2008-2011; Diemen: Centraal Plan Bureau. 2006.

61. Drummond MF, O’Brien B, Stoddart GL TG. Methods for the Economic Evaluation of Health Care Programmes. 2. Oxford: Oxford Medical Publications; 1997.

62. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL and TG. Methods for the Economic Evaluation of Health Care Programmes. Fourth. Oxford: Oxford University Press; 2015.

63. Essink-Bot M-L, Stuifbergen MC, Meerding W-J, Looman CWN, Bonsel GJ, VOTE group. Individual differences in the use of the response scale determine valuations of hypothetical health states: an empirical study. BMC Health Serv Res. 2007;7:62.

64. Euroqol. Euroqol. ttp://www.euroqol.org. Accessed October 20, 2014.

65. Euroqol. How to use EQ-5D. http://www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html. Accessed December 18, 2016.

66. Flynn TN. Using conjoint analysis and choice experiments to estimate QALY values: Issues to consider. Pharmacoeconomics. 2010;28(9):711-722.

67. Fobelets M, Pil L. bevolkingsonderzoek naar borstkanker in Vlaanderen : gezondheidseconomische evaluatie. 2015:1-26.

68. Fobelets M, Pil L, Annemans L. bevolkingsonderzoek naar dikkedarmkanker in Vlaanderen : gezondheidseconomische evaluatie. 2015:0-32.

69. Franken M, Heintz E, Gerber-Grote A, Raftery J. Health Economics as Rhetoric: The Limited Impact of Health Economics on Funding Decisions in Four European Countries. Value Heal. 2016:1-6.

70. Froberg DG, Kane RL. Methodology for measuring health-state preferences--II: Scaling methods. J Clin Epidemiol. 1989;42(5):459-471.

71. Gafni A, Torrance GW. Risk attitude and time preference in health. Manage Sci 1984; 30: 440–451.

72. Goodman LA. Latent class analysis: The empirical study of latent types, latent variables, and latent structures. In: Hagenaars JA, McCutcheon AL E, ed. Applied Latent Class Analysis. New York: Cambridge University Press; 2002:3–55.

73. Goren A, Mould-Quevedo J, daCosta DiBonaventura M. Prevalence of pain reporting and associated health outcomes across emerging markets and developed countries. Pain Med. 2014;15(11):1880-1891.

74. Green C, Brazier J, Deverill M. Valuing health-related quality of life. A review of health state valuation techniques. Pharmacoeconomics. 2000;17(2):151-165.

75. Greiner W, Claes C, Busschbach JJ V, von der Schulenburg J-MG. Validating the EQ-5D with time trade off for the German population. Eur J Health Econ. 2005;6(2):124-130.

76. Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ. 2011;20(2):212-224.

77. Groenewoud JH, Otten JDM, Fracheboud J, et al. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. Breast Cancer Res Treat. 2007;102(2):211-218.

78. Guest JF, Sladkevicius E, Gough N, Linch M, Grimer R. Utility values for advanced soft tissue sarcoma health States from the general public in the United kingdom. Sarcoma. 2013;2013:863056.

79. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-1736.

80. Gold MR, Siegel JE, Russell LB and MCW. Cost-Effectiveness in Health and Medicine. First. Oxford: Oxford University Press; 1996.

81. Gupta V, Srinivasan G, Mei SS, Gazzard G, Sihota R, Kapoor KS. Utility values among glaucoma patients: an impact on the quality of life. Br J Ophthalmol. 2005;89(10):1241-1244.

82. Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being: a World Health Organization Study in Primary Care. JAMA. 1998;280(2):147-151.

83. Hakim Z, Wolf A, Garrison LP. Estimating the effect of changes in body mass index on health state preferences. Pharmacoeconomics. 2002;20(6):393-404.

84. Handler RM, Hynes LM, Nease RF. Effect of locus of control and consideration of future consequences on time tradeoff utilities for current health. Qual Life Res. 1997;6(1):54-60.

85. Harzing A-W. Response Styles in Cross-national Survey Research: A 26-country Study. Int J Cross Cult Manag. 2006;6(2):243-266.

86. Heintz E, Gerber-Grote A, Ghabri S, et al. Is There a European View on Health Economic Evaluations? Results from a Synopsis of Methodological Guidelines Used in the EUnetHTA Partner Countries. Pharmacoeconomics. 2016;34(1):59-76.

87. Heintz E, Krol M, Levin L-Å. The impact of patients' subjective life expectancy on time tradeoff valuations. Med Decis Making. 2013;33(2):261-270.

88. Hobara M. Beliefs about appropriate pain behavior: cross-cultural and sex differences between Japanese and Euro-Americans. Eur J Pain. 2005;9(4):389-393.

89. Hoefman RJ, van Exel J, Brouwer W. How to include informal care in economic evaluations. Pharmacoeconomics. 2013;31(12):1105-1119.

90. Hsu PC, Krajden M, Yoshida EM, Anderson FH, Tomlinson GA, Krahn MD. Does cirrhosis affect quality of life in hepatitis C virus-infected patients? Liver Int. 2009;29(3):449-458.

91. Hsu PC, Federico CA, Krajden M, et al. Health utilities and psychometric quality of life in patients with early- and late-stage hepatitis C virus infection. J Gastroenterol Hepatol. 2012;27(1):149-157.

92. Ikeda, S., Shiroiwa, T., Igarashi, A., Noto, S., Fukuda, T., Saito, S., & Shimozuma K. Developing a Japanese version of the EQ-5D-5L value set. J Natl Inst Public Heal. 2015;64(1):47-55.

93. Izadpanah A, Sinno H, Vorstenbosch J, Lee BT, Lin SJ. Thigh laxity after massive weight loss: a utilities outcomes assessment. Ann Plast Surg. 2013;71(3):304-307.

94. Jakubiak-Lasocka J, Jakubczyk M. Cost-effectiveness versus Cost-Utility Analyses: What Are the Motives Behind Using Each and How Do Their Results Differ?-A Polish Example. Value Heal Reg Issues. 2014;4:66-74.

95. Johannesson M, Pliskin JS, Weinstein MC. A note on QALYs, time tradeoff, and discounting. Med Decis Making. 1994;14(2):188-193.

96. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? Med Care. 2005;43(3):221-228.

97. Kahneman, D. & Tversky A. Prospect theory: An analysis of decision under risk. Econometrica,. 1979;47:263-291.

98. Kattan MW, Fearn PA, Miles BJ. Time trade-off utility modified to accommodate degenerative and life-threatening conditions. Proceedings AMIA Symp. 2001:304-308.

99. Kind, P, Brooks, R, Rabin R. EQ-5D Concepts and Methods: A Developmental History. Dordrecht: Springer; 2005.

100. Kirsch J, McGuire A. Establishing health state valuations for disease specific states: an example from heart disease. Health Econ. 2000;9(2):149-158.

101. Knies S, Evers SMAA, Candel MJJM, Severens JL, Ament AJHA. Utilities of the EQ-5D: transferable or not? Pharmacoeconomics. 2009;27(9):767-779.

102. Komiyama O, Wang K, Svensson P, Arendt-Nielsen L, Kawara M, De Laat A. Ethnic differences regarding sensory, pain, and reflex responses in the trigeminal region. Clin Neurophysiol. 2009;120(2):384-389.

103. König H-H, Heider D, Lehnert T, et al. Health status of the advanced elderly in six European countries: results from a representative survey using EQ-5D and SF-12. Health Qual Life Outcomes. 2010;8:143.

104. Kontodimopoulos N, Niakas D. Overcoming inherent problems of preference-based techniques for measuring health benefits: an empirical study in the context of kidney transplantation. BMC Health Serv Res. 2006;6:3.

105. Krol M, Brouwer W, Sendi P. Productivity costs in health-state valuations : does explicit instruction matter? Pharmacoeconomics. 2006;24(4):401-414.

106. Krol M, Sendi P, Brouwer W. Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. Value Health. 2009;12(1):172-180.

107. Krol M, Brouwer W, Rutten F. Productivity costs in economic evaluations: past, present, future. Pharmacoeconomics. 2013;31(7):537-549.

108. Krol M, Brouwer W.How to estimate productivity costs in economic evaluations. Pharmacoeconomics. 2014 Apr;32(4):335-44.

109. Krol M, Attema AE, van Exel J, Brouwer W. Altruistic Preferences in Time Tradeoff: Consideration of Effects on Others in Health State Valuations. Med Decis Making. 2016;36(2):187-198.

110. Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJ V. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ. 2006;15(10):1121-1132.

111. Lamers LM, Stalmeier PFM, Krabbe PFM, Busschbach JJ V. Inconsistencies in TTO and VAS values for EQ-5D health states. Med Decis Making. 2006;26(2):173-181.

112. Leite WL, Cooper LA. Detecting Social Desirability Bias Using Factor Mixture Models. Multivariate Behav Res. 2010;45(2):271-293.

113. Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber A. The effect of search procedures on utility elicitations. Med Decis Making. 1998;18(1):76-83.

114. Lubeck DP, Grossfeld GD, Carroll PR. A review of measurement of patient preferences for treatment outcomes after prostate cancer. Urology. 2002;60(3 Suppl 1):72-7-8.

115. Lubke GH, Muthén B. Investigating population heterogeneity with factor mixture models. Psychol Methods. 2005;10(1):21-39.

116. Luo N, Wang Y, How CH, Tay EG, Thumboo J, Herdman, M. Interpretation and use of the 5-level EQ-5D response labels varied with survey language among Asians in Singapore. Journal of Clinical Epidemiology. 2015a:68(10), 1195–1204

117. Luo N, Wang Y, How CH, et al. Cross-cultural measurement equivalence of the EQ-5D-5L items for English-speaking Asians in Singapore. Qual Life Res. 2015b;24(6):1565-1574.

118. Martin AJ, Glasziou PP, Simes RJ, Lumley T. A comparison of standard gamble, time trade-off, and adjusted time trade-off scores. Int J Technol Assess Health Care. 2000;16(1):137-147.

119. Matza LS, Boye KS, Feeny DH, Johnston JA, Bowman L, Jordan JB. Impact of caregiver and parenting status on time trade-off and standard gamble utility scores for health state descriptions. Health Qual Life Outcomes. 2014;12:48.

120. Mazur DJ, Merz JF. Older patients' willingness to trade off urologic adverse outcomes for a better chance at five-year survival in the clinical setting of prostate cancer. J Am Geriatr Soc. 1995;43(9):979-984.

121. McCutcheon AL. Latent Class Analysis. Quantitative Applications in the Social Sciences Series. Volume 64. (Publications S, ed.). Thousand Oaks; 1987.

122. Meghani SH, Lee CS, Hanlon AL, Bruner DW. Latent class cluster analysis to understand heterogeneity in prostate cancer treatment utilities. BMC Med Inform Decis Mak. 2009;9:47.

123. Meltzer D, Weckerle C, Chang LM. Do people consider financial effects in answering quality of life questions. Med Decis Mak. 1999;19:517.

124. Mirowsky J. Subjective life expectancy in the US: correspondence to actuarial estimates by age, sex and race. Soc Sci Med. 1999;49(7):967-979.

125. Morimoto T, Fukui T. Utilities measured by rating scale, time trade-off, and standard gamble: review and reference for health care professionals. J Epidemiol. 2002;12(2):160-178.

126. Mrus JM, Leonard AC, Yi MS, et al. Health-related quality of life in veterans and nonveterans with HIV/AIDS. J Gen Intern Med. 2006;21 Suppl 5:S39-47.

127. Mrus JM, Sherman KE, Leonard AC, Sherman SN, Mandell KL, Tsevat J. Health values of patients coinfected with HIV/hepatitis C: are two viruses worse than one? Med Care. 2006;44(2):158-166.

128. Muthén LK MB. Mplus User's Guide. 6th ed. (Muthén & Muthén, ed.). Los Angeles

129. NICE. Guide to the Single Technology Appraisal Process.; 2009. London: National Institute of Health and Clinical Excellence. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/Guide-to-the-single-technology-appraisal-process.pdf. Accessed December 16, 2016

130. Neumann PJ, Zinner DE, Wright JC. Are methods for estimating QALYs in cost-effectiveness analyses improving? Med Decis Mak. 1997;17(4):402-408.

131. Norman R, Cronin P, Viney R. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. Appl Health Econ Health Policy. 2013;11(3):287-298.

132. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro, F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value in Health.2014:17(4), 445–453

133. Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ. 2006;15(7):653-664.

134. Péntek M, Brodszky V, Gulácsi ÁL, et al. Subjective expectations regarding length and health-related quality of life in Hungary: results from an empirical investigation. Health Expect. 2014;17(5):696-709.

135. PBAC. Guidelines for Preparing Submissions to the Pharmaceutical Benefits. Advisory Committee. Version 5.0.; 2016. Canberra: Pharmaceutical Benefits Advisory Committee. https://pbac.pbs.gov.au/ Accessed December 16, 2016.

136. Prieto L, Sacristán JA. Problems and solutions in calculating quality-adjusted life years (QALYs). Health Qual Life Outcomes. 2003;1:80.

137. Prosser LA, Payne K, Rusinak D, Shi P, Uyeki T, Messonnier M. Valuing health across the lifespan: Health state preferences for seasonal influenza illnesses in patients of different ages. Value Heal. 2011;14(1):135-143.

138. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. Med Care. December 2014.

139. Rappange DR, Brouwer WBF, van Exel J. A long life in good health: subjective expectations regarding length and future health-related quality of life. Eur J Health Econ. 2016;17(5):577-589.

140. Rabin R, Gudex C, Selai C, & Herdman, M. From translation to version management: A history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. Value in Health. 2014:17(1), 70–76.

141. Richardson J, Peacock SJ, Iezzi A. Do quality-adjusted life years take account of lost income? Evidence from an Australian survey. Eur J Health Econ. 2009;10(1):103-109.

142. Robinson A, Dolan P, Williams A. Valuing health status using VAS and TTO: what lies behind the numbers? Soc Sci Med. 1997;45(8):1289-1297.

143. Robinson A, Spencer A. Exploring challenges to TTO utilities: valuing states worse than dead. Health Econ. 2006;15(4):393-402.

144. Robinson A, Spencer AE, Pinto-Prades JL, Covey JA. Exploring Differences between TTO and DCE in the Valuation of Health States. Med Decis Making. 2017;37(3):273-284.

145. Ross CE, Mirowsky J. Family relationships, social support and subjective life expectancy. J Health Soc Behav. 2002;43(4):469-489.

146. Rowen D, Brazier J, Van Hout B. A Comparison of Methods for Converting DCE Values onto the Full Health-Dead QALY Scale. Med Decis Making. 2015;35(3):328-340.

147. Rowen D, Brazier J, Young T, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health. 2011;14(5):721-731.

148. Rutten-van Mölken MPMH, Hoogendoorn M, Lamers LM. Holistic preferences for 1-year health profiles describing fluctuations in health: the case of chronic obstructive pulmonary disease. Pharmacoeconomics. 2009;27(6):465-477.

149. Salomon JA, Patel A, Neal B, et al. Comparability of patient-reported health status: multicountry analysis of EQ-5D responses in patients with type 2 diabetes. Med Care. 2011;49(10):962-970.

150. Sarkisian CA, Steers WN, Hays RD, Mangione CM. Development of the 12-item Expectations Regarding Aging Survey. Gerontologist. 2005;45(2):240-248.

151. Scalone L, Cortesi PA, Ciampichini R, et al. Italian population-based values of EQ-5D health states. Value Health. 2013;16(5):814-822.

152. Schackman BR, Gold HT, Stone PW, Neumann PJ. How often do sensitivity analyses for economic parameters change cost-utility analysis conclusions? Pharmacoeconomics. 2004;22(5):293-300.

153. Scott NW, Fayers PM, Bottomley A, et al. Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses. Qual Life Res. 2006;15(6):1103-15-20.

154. Shah KK, Lloyd A, Oppe M, Devlin NJ. One-to-one versus group setting for conducting computer-assisted TTO studies: findings from pilot

studies in England and the Netherlands. Eur J Health Econ. 2013;14 Suppl 1:S65-73.

155. Sharma S, Brown GC, Brown MM, Hollands H, Robins R, Shah GK. Validity of the time trade-off and standard gamble methods of utility assessment in retinal patients. Br J Ophthalmol. 2002;86(5):493-496.

156. Shimizu F, Fujino K, Ito YM, et al. Factors associated with variation in utility scores among patients with prostate cancer. Value Health. 2008;11(7):1190-1193.

157. Shirado O, Doi T, Akai M, Fujino K, Hoshino Y, Iwaya T. An outcome measure for Japanese people with chronic low back pain: an introduction and validation study of Japan Low Back Pain Evaluation Questionnaire. Spine (Phila Pa 1976). 2007;32(26):3052-3059.

158. Souchek J, Byrne MM, Kelly PA, et al. Valuation of arthritis health states across ethnic groups and between patients and community members. Med Care. 2005;43(9):921-928.

159. Spencer A. The TTO method and procedural invariance. Health Econ. 2003;12(8):655-668.

160. Stening, B. W., & Everett JE. Response Styles in a Cross-Cultural Managerial Study. J Soc Psychol. 1984;122(2):151-156.

161. Stewart AL, Nápoles-Springer A. Health-related quality-of-life assessments in diverse population groups in the United States. Med Care. 2000;38(9 Suppl):II102-24.

162. Stiggelbout AM, Kiebert GM, Kievit J, Leer JW, Stoter G, de Haes JC. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. Med Decis Making. 1994;14(1):82-90.

163. Stiggelbout AM, Kiebert GM, Kievit J, Leer JW, Habbema JD, De Haes JC. The “utility” of the Time Trade-Off method in cancer patients: feasibility and proportional Trade-Off. J Clin Epidemiol. 1995;48(10):1207-1214.

164. Stolk E a, Oppe M, Scalone L, Krabbe PF. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. Value Heal. 2010;13(8):1005-1013.

165. Stolk EA, van Nooten FE. Values for resource allocation should expose the adaptation process, not the outcome. Virtual Mentor. 2005 Feb 1;7(2).

166. Stull DE, Houghton K. Identifying differential responders and their characteristics in clinical trials: innovative methods for analyzing longitudinal data. Value Health. 2013;16(1):164-176.

167. Stulz N, Thase ME, Klein DN, Manber R, Crits-Christoph P. Differential effects of treatments for chronic depression: a latent growth model reanalysis. J Consult Clin Psychol. 2010;78(3):409-419.

168. Sun S, Chen J, Kind P, Xu L, Zhang Y, Burström K. Experience-based VAS values for EQ-5D-3L health states in a national general population health survey in China. Qual Life Res. 2015;24(3):693-703.

169. Sutcliffe P, Connock M, Pulikottil-Jacob R, et al. Clinical effectiveness and cost-effectiveness of second- and third-generation left ventricular assist devices as either bridge to transplant or alternative to transplant for adults eligible for heart transplantation: systematic review and cost-effectivene. Health Technol Assess. 2013;17(53).

170. Szabo SM, Levy AR, Davis C, Holyoake TL, Cortes J. A multinational study of health state preference values associated with chronic myelogenous leukemia. Value Health. 2010;13(1):103-111.

171. Szende A, Devlin N, Oppe M, SpringerLink. EQ-5D Value Sets Inventory, Comparative Review and User Guide, EuroQol Group Monographs 2. Dordrecht: Springer; 2007.

172. Tamayama K, Kondo M, Shono A, Okubo I. Utility weights for allergic rhinitis based on a community survey with a time trade-off technique in Japan. Allergol Int. 2009;58(2):201-207.

173. Tilling C, Krol M, Tsuchiya A, Brazier J, Exel J van, Brouwer W. Does the EQ-5D reflect lost earnings? Pharmacoeconomics. 2012;30(1):47-61.

174. Tofighi D EC. Identifying the correct number of classes in growth mixture models. In: Hancock GR SK, ed. Advances in Latent Variable Mixture Models. Charlotte: Information Age Publishing, Inc.; 2008.

175. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res. 1972;7(2):118-133.

176. Torrance GW. Social preferences for health states. An empirical evaluation of three measurement techniques. Socio-Econ Plan. 1976;10:129-136.

177. Torrance GW. Toward a utility theory foundation for health status index models. Health Serv Res. 1976;11(4):349-369.

178. Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ. 1986;5(1):1-30.

179. Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care. 1989;5(4):559-575.

180. Tsang A, Von Korff M, Lee S, et al. Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders. J Pain. 2008;9(10):883-891.

181. Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ. 2002;11(4):341-353.

182. Uki J, Mendoza T, Cleeland CS, Nakamura Y, Takeda F. A brief cancer pain assessment tool in Japanese: the utility of the Japanese Brief Pain Inventory--BPI-J. J Pain Symptom Manage. 1998;16(6):364-373.

183. van Gils PF, Tariq L, Hamberg-van Reenen HH, van den B M. Kosteneffectiviteit van Preventie. Bilthoven; 2009.

184. van der Pol M, Shiell A. Extrinsic Goals and Time Tradeoff. Med Decis Mak. 2007;27(June):406-413.

185. van Nooten F, Brouwer W. The influence of subjective expectations about length and quality of life on time trade-off answers. Health Econ. 2004;13(8):819-823.

186. van Nooten FE, Koolman X, Brouwer WBF. The influence of subjective life expectancy on health state valuations using a 10 year TTO. Health Econ. 2009;18(5):549-558.

187. van Nooten FE, Koolman X, Busschbach JJ V, Brouwer WBF. Thirty down, only ten to go?! Awareness and influence of a 10-year time frame in TTO. Qual Life Res. 2014;23(2):377-384.

188. van Nooten FE, van Exel NJA, Koolman X, Brouwer WBF. “Married with children” the influence of significant others in TTO exercises. Health Qual Life Outcomes. 2015;13:94.

189. van Nooten FE, van Exel NJA, Eriksson D, Brouwer WBF. “Back to the future”: Influence of beliefs regarding the future on TTO answers. Health Qual Life Outcomes. 2016;14:4.

190. van Osch SMC, Wakker PP, van den Hout WB, Stiggelbout AM. Correcting biases in standard gamble and time tradeoff utilities. Med Decis Making. 2004;24(5):511-517.

191. van Osch S.M.C. SA. The development of the Health-Risk Attitude Scale. In: The Construction of Health State Utilities. Leiden: Leiden University Medical Center (LUMC),Leiden University; 2014.

192. van Wetering CR, Hoogendoorn M, Mol SJM, Rutten-van Mölken MPMH, Schols AM. Short- and long-term efficacy of a community-based COPD management programme in less advanced COPD: a randomised controlled trial. Thorax. 2010;65(1):7-13.

193. Vermunt JK MJ. Latent class cluster analysis. In: Hagenaars J MA, ed. Applied Latent Class Analysis. Cambridge University Press; 2002:89-106

194. Vermunt, JK. Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. Political Analysis. 2010: 18, 450-469.

195. Versteegh MM, Brouwer WBF. Patient and general public preferences for health states: A call to reconsider current guidelines. Soc Sci Med. 2016;165:66-74.

196. Wang XS, Cleeland CS, Mendoza TR, et al. Impact of cultural and linguistic factors on symptom reporting by patients with cancer. J Natl Cancer Inst. 2010;102(10):732-738..

197. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res. 2010;12(5):e53.

198. Wells CD, Murrill WB, Arguedas MR. Comparison of health-related quality of life preferences between physicians and cirrhotic patients: implications for cost-utility analyses in chronic liver disease. Dig Dis Sci. 2004;49(3):453-458.

199. Whitehead SJ, Ali S. Health outcomes in economic evaluation: The QALY and utilities. Br Med Bull. 2010;96(1):5-21.

200. Witney AG, Treharne GJ, Tavakoli M, et al. The relationship of medical, demographic and psychosocial factors to direct and indirect health utility instruments in rheumatoid arthritis. Rheumatology (Oxford). 2006;45(8):975-981.

201. Woloshin S, Schwartz LM, Moncur M, Gabriel S, Tosteson AN. Assessing values for health: numeracy matters. Med Decis Making. 2001;21(5):382-390.

202. Wouters S, van Exel NJA, Rohde KIM, Brouwer WBF. Are all health gains equally important? An exploration of acceptable health as a reference point in health care priority setting. Health Qual Life Outcomes. 2015;13:79.

203. Zarate V, Kind P, Chuang L-H. Hispanic valuation of the EQ-5D health states: a social value set for Latin Americans. Value Health. 2008;11(7):1170-1177.

204. Zax M, Takahashi S. Cultural influences on response style: comparisons of Japanese and American college students. J Soc Psychol. 1967;71(1):3-10.

205. ZiN. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. 2015;(november): 51. Diemen: Zorginstituut Nederland. https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg. Accessed December 16, 2016.

Summary

SUMMARY

Healthcare budgets are limited and decisions have to be made regarding which treatments to reimburse, and which not. To facilitate this decision-making process, economic evaluations are increasingly used. Economic evaluation concerns the assessment of the costs and benefits of an intervention, relative to an appropriate comparator. When performed in health care, costs are typically expressed in monetary terms while health benefits are expressed in terms of Quality-Adjusted Life-Years (QALYs). The QALY combines length and health-related quality of life into a single measure of outcome. In order to calculate QALYs, quality weights (or utility values) need to be attached to different health states. As explained in the chapter 1, the introduction to this thesis, various methods can be used to do so, including visual analogue scales (VAS), the standard gamble (SG) method, the time trade off (TTO) method, and discrete choice experiments (DCE). In each of these methods, respondents are asked to reveal their preferences between health states. Their responses are subsequently analysed in order to calculate quality of life weights or utilities for the valued health states. Different methods and different operationalisations of the same method have been shown to lead to different quality of life weights for the same health state. Consequently, the choice of method may have implications for reimbursement decisions, stressing the importance of improving our understanding responses to health state valuation exercises. This thesis focused on TTO, as its one of the preferred methods for obtaining quality of life weights. In a TTO exercise, a respondent is asked to trade off length of life against quality of life. This is typically done by deriving a point of indifference between two streams of health: one involving living a longer period in some impaired health state, the other living a shorter period in perfect health. The derived point of indifference gives the value of the impaired health state relative to perfect health.

Considering the potential implications for reimbursement decisions, understanding the validity of the TTO method and what influences TTO scores is important. In that context, understanding the influence of respondent characteristics on TTO scores is important for (i) sampling purposes, (ii) the ability to compare TTO scores across studies and populations, and (iii) understanding the observed TTO scores. Many respondent characteristics have the potential to influence TTO and corresponding quality-of-life weights, but in most cases their exact influence remains unclear. This is further discussed in Chapter 2.

Following this discussion, this thesis further explored the influence of several respondent characteristics and attitudinal factors on TTO scores, with the aim to contribute to a better understanding of responses to TTO exercises. To that end, the following research questions were addressed:

1. Does subjective life expectancy (SLE) impact the willingness to trade off (WTT) and the number of years traded-off for health state valuations?
2. How does the awareness of the reduced life-expectation, implied by a 10-year TTO, affect health state valuations?
3. What is the influence of beliefs regarding future health and death, as well as desires to witness certain life events, on health state valuations?
4. Which responder characteristics influence TTO responses, with an emphasis on consideration of significant others, such as partners and children?
5. Do cultural differences in TTO responses exist?
6. Can classes of respondents be identified based on their response patterns in TTO exercises, and which respondent characteristics predict membership of the identified subgroups?

SIx chapters in this thesis (chapters 3-8) provided answers to these questions, through several empirical TTO studies. In all these studies, respondents from the general public answered TTO questions to value hypothetical health states. In addition, several respondent characteristics and attitudinal variables were collected in these studies, including demographic characteristics, future expectations regarding health and age of death, as well as beliefs about life and death.

Research question 1 was investigated in chapters 3 and 4, using a sample of the general public from the Netherlands. Subjective life expectancy was shown to affect health state valuations. This result was evident in TTO exercises using a 10-year time frame as well as in those using a lifetime time frame. SLE was negatively correlated to willingness to trade, meaning that the number of years respondents were willing to trade was lower when SLE was higher; the implication is that it is important to capture respondents' SLE when performing TTO exercises in order to be able to account for the effect of SLE on health state valuations. Chapter 4 also explored the influence a respondents' awareness of reduced life-span in TTO exercises (research question 2). Trading off more years to regain quality of life implies reducing one's lifespan in a TTO exercise and bringing closer the time of death. The results in chapter 4 indicated that explicating and emphasizing the reduced life span is important to make respondents aware of this fact. Respondents who were made explicitly aware of the shortened life expectancy were less willing to trade off years compared with those who were not made explicitly aware of this. Providing explicit information on the reduced life expectancy may thus impact TTO responses. Next, considering that shortening the remaining lifespan seems to play an important role in TTO exercises, in Chapter 5 we investigated the role of respondents' beliefs regarding death and death-related issues (such as opinion about euthanasia, fear of death, desire to witness certain life events, future health expectations) (research question 3). Our results showed that beliefs regarding death and future health expectations were associated with TTO responses as well. Greater fear of death was associated with less years traded-off, and people who do not oppose euthanasia were more willing to trade off years. Hence, chapters 3 to 5 indicate that respondents' beliefs regarding future health and death may have an impact on health state valuations based on TTO, and need to be considered when conducting TTO exercises.

Other respondent characteristics were also investigated. Chapter 6 focused on the consideration of significant others, such as partners and children, in TTO exercises (research question 4). Using a sample from the Dutch general public, we found that respondents indeed take (the impact on) others into consideration in TTO exercises. The results showed that having a partner or children matters, although the signs of the associations differed, indicating that some of these relational factors may have more complex relationships with TTO scores. Chapter 7 explored whether cultural differences can impact TTO scores (research question 5) by comparing TTO responses valuing EQ-5D-5L scores obtained in samples from the UK and Japan. The results showed that compared with respondents in the UK, Japanese respondents were more willing to trade off years for the mildest health states and less willing to trade off years for severe states. However, although the TTO scores for the same health states differed between the UK and Japan, no specific patterns emerged. The influence of cultural differences may thus be complex and requires further study. Chapter 8 investigated whether patterns can be observed in responses to TTO exercises (research question 6), using latent class analyses on TTO data of a sample from the Dutch general public. Two main groups of TTO responders were identified: low and high traders, where low traders on average trade off less years than high traders. Compared to individuals in the high-trader group, those in the low-trader group were more likely to be female, had a higher subjective life expectancy, and self-reported a stronger preference for quantity over quality of life. Interestingly, low and high traders could be identified upfront by asking respondents directly to indicate whether they preferred quality of life over quantity of life.

Finally, chapter 9 presented the discussion of this thesis. Several limitations of the different studies were mentioned. Some of the more important ones concern that most of the empirical studies in this thesis did not perform an iterative elicitation (ping-pong) method to derive indifference points, used web-based questionnaires to obtain the data, and were performed in The Netherlands. These limitations are important to note and emphasize the need for replication and future research.

Notwithstanding these limitations, the main aim of this thesis was to contribute to a better understanding of which respondent characteristics and attitudinal factors influence TTO responses. Overall, this thesis provided additional supportive evidence that TTO responses are affected by several respondent characteristics, in particular subjective life expectancy, but also attitudes and beliefs regarding future health, significant others, and death may also influence TTO scores. Besides increasing our understanding of TTO responses, such respondent characteristics may also be important in the context of sampling, especially when comparing results from different TTO studies, or aiming to derive representative health state valuations, for instance to be used for national tariffs.

In conclusion, while much more work can and should be done in this area, this thesis hopes to have contributed to a better understanding of the influence of ‘the person behind the TTO’ on observed health state valuations, and the relevance of doing so even better in the future.

Samenvatting

SAMENVATTING

Omdat het zorgbudget eindig is, moeten er beslissingen genomen worden over welke behandelingen vergoed kunnen worden en welke niet. Om dit besluitvormingsproces te faciliteren, worden in toenemende mate economische evaluaties gebruikt. In een economische evaluatie worden de kosten en baten van een interventie berekend, ten opzichte van een vergelijkbare reeds bestaande interventie. Binnen de gezondheidszorg worden in economische evaluaties kosten meestal gemeten in geld, terwijl de baten, de gezondheidseffecten, worden gemeten in zogenaamde "Quality-Adjusted-Life-Years" (QALYs). De QALY combineert lengte en kwaliteit van leven en drukt deze beide uit in een enkele waarde. Om QALYs te kunnen berekenen, moeten waarderingen (of utiliteiten) worden toegekend aan verschillende gezondheidstoestanden. Zoals in hoofdstuk 1, de introductie van dit proefschrift, wordt uitgelegd, bestaan hiervoor verschillende methodes, waaronder Visuele Analoge Schalen (VAS), de "Standard Gamble" (SG) methode, de "Time Trade-Off" (TTO) methode en discrete keuze experimenten (DCE). In elk van deze methoden wordt aan respondenten gevraagd om hun voorkeur voor verschillende gezondheidstoestanden aan te geven. De antwoorden worden vervolgens geanalyseerd om de waarderingen van de voorgelegde gezondheidstoestanden te berekenen. Het is gebleken dat verschillende methodes en verschillende operationalisaties van dezelfde methode tot verschillende waarderingen van dezelfde gezondheidstoestand kunnen leiden. Als gevolg hiervan kan de keuze voor een bepaalde uiteindelijk waarderingsmethode ook van invloed zijn op vergoedingsbeslissingen. Dit onderstreept het belang van het verbeteren van ons begrip van waarderingen van gezondheidstoestanden en de factoren die daarbij een rol spelen. Dit proefschrift concentreerde zich op de TTO methode, omdat dit een veel gebruikte en aanbevolen methode is om gezondheidstoestanden mee te waarderen. In een TTO wordt een respondent gevraagd om lengte van leven uit te ruilen tegen kwaliteit van leven. Meestal gebeurt dit door te onderzoeken hoeveel korter leven in een perfecte gezondheid voor een respondent even goed is als een bepaalde periode leven in een slechtere gezondheid. Uit dit zogeheten indifferentiepunt kan vervolgens de waarde van de verminderde gezondheidstoestand in verhouding tot die van perfecte gezondheid worden afgeleid.

Gezien de mogelijke implicaties op vergoedingsbeslissingen, is het belangrijk om de antwoorden op TTO vragen te begrijpen en de factoren die deze antwoorden mogelijk beïnvloeden. Hierbij is ook de invloed van respondentkenmerken belangrijk, bijvoorbeeld voor (1) het samenstellen van een goede steekproef uit het algemene publiek, (2) de mogelijkheid om TTO-uitkomsten te vergelijken tussen verschillende studies en bevolkingsgroepen, en (3) het begrijpen van geobserveerde TTO uitkomsten. Veel kenmerken van respondenten hebben de potentie om TTO antwoorden te beïnvloeden, en dus de afgeleide waardering van gezondheidstoestanden. Hun exacte invloed is echter veelal onduidelijk. Dit is in hoofdstuk 2 uitvoeriger besproken.

Vervolgens is in dit proefschrift dieper ingegaan op de invloed van een aantal persoonskenmerken en gedragsaspecten van respondenten op TTO antwoorden, met als doel om bij te dragen aan een beter begrip van antwoorden op TTO vragen. Daartoe werden de volgende vragen onderzocht:

1. Heeft de subjectieve levensverwachting (SLE) invloed op de bereidheid om levensduur op te geven (WTT) en op het aantal uitgeruilde jaren in een TTO?
2. Hoe beïnvloedt in een TTO met een tijdshorizon van 10 jaar het besef van een gereduceerde levensverwachting de waarderingen van gezondheidstoestanden?
3. Wat is de invloed van overtuigingen betreffende toekomstige gezondheid en dood, en de wens bepaalde levensgebeurtenissen nog mee te maken, op de waarderingen van gezondheidstoestanden?
4. Welke kenmerken van respondenten, in het bijzonder de belangen van anderen zoals partners of kinderen, beïnvloeden TTO-antwoorden?
5. Bestaan er culturele verschillen in TTO-uitkomsten?
6. Kunnen er klassen van respondenten worden onderscheiden op basis van hun antwoordpatronen in TTO onderzoek en hoe hangen die klassen samen met respondentkenmerken?

In zes hoofdstukken in dit proefschrift (hoofdstukken 3 tot en met 8) werden deze vragen beantwoord met behulp van empirische TTO onderzoeken. Hierbij werden respondenten uit de algemene bevolking gevraagd om hypothetische gezondheidstoestanden te waarderen. Ook werd in deze studies informatie verzameld over allerlei kenmerken en opvattingen van de respondenten, zoals demografische kenmerken, verwachtingen betreffende toekomstige gezondheid en levensverwachting, en overtuigingen aangaande aspecten van leven en dood.

Onderzoeksvraag 1 werd in de hoofdstukken 3 en 4 onderzocht, gebruikmakend van een steekproef van de algemene Nederlandse bevolking. Er werd aangetoond dat de waarderingen van gezondheidstoestanden samenhing met de subjectieve levensverwachting van respondenten. Dit resultaat werd gevonden voor TTO's met een tijdshorizon van tien jaar en die met tijdshorizon die gelijk was aan de (gemiddelde) levensverwachting. De subjectieve levensverwachting (SLE) was negatief gecorreleerd met de bereidheid om levensduur uit te ruilen. Het aantal jaren dat respondenten wilden opgeven om hun gezondheid te verbeteren was dus lager naarmate hun subjectieve levensverwachting hoger was. Het is dus belangrijk om de SLE van een respondent vast te stellen om rekening te kunnen houden met het effect van SLE op de waarderingen van gezondheidstoestanden. In hoofdstuk 4 werd ook de invloed op de TTO antwoorden van het besef van respondenten aangaande de verminderde levensverwachting onderzocht (onderzoeksvraag 2). Als door een respondent in een TTO exercitie meer jaren worden uitgeruild om een betere levenskwaliteit te krijgen, betekent dit in de context van een TTO een kortere levensduur en dus dat het tijdstip van overlijden dichterbij wordt gebracht. De resultaten in hoofdstuk 4 laten zien, dat het uitleggen en benadrukken hiervan van invloed is op het aantal jaren dat respondenten uitruilen. Respondenten aan wie de verkorte levensduurverwachting expliciet duidelijk was gemaakt, waren daarna minder bereid om jaren uit re ruilen, ten opzichte van respondenten die deze uitleg niet hadden gekregen. Dit suggereert dat expliciete informatieverstrekking over de kortere levensduur invloed op TTO antwoorden kan hebben. In hoofdstuk 5 werd, gezien de aard van een TTO exercitie, vervolgens nader ingegaan op de rol van overtuigingen van respondenten betreffende toekomstige gezondheid en de dood, zoals euthanasie, angst om te sterven, en de wens om bepaalde levensgebeurtenissen nog mee te maken, (onderzoeksvraag 3). De resultaten lieten zien, dat overtuigingen met betrekking tot de dood en toekomstige gezondheidsverwachtingen ook verband hielden met TTO antwoorden. Grotere angst voor de dood hing samen met een lagere uitruil van jaren. Een positievere houding tegenover euthanasie hing samen met een grotere bereid om jaren op te geven voor een betere kwaliteit van leven. Samengevat kan er gesteld worden dat overtuigingen van respondenten met betrekking tot toekomstige gezondheid en de dood van invloed kunnen zijn op waarderingen van gezondheidstoestanden in een TTO.

Er werd ook onderzoek gedaan naar de invloed van andere kenmerken van respondenten op TTO antwoorden. Hoofdstuk 6 onderzocht of de invloed van belangrijke anderen, zoals partners en kinderen, op TTO antwoorden (onderzoeksvraag 4). In een TTO exercitie in een steekproef uit de Nederlandse bevolking bleken respondenten beïnvloed te worden door het hebben van een partner of kinderen. Echter, het teken van de correlaties varieerde, Dit duidt erop dat deze relationele factoren mogelijk in een complexe verhouding staan tot TTO antwoorden. In hoofdstuk 7 werd uitgezocht of culturele verschillen van invloed kunnen zijn op TTO uitkomsten (onderzoeksvraag 5). Dit werd gedaan door TTO antwoorden voor EQ-5D-5L waarderingen te vergelijken tussen steekproeven uit het Verenigd Koninkrijk (VK) en Japan. De resultaten toonden aan dat vergeleken met het VK, Japanse respondenten eerder geneigd waren om jaren te ruilen voor de mildste gezondheidstoestanden maar juist minder geneigd jaren uit te ruilen voor ernstige gezondheidstoestanden. Hoewel de TTO uitkomsten verschilden tussen beide landen voor dezelfde gezondheidstoestanden, kwamen er geen specifieke patronen naar boven. De invloed van culturele verschillen kan dus complex zijn en vereist verder onderzoek. Hoofdstuk 8 onderzocht of bepaalde patronen kunnen worden waargenomen in TTO antwoorden (onderzoeksvraag 6), met behulp van ”latente klasse analyse”. In van een steekproef uit de algemene Nederlandse bevolking konden twee hoofdgroepen worden geïdentificeerd onder de respondenten: lage en hoge uitruilers. Vergeleken met de personen in de hoge uitruil groep, waren die in de lage uitruil groep eerder vrouw, hadden zij een hogere subjectieve levensverwachting en rapporteerden ze zelf een sterkere voorkeur voor een langer leven in plaats van een kwalitatief goed leven. Belangwekkend was, dat veel lage en hoge uitruilers van tevoren geïdentificeerd konden worden, door rechtstreeks naar hun voorkeur betreffende kwaliteit versus lengte van leven te vragen.

Ten slotte bevatte hoofdstuk 9 de discussie van dit proefschrift. Verscheidene beperkingen van de verschillende studies werden geïdentificeerd. Enkele van de belangrijkste waren (i) dat de meeste empirische studies in dit proefschrift geen iteratief waarderingsproces gebruikten (zoals een ping-pong methode) om indifferentiepunten te achterhalen, (ii) het gebruik van internet-gebaseerde vragenlijsten om de gegevens te verkrijgen, en (iii) dat de studies hoofdzakelijk in Nederland werden uitgevoerd. Het is belangrijk om deze beperkingen te onderkennen. Deze onderstrepen ook de noodzaak om de studies te herhalen en verder onderzoek te doen.

Ondanks de genoemde beperkingen was het voornaamste doel van dit proefschrift om bij te dragen aan een beter begrip van welke kenmerken en overtuigingen van respondenten TTO antwoorden beïnvloeden. Samenvattend leverde dit proefschrift aanvullend ondersteunend bewijs dat TTO antwoorden samenhangen met verschillende kenmerken van respondenten, in het bijzonder de subjectieve levensverwachting, maar ook overtuigingen met betrekking tot toekomstige gezondheid, de dood, en attitudes ten aanzien van belangrijke anderen, zoals partner of kinderen. Naast het vergroten van begrip omtrent TTO antwoorden, kunnen dergelijke kenmerken van respondenten ook belangrijk zijn in de context van het samenstellen van representatieve steekproeven (bijvoorbeeld bij het bepalen van nationale tarieven), en bij het vergelijken van resultaten van verschillende TTO studies.

Tot slot, hoewel er op dit belangrijke onderzoeksterrein nog veel meer werk gedaan kan en moet worden, hoopt dit proefschrift bij te dragen aan een beter begrip van de invloed van 'de persoon achter de TTO' op waargenomen waarderingen van gezondheidstoestanden, en de relevantie van toekomstig onderzoek op dit gebied.

Curriculum Vitae

PhD Portfolio

Dankwoord

CURRICULUM VITAE

Floortje Eline van Nooten was born in 1976 in Heilbronn, Germany, as daughter of Sebastiaan Eliza van Nooten and Nellie Antonia van Gijn-van Nooten. She graduated from secondary school at the European School Munich in Munich, Germany in 1994. Subsequently she went to study pharmacy and obtained her pharmacy degree from the University of Utrecht, the Netherlands in 2001. Following her graduation as a pharmacist she went to study Health Economics at the University of York, York, United Kingdom and concluded in 2002 the MSc in Health Economics. After finalization of her MSc she started working as a researcher at the institute for Medical Technology Assessment in 2002. Her research focused on cost-effectiveness of medical interventions. She changed employer in 2006 and joined as associate researcher at United BioSource Corporation (UBC) (now Evidera) first in Brussels, Belgium later in London, UK. She continued working for UBC until 2012. After leaving UBC, she continued her career at Astellas as associate director health economics in Leiden, the Netherlands. In 2016 she joined Dompé Farmaceutici as head of market access global in Milan, Italy.

She has published papers in peer-reviewed journals, such as Health Economics, Value in Health, Health Quality of Life Outcomes, Pharmacoeconomics and European Journal of Health Economics.

PHD PORTFOLIO

Publications

Finkelstein FO, van Nooten F, Wiklund I, Trundell D, Cella D. Measurement properties of the Short Form-36 (SF-36) and the Functional Assessment of Cancer Therapy - Anemia (FACT-An) in patients with anemia associated with chronic kidney disease. Health Qual Life Outcomes. 2018 May 31;16(1):11

van Nooten F, Busschbach J, van Agthoven M, van Exel J, Brouwer W.What should we know about the person behind a TTO? Eur J Health Econ. 2018 Apr 18 (Epub ahead of print)

Hwang S, van Nooten F, Wells T, Ryan A, Crawford B, Evans C, English M. Neuropathic pain: A patient-centred approach to measuring outcomes. Health Expect. 2018 Aug;21(4):774-78

van Nooten FE, Cline J, Elash CA, Paty J, Reaney M. Development and content validation of a patient-reported endometriosis pain daily diary.Health Qual Life Outcomes. 2018 Jan 4;16(1):3

van Nooten FE, Houghton K, van Exel J, van Agthoven M, Brouwer WBF, Stull DE. A (Latent) Class of Their Own: Response Patterns in Trading Off Quantity and Quality of Life in Time Trade-Off Exercises. Value Health. 2017 Dec;20(10):1403-1410.

van Nooten F, Treur M, Pantiri K, Stoker M, Charokopou M. Capsaicin 8% Patch Versus Oral Neuropathic Pain Medications for the Treatment of Painful Diabetic Peripheral Neuropathy: A Systematic Literature Review and Network Meta-analysis. Clin Ther. 2017 Apr;39(4):787-80

Feng Y, Herdman M, van Nooten F, Cleeland C, Parkin D, Ikeda S, Igarashi A, Devlin NJ. An exploration of differences between Japan and two European countries in the self-reporting and valuation of pain and discomfort on the EQ-5D. Qual Life Res. 2017 Mar 25. doi: 10.1007/s11136-017-1541-5. [Epub ahead of print]

Korol E, Wang S, Johnston K, Ravandi-Kashani F, Levis M, van Nooten F, Health-Related Quality of Life of Patients with Acute Myeloid Leukemia: A Systematic Literature Review. Oncology and Therapy, January 2017

van Nooten F, Trundell D, Staniewska D, Chen J, Davies EW, Revicki DA. Evaluating the Measurement Properties of the Self-Assessment of Treatment Version II, Follow-Up Version, in Patients with Painful Diabetic Peripheral Neuropathy.Pain Res Treat. 2017;2017:6080648. doi: 10.1155/2017/6080648. Epub 2017 Jan 16.

Eriksson D, Goldsmith D, Teitsson S, Jackson J, van Nooten F. Cross-sectional survey in CKD patients across Europe describing the association between quality of life and anaemia.BMC Nephrol. 2016 Jul 26;17(1):97.

Van Nooten FE, van Exel NJ, Eriksson D, Brouwer WB. "Back to the future": Influence of beliefs regarding the future on TTO answers.Health Qual Life Outcomes. 2016 Jan 12;14:4

Tomaszewski EL, Fickley CE, Maddux L, Krupnick R, Bahceci E, Paty J, van Nooten F.
The Patient Perspective on Living with Acute Myeloid Leukemia. Oncol Ther. 2016;4(2):225-238

van Nooten FE, van Exel NJ, Koolman X, Brouwer WB. "Married with children" the influence of significant others in TTO exercises. Health Qual Life Outcomes. 2015 Jul 2;13:94

Yingxin Xu, Kyle Fahrbach, Floortje van Nooten, Grace Jennings. Early Network Meta-Analyses (NMAs): Filling a Need in Clinical Trial Designs and HTA Submissions. Evidence Forum (Evidera external newsletter), March 2015

Alleman CJ, Westerhout KY, Hensen M, Chambers C, Stoker M, Long S, van Nooten FE. Humanistic and economic burden of painful diabetic peripheral neuropathy in Europe: A review of the literature. Diabetes Res Clin Pract. 2015 May 6.

van Nooten FE, Winnette R, Stein R, Kissner M, Schröder A, Jöckel M, Raluy-Callado M, Lambrelli D, Meinhardt M, Wasiak R Resource utilization and productivity loss in persons with spina bifida—an observational study of patients in a tertiary urology clinic in Germany. Eur J Neurol. 2015 Jan; 22(1):53-8.

Bowles D, Wasiak R, Kissner M, van Nooten F, Engel S, Linder R, Verheyen F, Greiner W. Economic burden of neural tube defects in Germany. Public Health. 2014 Mar;128(3):274-81

F van Nooten , J Jaime Caro .Use of relative effectiveness information in reimbursement and pricing decisions in Europe. Journal of Comparative Effectiveness Research, January 2013, Vol. 2, No. 1, Pages 33-44.

van Nooten F, Stern S, Braunstahl GJ, Thompson C, Groot M, Brown RE. Cost-effectiveness of omalizumab for uncontrolled allergic asthma in the Netherlands.J Med Econ. 2013;16(3):342-8.

van Nooten F, Holmstrom S, Green J, Wiklund I, Odeyemi IA, Wilcox TK.Health economics and outcomes research within drug development: challenges and opportunities for reimbursement and market access within biopharma research. Drug Discov Today. 2012 Jun;17(11-12):615-22

Petrillo J, van Nooten F, Jones P, Rutten-van Mölken M.Utility Estimation in Chronic Obstructive Pulmonary Disease (COPD): A Preference for Change? Pharmacoeconomics 2011

van Nooten F, Davies GM, Jukema JW, Liem AH, Yap E, Hu XH. Economic evaluation of ezetimibe combined with simvastatin for the treatment of primary hypercholesterolaemia. Neth Heart J. 2011

van Nooten FE, Green J, Brown R, Finkelstein FO, Wish J. Burden of illness for patients with non-dialysis chronic kidney disease and anemia in the United States: review of the literature.
J Med Econ. 2010;13(2):241-56.

Floortje van Nooten. Reimbursement in a World of Uncertainty. Evidence Matters (UBC external newsletter), October 2010 (Also in special booklet UBC created!!!)

Green J, van Nooten F, Thwaites R, Lenderking W, The growing importance of health technology assessments (HTAs) in reimbursement decision-making, Evidence Matters (UBC external newsletter), March 2010 van Wetering CR, van Nooten FE, Mol SJM, Hoogendoorn M, Rutten-van Mölken MPMH, Schols AM. Systemic impairment in relation to disease burden in patients with moderate COPD eligible for a lifestyle program. Findings from the INTERCOM trial. Int J Chron Obstruct Pulmon Dis. 2008; 3(3):443-51.

van Nooten FE, Koolman AHE, Brouwer WBF. The influence of subjective life expectancy on TTO utilities: a new experiment. Health Econ. 2009 May;18(5):549-58.

Rutten-van Mölken MPMH, van Nooten FE, Lindemann M, Caeser M, Calverley PMA. A 1-Year Prospective Cost-Effectiveness Analysis of Roflumilast for the Treatment of Patients with Severe Chronic Obstructive Pulmonary Disease (COPD). Pharmacoeconomics 2007;25(8):695-711.

van Nooten FE, van Agthoven M. Diagnose Behandeling Combinaties (DBC's): Visies en verwachtingen van het veld. ZM Magazine 2006; 1: 10-13.

van Nooten FE, Stolk EA, Steenhoek A. De definitie van specialistische geneesmiddelen. Institute for Medical Technology Assessment, Erasmus MC, Rotterdam, 2006.

Stolk EA, van Nooten FE. Values for resource allocation should expose the adaptation process, not the outcome. Virtual Mentor. 2005 Feb 1

van Nooten FE, Brouwer WBF. The influence of subjective expectations about length and quality of life on time trade-off answers. Health Econ 2004; 13: 819-823.

Brouwer WBF, Van Nooten FE, Rutten FFH. Nog veel te winnen met betere richtlijnen. TSG 2004; 8: 523-526.

Brouwer WBF, Hulst E, van Nooten FE, Rutten FFH. Generieke substitutie. Wat zijn de bevoegdheden van arts, apotheker en patiënt? TSG 2003; 3: 162-165.

Rendering JA, van Nooten FE, de Jong JGAM. Naar een uniforme interne markt. Verkrijgbaarheid van geneesmiddelen binnen de Europese Unie. Pharmaceutisch Weekblad 2001; 48: 1798-1802.

DANKWOORD

Het spreekwoord dat altijd in mij opkwam als ik over mijn proefschrift nadacht, was “lang verwacht, stil gezwegen, nooit gedacht, toch *af*gekregen”. Ooit kreeg ik te horen, dat ik geen volhoudingsvermogen zou hebben, misschien dat dit proefschrift daarom wel zo lang duurde tot voltooiing.

Na zoveel jaar aan dit proefschrift te hebben gewerkt, zijn er heel veel mensen om te bedanken voor hun bijdrage aan dit proefschrift.

Werner, als allereerste wil ik natuurlijk jou bedanken. Het begon met een afstudeerscriptie voor de universiteit van York. Ik zocht een plek en via Eddy kwam ik bij jou terecht, want jij had een idee dat ik uit kon werken. Ons eerste gesprek vond (samen met Job) op jullie kamer plaats. Ik herinner met het nog goed (je was niet meneer Brouwer, maar Werner). Dat idee is nu de basis van dit hele proefschrift geworden. En daar ben ik je dankbaar voor. De data voor mijn afstudeerscriptie en het allereerste artikel heb ik toen al fietsende door Nederland verzameld. Daarna was je ook bereid mij te ondersteunen, toen ik voorstelde er een proefschrift van te maken en je zal mij sindsdien best vervloekt hebben, kan ik mij zo voorstellen (hoop dat niet alle grijze haren door mij komen). Ik vond onze meetings op allerlei plekken in Nederland of telefonisch altijd leerzaam en gezellig. Zal ze ook wel missen. Bedankt voor alles.

Job, jij was er ook sinds het begin bij en ben dan ook blij dat je er ook op het einde bij bent. Hartelijk voor al je steun gedurende alle jaren.

Aan de promotiecommissie wil ik ook mijn dank uitspreken voor het lezen, beoordelen en opponeren van dit proefschrift.

De meeste hoofdstukken in dit proefschrift heb ik niet alleen gedaan en wil dan ook alle co-auteurs bedanken voor hun bijdrage en inzicht. Met name Xander ben ik dank verschuldigd voor de vele uren die we samen hebben doorgebracht om de analyses te bespreken en te runnen. Mies, ook jou wil ik bedanken voor de vele gezellige uren die we besteed hebben aan het brainstormen over de analyse vragen, doorlezen van de hoofdstukken en het doorspitten van artikelen. Fijn dat je nu ook mijn paranimf wil zijn.

Verder wil ik Maureen hartelijk danken voor het mij wegwijs maken in de wetenschappelijke wereld tijdens mijn iMTA dagen. Je hebt mij geleerd hoe je het goed moet doen.

Daarnaast wil ik alle talloze ex- en collega's bedanken voor hun kennis die zij mij met mij wilden delen, wat mij altijd weer verder hebben geholpen.

Mijn vrienden wil ik danken voor hun nooit aflatende belangstelling en steun voor mijn proefschrift (en in mij). Hoe vaak ik de afgelopen vele jaren wel niet de vraag te horen heb gehad "is het af?" kan ik niet meer tellen. Dank jullie wel, thank you, danke schön, grazie.

Mijn familie wil ik ook hartelijk danken voor hun steun. Ze zijn er altijd als ik ze nodig heb, maakt niet uit hoe groot de fysieke afstand is. Pap dank voor het helpen van het voltooien van het proefschrift, en mam dank voor de vele uren aan de telefoon. Lieve pap en mam en Bastiaan, jullie zijn mijn grote helden.