Examining Sources Of Variation In Developing A Societal Health State Value Set

BY

ERNEST H. LAW
B.Sc.(Pharm), University of Alberta, 2009
Pharm.D., University of British Columbia, 2014

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Pharmacy
in the Graduate College of the
University of Illinois at Chicago, 2018

Chicago, Illinois

Defense Committee:

A. Simon Pickard, Chair and Advisor
Todd A. Lee
Surrey M. Walton
Alan Schwartz, Medical Education
Feng Xie, McMaster University

To Su and Caeden.

ACKNOWLEDGMENTS

First and foremost, I wish to thank my dissertation committee - Drs. Simon Pickard (chair), Todd Lee, Surrey Walton, Alan Schwartz, and Feng Xie – for pushing me to be better with each discussion and communication. I am sincerely grateful for the time and energy they have spent supporting this work that I am very proud of.

I am especially indebted to Simon for his mentorship and advice over the years. I will miss our regular chats where we often traversed the continuum of music, comedy, pop culture, quantum physics, IPAs, and, of course - research.

I am grateful to all the faculty in the department of Pharmacy, Systems, Outcomes, and Policy for the example they project as scientists, resolute in their passion for inquiry and guiding the "next generation" of researchers.

To my fellow graduate students and the research assistants I've had the pleasure of meeting. I consider it an honour to get to know each of them. It's safe to say that they are all a little… weird… but they have become my family over the years – so, *cheers to that*.

To my parents, James and Shirley Law - I am convinced I will never know the true depths of their sacrifices that have allowed me to walk this path, but I do know that they will always have my gratitude and admiration.

Lastly, my boundless love and appreciation to Su and Caeden. Su - for all the thoughtfulness, care, and patience you have given me during this long pursuit of knowledge. I am reconstituted daily by your humour and grace. Caeden - your interminable curiosity never fails to inspire me and there is not a worry I carry in my shoulders that cannot be lightened by your laughter. Both of you are the best things to happen to me and I look forward to future adventures together. - EHL

CONTRIBUTION OF AUTHORS

I (EHL) hold primary responsibility for all aspects of the research in this dissertation and should be considered lead author for all chapters. I prepared all initial drafts for all chapters and conducted all analyses for each study.

ASP provided guidance and critical review throughout my doctoral studies and dissertation preparation; he contributed to the design and interpretation of all studies and should be considered secondary author for Chapters 2, 3, and 4.

TL provided guidance and review of the design and interpretation of the results for all three studies (Chapters 2, 3, and 4).

SW provided guidance and review of the design and interpretation of the results for all three studies (Chapters 2, 3, and 4).

AS provided guidance and review of the design and interpretation of the results for all three studies (Chapters 2, 3, and 4).

FX provided guidance and review of the design and interpretation of the results for all three studies (Chapters 2, 3, and 4).

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF ABBREVIATIONS

3L	EQ-5D-3L
5L	EQ-5D-5L
AD	Anxiety/Depression
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BTD	Better-Than-Dead
CBA	Cost Benefit Analysis
CEA	Cost-Effectiveness Analysis
CI	Confidence Interval
cTTO	Composite Time Trade-Off
CUA	Cost Utility Analysis
DC	Discrete Choice
DCE	Discrete Choice Experiment
EQ-VT	EuroQol Valuation Technology
HRQOL	Health-Related Quality of Life
HTA	Health Technology Assessment
HUI	Health Utilities Index
ICC	Intra-class Correlation Coefficient
IQR	Interquartile Range
LT-TTO	Lead Time Time Trade-Off
MAE	Mean Absolute Error

LIST OF ABBREVIATIONS (continued)

MO	Mobility
NICE	the National Institute of Health and Care Excellence
OR	Odds Ratio
PD	Pain/Discomfort
QALY	Quality-Adjusted Life Year
SC	Self-Care
SD	Standard Deviation
SE	Standard Error
SG	Standard Gamble
TTO	Time Trade-Off
UA	Usual Activities
US	United States
VAS	Visual Analogue Scale
WTD	Worse-Than-Dead

SUMMARY

This dissertation aimed to examine sources of variation in developing a societal value set for health preferences using a generic preference-based measure. The EQ-5D was selected as a case study for this dissertation because of its prominence among health technology assessment (HTA) agencies worldwide, and because it is the most widely used measure for producing quality-adjusted life years used for economic evaluations. Three studies were conducted to: (1) compare value sets developed using two different descriptive systems; (2) compare health values elicited at different points in time; and (3) determine if respondents with advance directives, who have reflected about the experience and consequences of a range of health states, have different stated preferences for health states.

This dissertation is composed of five chapters. **Chapter 1** provides an introduction of this dissertation before presenting an overview of several concepts and issues essential to understanding the significance of this dissertation: economic evaluations and the role of health valuation, societal health value sets, direct and indirect approaches to eliciting health state values, and potential sources of variation within and between societal value sets. A conceptual framework and list of specific aims for each study are also introduced.

Chapter 2 details the first study which aimed to compare and contrast EQ-5D-5L ('5L') and EQ-5D-3L ('3L') societal value sets derived from a common sample in order to better understand how modifications to a descriptive system potentially impacts cost utility analyses. This study utilized data from the 2017 US EQ-5D-5L valuation study where respondents valued 3L and 5L health states using similar study designs. Value sets were modeled with random-effects linear regression. Properties of the descriptive system and value set characteristics were compared by examining distributions of predicted index scores, ceiling effects, and single-level transition values from adjacent corner health states. Mean single-level transition values were calculated for all predicted 3L and 5L health states and plotted against baseline index scores. A total of 1,062 respondents were included in the analysis. The ranges of scale for the 5L and 3L were 0.973 ("11111") to -0.356 ("55555") and 0.921 ("11111") to -0.662 ("33333"), respectively. Values for the mildest 5L health states ranged from 0.888 to 0.924 and were similar to 11111 (0.922) for the 3L. Parameter estimates for matched dimension-levels differed by < |0.07| except for the most severe level of Mobility (3L-level 3: -0.525 vs. 5L-level 5: -0.262). Mean transition values calculated using the 3L value set were greater for lower baseline 3L index scores, whereas the mean transition value remained constant irrespective of the baseline 5L index score. This study provides insight into the characteristics of values derived using the 3L and 5L descriptive systems based on data that minimizes structural confounding. Compared to the 3L, the value set derived using the 5L descriptive system exhibited improved measurement properties and a reduced ceiling effect. The 3L value set had a larger range of scale; however, this difference was driven by the difference in weights assigned to the most severe level of problems in Mobility for the 3L ("confined to bed") and 5L ("unable to walk about").

The second study (**Chapter 3**) asked, "*Do health preferences differ between individuals with and without advance directives?*" This study sought to better understand the implications of informed preferences by examining the association between advance directives for health care and health preferences. Data from the 2017 US EQ-5D-5L valuation study was analyzed. Using advance directive status and values obtained from 10 TTO exercises, regression models fitted health values to estimate the impact of advance directives, adjusting for respondent characteristics. Logistic regression models examined the association between advance directive status and likelihood of generally valuing health states as worse-than-dead (WTD) and at least one of the 10 health states as WTD during the TTO exercises. Of 1061 respondents, 27.4% had

an advance directive. Advance directives were associated with significantly lower mean values assigned to overall health states (difference=-0.101; 95%CI -0.175 to -0.028) and specific dimension-levels: severe and extreme problems with Mobility, slight problems with Usual Activities, severe problems with Pain/Discomfort, and severe problems with Anxiety/Depression. Advance directive status was not associated with an increased likelihood of valuing health states as WTD (OR=1.62; 95%CI 0.99-2.60; p=0.053) or at least one of the 10 health states as WTD (OR=1.29; 95%CI 0.91-1.84; p=0.146). This study provides evidence of how values obtained from those with advance directives may differ from those who have not considered end-of-life health and treatments. If advance directives are an indicator of a substantial proportion of the population with “informed” preferences, it has implications for resource allocation based on cost-utility analysis.

The third and final study is presented in **Chapter 4** and asked, “*What are the time-specific differences in societal preferences elicited between two eras?*” The need to update value sets of preference-based measures of health can be motivated in part by identifying whether the values of the target population have changed. Because of differences in methodology, it is challenging to compare value sets. However, the recently completed US EQ-5D-5L valuation study included a sub-study valuation of the EQ-5D-3L similar to the 2002 EQ-5D-3L valuation study. This study compared EQ-5D-3L valuation studies in 2002 and 2017 to identify if there were time-specific differences in stated preferences for health. Data from 2002 and 2017 EQ-5D-3L valuation studies were combined. The primary focus was to compare valuations of better-than-dead (BTD) states, as both studies used the same time trade-off (TTO) approach. For worse-than-dead (WTD) states, the 2017 study used lead-time TTO, whereas the 2002 study used a modified conventional TTO which necessitated transformation of WTD values. Unadjusted mean TTO valuations were

compared for 16 common EQ-5D-3L health states. Regression models were fitted to BTD values to estimate time-specific differences, adjusting for respondent characteristics. Secondary analyses examined models that fitted WTD values (using linear and non-linear transformations of the 2002 data) and all TTO values. Unadjusted BTD-only mean values in 2017 were significantly higher than 2002 values for 1 of 16 common health states. In the adjusted BTD-only model, mean valuations were significantly higher for 2017 compared to 2002 ($\beta 2017=0.05$, $p<0.001$). Models using WTD data showed negative changes across time that were dependent on the transformation method ($\beta 2017=-0.72$ with linear transformation, $\beta 2017=-0.35$ with nonlinear, both $p<0.0001$). Using all values, mean valuations were lower in 2017 compared to 2002 using a linear transformation ($\beta 2017=-0.11$; $p<0.001$) but did not differ with the non-linear transformation ($\beta 2017=-0.01$; $p=0.5$). For the most methodologically comparable data (BTD only), values in 2017 were modestly higher, which implies people were less willing to trade time for quality of life than in 2002, i.e. 6 months over 10 years. The large differences between 2002 and 2017 when WTD data were included appeared to be driven by differences in methodology, and also illustrated the profound impact of choice of transformation on value sets. Overall, results suggest that the time period when values were elicited may be important and may be one reason to consider updating societal value sets. These results are relevant on a broader global basis, given that some value sets were developed decade(s) ago.

Lastly, **Chapter 5** of this dissertation summarizes the findings of the three papers and provides a general discussion of the implications of each study with respect to policy and future research.

1. INTRODUCTION

1.1 Purpose and Significance

As health care expenditures increase, so has the recognition from many stakeholders that costs must be contained and that health technologies produce value (Anderson et al., 2014; Berwick, Nolan, & Whittington, 2008; Schnipper et al., 2015). The quality-adjusted life year (QALY) garners support as an outcome measure used in the assessment of health technologies to inform healthcare decision-making, though not without its criticisms (F. R. Johnson, 2009; Lipscomb, Drummond, Fryback, Gold, & Revicki, 2009; Nord, Daniels, & Kamlet, 2009). Nevertheless, the number of published cost-utility analyses (CUAs) continues to grow over the past decades even when cost per QALY thresholds are prohibited in the US by the Affordable Care Act legislation (P. Neumann & Weinstein, 2010; P. J. Neumann, Thorat, Shi, Saret, & Cohen, 2015), in part because third-party payers in the US incorporate clinical and economic evaluations in their reimbursement decision-making, which includes CUAs that estimate cost per QALY gains associated with new technology (Ahlstrom et al., 2017; AMCP, 2010; Husereau, Culyer, Neumann, & Jacobs, 2015). Moreover, influential stakeholders, including clinical and independent advisory/guideline setting groups, have endorsed the cost per QALY as important to evaluating the benefit of health technologies (Anderson et al., 2014; ICER, 2017; Sanders et al., 2016).

Preference-based measures are part of the utility-approach to capturing changes in health-related quality of life (HRQOL) and facilitate the calculation of QALYs (Brazier, Ratcliffe, Saloman, & Tsuchiya, 2017). The use of preference-based measures has increased as economic evaluations continue to inform health policy. QALYs generated from preference-based measures can vary depending on which measure is employed, and not just due to the attributes included in the descriptive system (Lipscomb et al., 2009), but many other factors such as the target population from whom the values are obtained (Rowen et al., 2017; Versteegh & Brouwer, 2016), the study design employed (Lamers, McDonnell, Stalmeier, Krabbe, & Busschbach, 2006), and the country/culture of the population (Feng et al., 2017; J. A. Johnson, Luo, Shaw, Kind, & Coons, 2005; Olsen, Lamu, & Cairns, 2017). Additionally, there may be societal shifts in demographics and values over time (A. S. Pickard, 2015). Meanwhile, HTA agencies such as the National Institutes of Health and Care Excellence (NICE) in the UK, have recommended the use of the same preference-based measure for CUA, i.e., the EQ-5D-3L, and its societal value set (Dolan, 1997; NICE, 2017) in order to retain consistency when applying thresholds for reimbursement decisions of new health technologies. There is a tension between the need for consistency and improving the science and the underlying value sets upon which decisions are based.

When new measures and techniques are introduced as the science of measurement and valuation progresses, understanding the back compatibility with existing measures is essential to policy makers. Discrete choice experiments and modifications to the time trade-off have been introduced to the health valuation field over the past decade; many existing preference-based scoring algorithms/value sets were derived decades ago and may not represent current societal preferences; and descriptive systems have been modified for greater content coverage. All of these issues are relevant to the EQ-5D, which is the globally predominant generic preference-based measure used in HTA and available in hundreds of languages and dozens of country-specific value sets ("EQ-5D Instruments | About EQ-5D," ; Richardson, McKie, & Bariola, 2011; Wisløff et al., 2014).

This dissertation examined sources of variation in developing a societal value set for health preferences using a generic preference-based measure. The EQ-5D was selected because of its prominence among HTA agencies worldwide, and because it is the most widely used measure for producing QALYs used in economic evaluations (Richardson et al., 2011; Wisløff et al., 2014).

1.2 Background

The following sections describe key concepts and issues relevant to understanding the significance of this dissertation work.

1.2.1 Economic evaluations and health-related quality of life

In striving to maximize health in the context of limited resources, decision makers in health systems often utilize economic evaluations to inform decisions involving the adoption of health technologies (P. J. Neumann et al., 2015). Performing economic evaluations of health technologies requires quantifying the costs and consequences of adopting an intervention and comparing it to the costs and consequences of an appropriate alternative (Drummond, Sculpher, Claxton, Stoddart, & Torrance, 2015).

The three most common types of economic evaluations are cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and CUA (Drummond et al., 2015). Each type differs in how health consequences are measured. In a CBA, both costs and consequences are quantified in monetary terms. While measuring all costs and consequences in monetary terms simplifies decision-making (i.e., if benefits exceed costs, then the health technology should be adopted), there may be considerable pushback from important healthcare stakeholders (e.g., patients and clinicians) to valuing health in dollars as well as the challenge of translating health benefits in monetary terms. In a CEA, health consequences are measured in natural (or clinical) units, such as life-years gained, fatal and non-fatal myocardial infarctions avoided, or degree to which blood pressure is lowered. One of the major limitations of the CEA is that unless consequences of treatment are measured in common units of health, the evaluations cannot be used to allocate resources across the health care system without additional judgements about trade-offs using different natural units.

Due to the limitations of CBA and CEA, CUAs has gained popularity in healthcare jurisdictions around the world. CUA is a type of CEA where health consequences are measured in terms of QALYs gained (described in detail in the next section). Two or more interventions are compared relative to the standard of care in terms of incremental cost per QALYs gained. For example, NICE typically compares the estimated incremental cost effectiveness ratios (ICER) for a given health technology to a threshold of ~30,000 GBP / QALY (McCabe, Claxton, & Culyer, 2008). By using the QALY as a common currency or measurement of health benefits, CUAs allow the potential for comparing different interventions across different disease contexts and settings.

1.2.2 <u>The Quality-Adjusted Life Year</u>

The QALY integrates two major health benefits that health care systems typically prioritize: reducing mortality and improving HRQOL (Weinstein, Torrance, & McGuire, 2009). To calculate QALYs, a period of time within a specific health state is multiplied by a weight or value that corresponds to the "utility" or "health preference" associated with the health state. There is a debate among researchers as to whether values derived without the use of the standard gamble (SG) technique, a specific method for direct preference elicitation, can be considered "*utilities*", as they are not elicited under uncertainty per von Neumann-Morgenstern's expected utility theory (Drummond et al., 2015). Therefore, in this dissertation, "health value" is used to signify the numerical strength of preference for a given health state.

Within the QALY framework, health values are standardized as index scores that lie on a cardinal scale anchored from 1 (full health or absence of health problems) and 0 (dead) (Brazier, Ratcliffe, et al., 2017; Drummond et al., 2015), with the possibility of negative values if health states are allowed to be deemed unlivable or considered “worse than dead” (WTD). Positive values indicate health states considered better than dead (BTD) (Patrick, Starks, Cain, Uhlmann, & Pearlman, 1994). Constant proportionality between quantity and quality of life is assumed. For example, one QALY equals 1 year in full health (i.e., health value = 1) or four years in a health state valued at 0.25. The (positive) difference in average QALYs between, for instance, a group of patients receiving an intervention and the control group, represents the “QALY gain”, or health benefit. In a CUA, the difference in costs can be divided by the QALY gain to determine the incremental cost per QALY.

In 2016, the Second Panel on Cost-Effectiveness in Health and Medicine (Peter J Neumann, Sanders, Russell, Siegel, & Ganiats, 2016; Sanders et al., 2016) provided recommendations for the “conduct and reporting of cost-effectiveness analyses”, building on recommendations from the First Panel (Gold, 1996; Sanders et al., 2016). A “reference case” was presented to set a standard for all CEA/CUAs to improve quality and comparability of economic evaluations. Several features of the reference-case are notable and reflect the need to for a “common currency” for health benefit that can be used in CUAs. First, the referent CUA should measure health effects in terms of QALYs. Second, “quality weights” should be preference-based and interval-scaled. Third, preferences should be sourced from a “community-based” (or societal) sample, with sensitivity analysis to include patients and other sources where differences in health preferences may be important. Support for using these criteria to improve comparability is echoed by the 2017 Value Framework developed by the Institute for Clinical and Economic Review, an increasingly influential non-profit HTA organization in the US (ICER, 2017). With these recommendations, some payers in the US may join other HTA agencies around the world, such as those in the United Kingdom, Canada, and Australia, that view societal value sets as central to informing health care decision-making (CADTH, 2017; NICE, 2013; PBAC, 2016).

1.2.3 Obtaining health state preferences

Health values can either be directly or indirectly elicited. Direct preference elicitation, is performed by describing one or more health states to an individual and valuing each health state using one of several valuation methods (e.g., time trade-off [TTO] or SG) so that the strength of preference, or value, for each health state can be quantified (Brazier, Ratcliffe, et al., 2017; Drummond et al., 2015). For societal value sets, stated preferences for hypothetical health states are required and therefore direct preference elicitation of one's own health state is not appropriate. Moreover, even if the individual's current health state value is desired, the feasibility of direct preference elicitation as the sole approach to obtaining health values for QALY estimates is limited by the amount of resources required to value the full range of health states described by a health classification system. For instance, the EQ-5D-5L describes 3,125 unique health states, a prohibitive number of health states that can be valued in any given study. The second, more popular, and much less resource intensive approach to eliciting health preferences is through an indirect preference-based measure (Arnold, Girling, Stevens, & Lilford, 2009; Brazier, Ratcliffe, et al., 2017). With this approach, respondents complete a questionnaire that asks them to describe their own health state (or another's in the case of valuation by proxy). This health state is assigned a health value drawn from a "value set" and can be used in an economic evaluation. The value set is analogous to a "catalog" and obviates the need for respondents to complete direct preference elicitation exercises.

Compared to the direct measurement approach, indirect measures are more straightforward once a value set is established. However, developing a value set requires a representative sample of the desired population to directly value a subset of health states. The directly measured (or observed) values are used to estimate values for all possible health states defined by the descriptive system of the preference-based measure. While the indirect measurement approach and the use of “off-the-shelf” value sets to generate health values is relatively straightforward, conducting health valuation studies to develop a value set requires considerable resources, expertise, effort, and is accompanied by several methodological issues. However, CUAs that use the same value set to obtain preference weights or health values to derive QALYs bolsters the comparability of cost per QALY estimates across studies and disease areas.

The development of an indirect preference-based measure can be decomposed into three broad steps: (1) defining the health state descriptive system; (2) directly eliciting preferences for a subset of health states from a sample of the target population; and (3) developing a scoring algorithm that assigns a health value to all possible health states classified by the preference-based measure (Brazier, Ratcliffe, et al., 2017; McDonough & Tosteson, 2007). Each of these steps is described below.

1.2.4 Health state description

The description of a health state is concerned with identifying aspects (or “dimensions”) of HRQOL (e.g., mobility, pain, mental health) that may be affected by a health intervention. Health state dimensions are assigned two or more levels to represent the place at which a respondent falls within the dimensions (e.g., “I have no pain or discomfort” vs. “I have extreme pain or discomfort”). By identifying all relevant dimensions and levels, a descriptive system is established so that a “universe” of all unique health states is characterized. For example, the original EQ-5D descriptive system (now known as the EQ-5D-3L) is one that consists of five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and three levels for each dimension (no problems, some/moderate problems, unable to/extreme problems) which results in a preference-based measure with 243 possible health states (3^5) (Rabin, Oemar, Oppe, Janssen, & Herdman, 2011). Descriptive systems for preference-based measures can also be classified as "generic" (e.g., EQ-5D and HUI) or condition-specific (e.g., DHP-5D for diabetes) (D. Feeny, Furlong, Saigal, & Sun, 2004; Mulhern et al., 2018). The former focuses on dimensions that are considered relevant across most disease contexts, while latter may be more relevant to specific health conditions (e.g., presence/absence of hyperglycemic symptoms in diabetes).

A generic set of domains in a measure often forms the core for most descriptive systems. Most systems contain domains for common symptoms (e.g., pain, depression) and function (physical, social or role) but still may differ in other domains. In other situations, a descriptive system is modified so that the newer version contains the same dimensions, but the number of response levels are different. For example, the aforementioned EQ-5D-3L, which includes three levels for each of the five dimensions, was expanded to the 5-level version, the EQ-5D-5L (M. Herdman et al., 2011). The 5L version has since shown improved discriminative properties and a reduction in ceiling effects compared to the 3L version (Feng, Devlin, & Herdman, 2015; M. Janssen et al., 2013; A. S. Pickard, Kohlmann, et al., 2007).

Taken together, characteristics of a descriptive system may impact the psychometric properties of the preference-based measure, including the ability to measure HRQOL as a numerical index score, and to detect meaningful changes in health.

1.2.5 Health preference elicitation

The next broad step in developing a societal value set is eliciting preferences for a health states described by the classifier system among a sample of the target population. There are several direct preference elicitation methods available to researchers. The most common methods are the visual analogue scale (VAS), SG, TTO, and more recently, the discrete choice experiment (DCE). Different valuations methods have been chosen by developers in developing a societal value sets. For example, the developers of the Health Utility Index (HUI) and SF-6D employ the SG, while the EQ-5D Group has traditionally used the TTO. The choice of valuation method can lead to differences in health values and ultimately impact cost per QALY estimates from CUAs (Bleichrodt, 2002; Doctor, Bleichrodt, & Lin, 2010).

Both DCE and TTO were incorporated by the EuroQol Group into the international protocol developed for valuing EQ-5D-5L health states. Specifically, a variant of the TTO, the composite TTO (cTTO) which uses the conventional TTO to elicit BTD values and the lead-time TTO (LT-TTO) to elicit WTD values (**Figure 1**), and DCE without duration (i.e., time spent in the health state as an attribute) (Brazier, Ara, Rowen, & Chevrou-Severac, 2017; Krabbe et al., 2014; Oppe, Devlin, van Hout, Krabbe, & de Charro, 2014). Each cTTO task begins with the conventional TTO in which the respondent chooses between Life B (10 years in the suboptimal health state being valued) or Life A (10 years in full health). For health states considered severe, individuals may choose to trade off all 10 years in full health. Therefore, the value for that state is *no greater than* 0, which means it cannot be better than dead (BTD) and suggests instead that it may be worse than dead (WTD). In this situation, if the respondent states that the health state is WTD, then the LT-TTO is introduced to elicit values < 0. The respondent is then provided another 10 years in full health to trade, preceding the 10-years in the health state being valued in Life B. Depending on the answer the participant provides, the amount of time in full health in Life B will change using an iterative process. This is continued until the individual indicates indifference between Life A and Life B.

DCEs is a preference elicitation technique that assumes that all goods or services (e.g., health) can be decomposed into its characteristics (or attributes) (Lancsar & Louviere, 2008). Moreover, each attribute can be further defined by levels. The technique involves presenting several tasks or choice sets to a respondent that vary systematically by attribute-levels. Respondents are asked which life they prefer from the choice set (e.g., 'Life A' or 'Life B'). Responses provide information on the relative importance of different attributes by allowing one to compare the size of preference weights between attribute-levels. In the DCE developed for the EQ-5D-5L international protocol, participants are guided through seven discrete choice (DC) tasks. Each task asks the respondents to consider two different EQ-5D-5L health states in a pairwise comparison. They are asked to indicate which state they prefer (**Figure 2**). No opt-out options (e.g., "both state A and B are about the same", "neither state A nor B is preferred", "prefer not to answer", etc.) are provided.

1.2.6 <u>Issues in constructing societal value sets</u>

The following section describe two normative questions in developing health state value sets from a societal perspective: (1) what individual-level factors influence societal preferences; and (2) do societal preferences change over time?

Individual-level variables that influence health valuation

There continues to be a debate among researchers and policy-makers over whose preferences should be captured in deriving a health state value set for the purposes of calculating

QALYs for informing resource allocation. In the context of a decision-maker adopting a societal perspective, the key is to ensure that the health values (and subsequently calculated QALYs) are sourced from a sample of the target population. With the emergence of the patient-centered outcomes movement, there is cachet to the notion of using patients as the source of preferences regarding health care. However, there is may be no clear reason to think a subgroup has values different from the general population unless there are characteristics that are known to contribute to different views and values about health. Therefore, it is important to understand the relationship between individual characteristics and their effect on health values to guide interpretation and application of value sets.

Numerous studies have identified characteristics associated with differences in health values, including age, sex, race/ethnicity, geographical setting, and illness experience on general health preferences (Craig et al., 2014; Kind & Dolan, 1995; Sayah et al., 2016; James W Shaw, Johnson, Chen, Levin, & Coons, 2007). However, many studies lack additional information on family status, religious beliefs, and attitudes towards end-of-life care. Such non-health considerations may be important as a recent qualitative investigation suggests that these factors may play an important role in how individuals form preferences in the context of health valuation. For instance, Karimi et al. interviewed members of the public in the United Kingdom to explore the process by which individuals value health, including the role of the non-health factors in developing health state preferences (M Karimi, J Brazier, & S Paisley, 2017). The authors concluded that individuals consider personal and social circumstances (e.g., availability of support, especially for severe health) when reflecting upon the consequences of living in a hypothetical health state; these circumstances were called “conversion factors”. In a Canadian EQ-5D-5L valuation sub-study, Al-Sayah et al. also noted the importance of considering the non-health factors, such as the burden of the respondent’s health on his/her family (Al Sayah, Mladenovic, Gaebel, Xie, & Johnson, 2016). While these studies provide important insight into how individuals value health through qualitative interviews, few studies have quantified the extent that additional factors influence health value, independent of sociodemographic factors typically collected in valuation studies. Understanding how individual characteristics, including potential conversion factors, impact health values is important to researchers and decision-makers for ensuring value sets adequately represent the preferences of the population.

Individuals who have completed advance care planning constitute a group within the general population who may have reflected substantially on the experience and consequences of a range of health states. Advanced care planning and the use of advance directives involves the discussion and documentation of an individual's preferences concerning their goals of care in the event they lose capacity or communication ability (Emanuel, 1995). Advance directives are meant to increase the autonomy of patients and express their wishes when unable to do so themselves. By completing an advance directive, an individual has made a legally binding, actionable decision based on their health preferences as it pertains to a meaningful life. Past studies have shown high rates of concordance between treatment preferences and severe health states often discussed in advance care planning (Brinkman-Stoppelenburg, Rietjens, & van der Heide, 2014; Patrick et al., 1997). Further, treatment preferences for end-of-life (EOL) were demonstrated to be stable over time across several illness contexts and sufficient in guiding treatment when the patient’s preferences cannot be directly confirmed when a medical decision is required (Auriemma et al., 2014; Barrio-Cantalejo et al., 2013). In the US, all 50 states and the District of Columbia regulate their AD legalities individually, but the process typically involves the completion of a living will and durable power of attorney. National estimates of AD completion vary from 16 to 26% (Rao, Anderson, Lin, & Laux, 2014; Wilkinson, Wenger, & Shugarman, 2007). Taken together, existing literature supports advance directives completion as a valid indicator of reflecting on health preferences and decision-making that remains stable over time.

Do societal preferences change over time?

The selection and availability of value sets for the purposes of economic evaluation should reflect health preferences of the target population. Fifteen years have passed since the valuation of the EQ-5D-3L among the US general public (J. W. Shaw, Johnson, & Coons, 2005). Since 2002, the US population has undergone demographic shifts and values may have changed (Prochaska, 2013). For instance, the percentage of the US population that is 65 years or older is expected to double by 2050 (from 2012), and increasing age is a well-established determinant of health across many settings (Luo, Johnson, Shaw, Feeny, & Coons, 2005; Ortman, Velkoff, & Hogan, 2014; Sayah et al., 2016). Moreover, Americans' personal views on health, particularly in the context of end-of-life medical treatments, is changing (Prochaska, 2013). These changes in population demographics and attitudes towards (severe) health over time are likely to lead to changes in health preferences.

However, no valuations studies have been rationalized on the basis of updating societal value sets. Instead, value sets are often repeated in the same population or country due to the introduction of a new preference-based measure (e.g., EQ-5D-3L to the-5L version), the evolving state of the science in preference elicitation approaches (e.g., conventional TTO to cTTO and DCE), and other methodological improvements that reduce bias (A. S. Pickard, 2015; Feng Xie et al., 2015). Most studies comparing differences in societal value sets focus on cross-country comparisons, but few empirical studies have explored the magnitude and determinants of changes in health preferences over time (Greiner et al., 2003; J. A. Johnson et al., 2005; Olsen et al., 2017). Given that some value sets were developed decade(s) ago (Dolan, 1997; J. W. Shaw et al., 2005), it is important to recognize how changes have occurred to better inform the need to pursue an “update.”

1.3 <u>Conceptual Framework</u>

This section introduces and describes a conceptual framework for this dissertation. First, the process by which an individual forms his/her health state preferences is described. Then, the framework is modified and adapted to a broader framework that outlines the development of a preference-based measure and its value set, identifying each study in the dissertation and its role in informing this framework.

The process by which an individual forms stated preferences for hypothetical health states is theorized to involve several components (**Figure 3**) (M Karimi et al., 2017). The process begins with the presentation of a health state (e.g., EQ-5D health state) to the respondent. The respondent considers the health state description and imagines what that health state would be like to experience (*Link 1*). This linkage forms a concrete interpretation of the health state. The next linkage (*Link 2*) occurs between the concrete interpretation of the health state and “conversion factors”. The term “conversion factors” is a relatively broad construct coined by Karimi and colleagues that refers to four categories of personal and social circumstances that affect how respondents value health: personal interests and circumstances, other people's reaction to ill health, the ability to adapt or change expectations of life, and available support (M Karimi et al., 2017). Once this linkage is made, the individual views the consequences of being in that health state through these conversion factors in the final linkage (*Link 3*). As a result of this linkage, health preferences are formed.

The framework established by Karimi et al. was modified and adapted to the broader process of developing a societal value set using a generic preference-based measure (**Figure 4**).

From this modified framework, sources of variation discussed in Chapter 2 are highlighted and the studies proposed in this dissertation are linked. **Study 1** evaluated how modifications to health state descriptive systems (3L vs. 5L) impact health preferences. As the health state is the point at which the health valuation process begins, one can expect the descriptors of the health state to have downstream effects on the health preferences elicited from respondents. **Study 2** examined whether the stated preferences of individuals with advance directives are different, as advance directive completion was hypothesized to be a conversion factor that may be a reflective indicator of informed preferences. Lastly, **Study 3** evaluated whether stated preferences for health are different between across two points in time, adjusting for known demographic shifts, in order to determine if value sets should be updated because values change.

1.4 Research Questions

This dissertation addressed the following research questions:

1) In what ways do value sets developed using the 3L and 5L version of the EQ-5D descriptive system differ?

2) Do health preferences differ between individuals with and without advance directives?

3) Do societal value sets differ over time?

1.5 Specific Aims

The study aims for each study proposed in this dissertation are outlined below.

Study 1 – In what ways do value sets developed using the 3L and 5L version of the EQ-5D descriptive system differ?

- **Aim 1.1**: Compare the range of scale for TTO values generated using the -5L versus -3L versions of the EQ-5D
  - **Hypothesis 1.1**: A wider range of scale is observed for all predicted EQ-5D-3L health states (i.e., 11111 to 33333) compared to the 5L health states (i.e., 11111 to 55555).
    - **Rationale**: It was anticipated that labelling difference for the most severe level of problems in Mobility between the 3L ("Confined to bed") and the 5L ("Unable to walk about") would result in an overall greater disutility for the 3L, resulting in a larger range of scale.

- **Aim 1.2**: Compare relative preference weights for each health dimension generated using 5L and 3L descriptive systems.
  - **Hypothesis 1.2a**: Relative preference weights for the "Mobility" dimension will be significantly larger for the 3L compared to the 5L version, due to differences in labelling for the most severe level.
  - **Hypothesis 1.2b**: The ordering of preference weights for each dimension will be different between the 3L and 5L version with Mobility being the most important dimension for the 3L value set, due to the label for level 3 ("Confined to bed").
    - **Rationale**: Similar to the rationale for Hypothesis 1.1, it was expected that labelling difference for the most severe level of problems in Mobility would result in the Mobility dimension having the largest preference weight in the 3L compared to the 5L.

- **Aim 1.3**: Compare number of health states valued as worse than dead (WTD)

- **Hypothesis 1.3**: The 3L value set will produce a higher proportion of health states with negative values due to the disutility for Mobility level 3 (“Confined to bed”).

- **Aim 1.4**: Compare differences in utility between adjacent health states (i.e., differences between health states with only one dimension and on-level change.
  - **Hypothesis 1.4 (null)**: The change in adjacent states will be similar for the 3L compared to the 5L value set across all health dimensions.

- **Aim 1.5**: Compare differences in overall mean transition values for all possible transitions for all possible healths states.
  - **Hypothesis 1.5 (null)**: The mean transition values will be similar for the 3L compared to 5L.

Study 2 – Do health preferences differ between individuals with and without advance directives?

- **Aim 2.1**: To identify respondent characteristics independently associated with advance directive completion.
  - **Hypothesis 2.1**: Individual who are White, male, and older with post-secondary education are more likely to complete advance directives based on previous literature.

- **Aim 2.2**: To examine mean differences in health values between individuals with and without advance directives, accounting for respondent characteristics.
  - **Hypothesis 2.2 (null)**: There is no overall differences in mean values for EQ-5D-

5L health states between those with and without ADs.

- **Aim 2.3**: To examine differences in dimension-level preference weights for EQ-5D-5L health states between those with and without advance directives for TTO valuations.
    - **Hypothesis 2.3 (null)**: There are no differences in dimension-level preference weights for EQ-5D-5L health states between those with and without.

- **Aim 2.4**: To assess the likelihood of valuing health states as worse-than-dead by advance directive completion status.
    - **Hypothesis 2.4 (null)**: There is no difference in the likelihood of valuing health states as WTD between those with and without advance directives.

- **Aim 2.5**: To compare the overall mean transition values derived from individuals with and without advance directives
    - **Hypothesis 2.5 (null)**: The overall mean transition value will be similar for respondents with advance directives relative to those without advance directives

Study 3 – Do societal value sets differ over time?

- **Aim 3.1**: To compare mean values estimated from valuations obtained in 2002 and 2017.
    - **Hypothesis 3.1 (null):** There are no differences in mean TTO values between 2002 and 2017.

- **Aim 3.2**: To compare relative preference weights for each health dimension between 2002 and 2017.

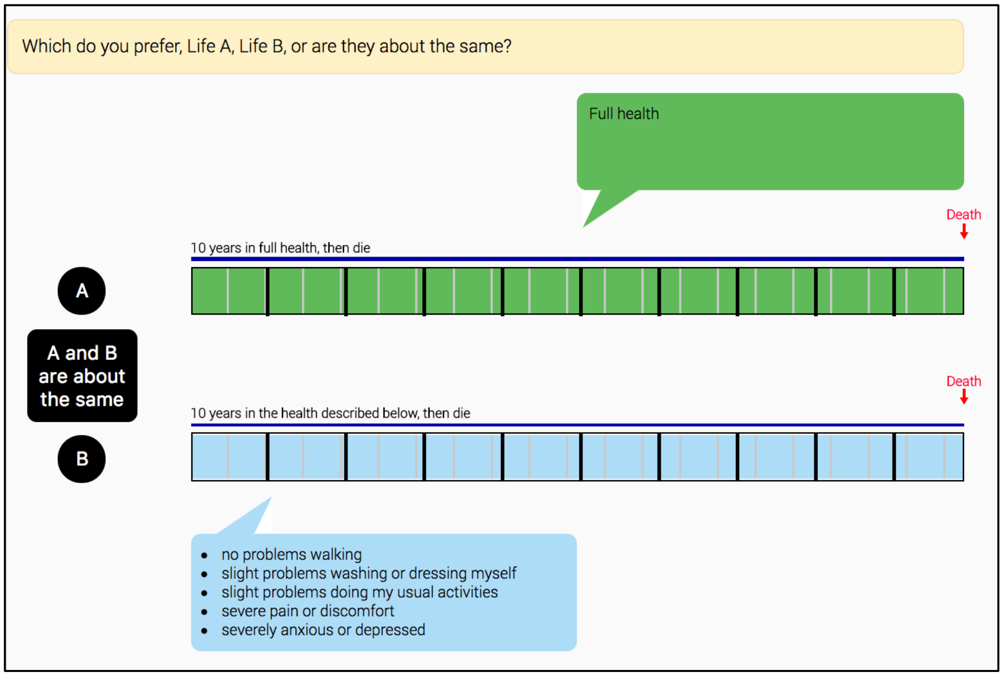
- **Hypothesis 3.2 (null):** There will be no differences in relative preference weights for each dimension-level dimension between 2002 and 2017.

- **Aim 2.3**: To compare the likelihood of valuing healths states as WTD between 2002 and 2017.
  - **Hypothesis 3.3 (null):** There will be no differences in relative preference weights for each dimension-level dimension between 2002 and 2017.

FIGURES

Figure 1 Examples of composite time trade-off tasks for better than dead and worse than dead health states used in the EuroQol Valuation Technology platform

a) Better than dead health state



b) Worse than dead health state

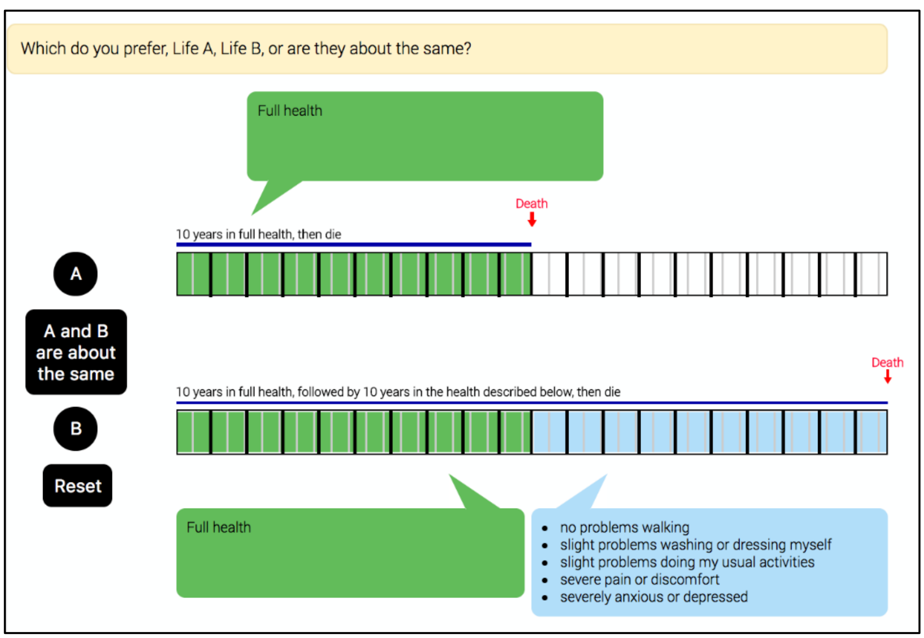


Figure 2 Example of discrete choice task used in the EuroQol Valuation Technology

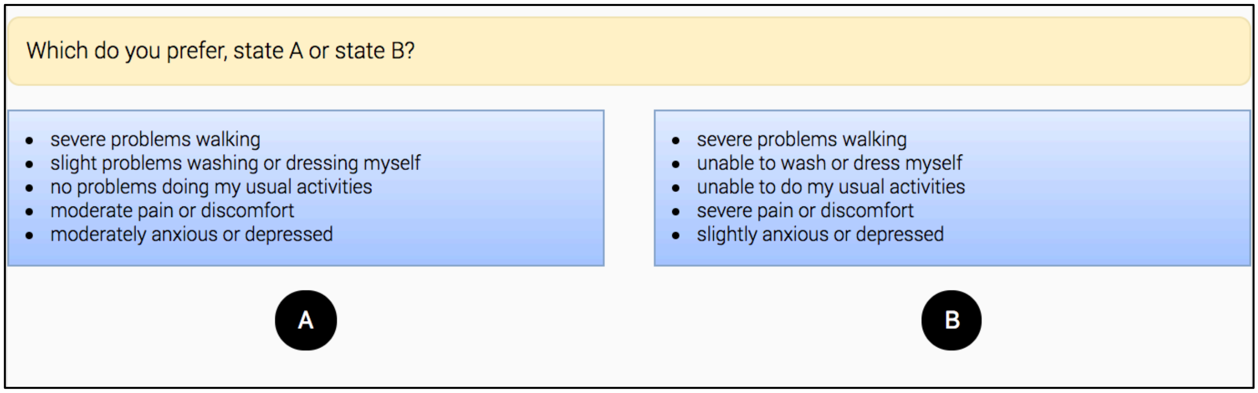


Figure 3 Process by which individuals value health (M Karimi et al., 2017)

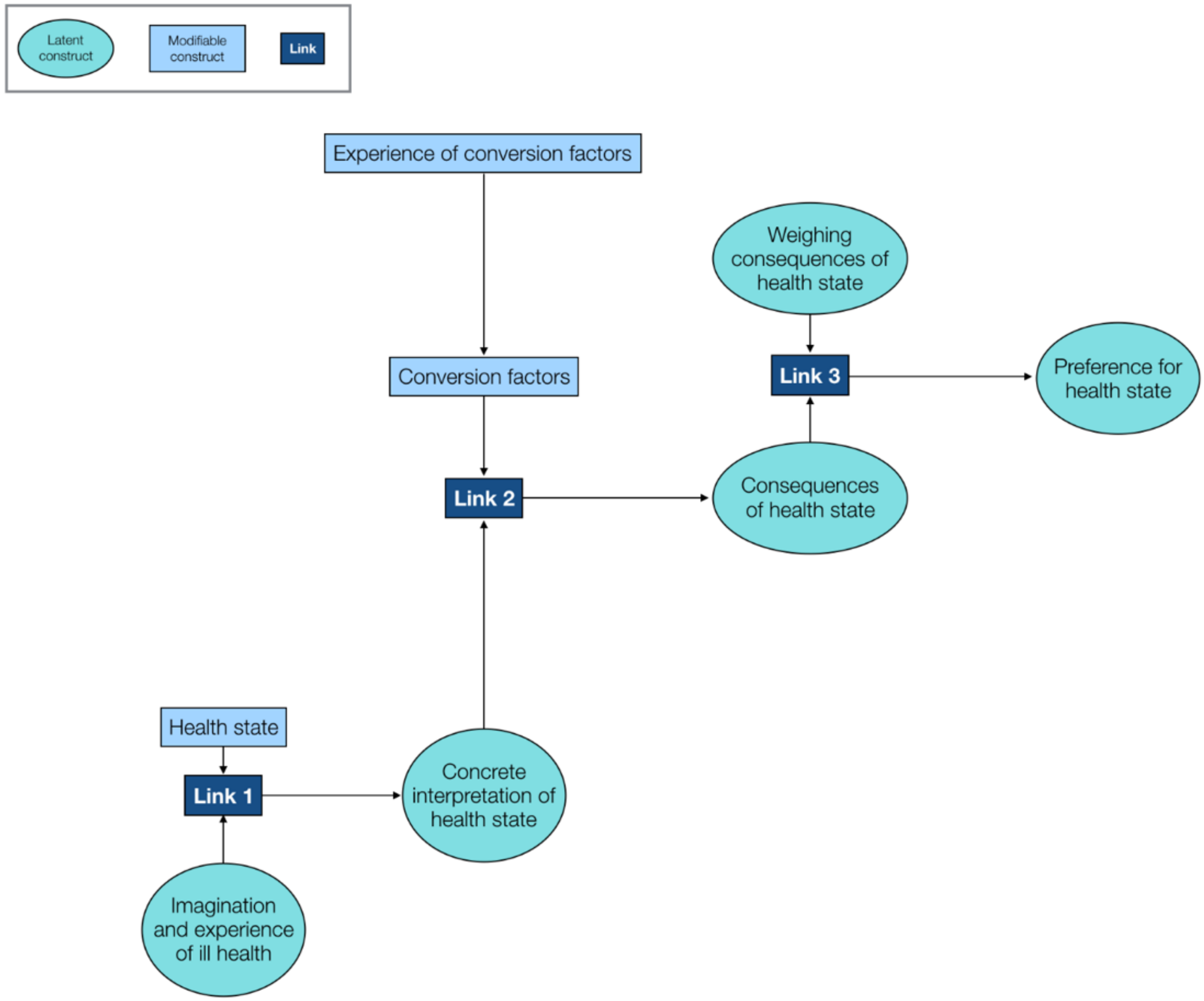
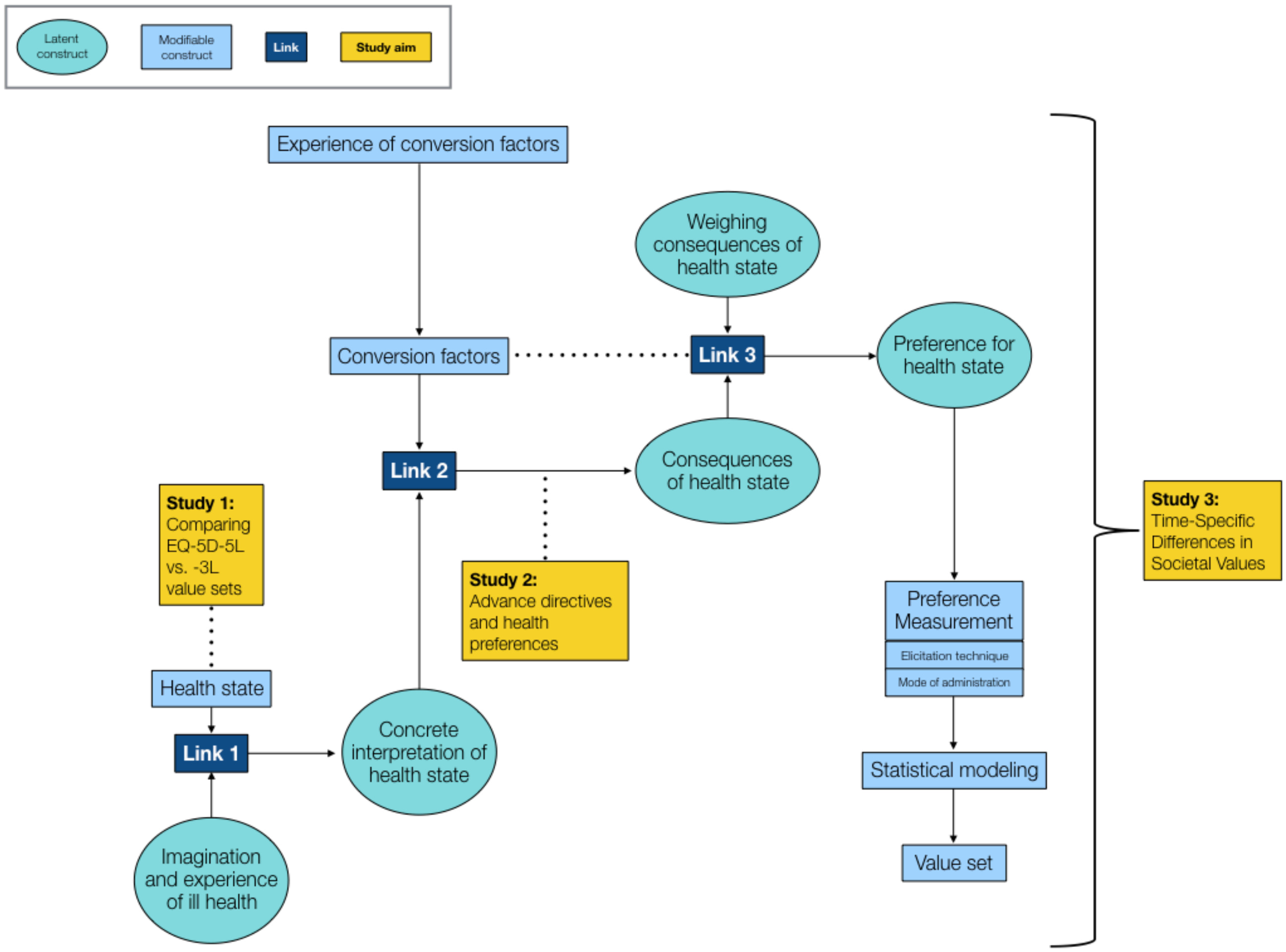


Figure 4 Modified framework outlining the process by which societal value sets are developed using individual health preferences



(Adapted from Karimi et al. 2017)

2. PARALLEL VALUATION: A DIRECT COMPARISON OF EQ-5D-3L AND EQ-5D-5L SOCIETAL VALUE SETS

Authors: Ernest H. Law, PharmD; A. Simon Pickard, PhD; Feng Xie, PhD; Surrey M. Walton, PhD; Todd A. Lee, PharmD, PhD; Alan Schwartz, PhD (***submitted to Medical Decision Making)***

2.1 Abstract

Objective: To compare and contrast EQ-5D-5L ('5L') and EQ-5D-3L ('3L') societal value sets derived from a common sample. **Methods**: Data from the 2017 United States EQ-5D valuation study was analyzed. Respondents provided composite time trade-off (cTTO) valuations for 3L and 5L health states. Value sets were modeled with random-effects linear regression. Properties of the descriptive system and value set characteristics were compared by examining distributions of predicted index scores, ceiling effects, and single-level transition values from adjacent corner health states. Mean single-level transition values were calculated for all predicted 3L and 5L health states and plotted against baseline index scores. **Results**: A total of 1,062 respondents were included in the analysis. The ranges of scale for the 5L and 3L were 0.973 ("11111") to -0.356 ("55555") and 0.921 ("11111") to -0.662 ("33333"), respectively. Values for the mildest 5L health states ranged from 0.888 to 0.924 and were similar to 11111 (0.922) for the 3L. Parameter estimates for matched dimension-levels differed by < |0.1| except for the most severe level of Mobility (3L-level 3: -0.525 vs. 5L-level 5: -0.262). Mean transition values calculated using the 3L value set were greater for more severe baseline 3L index scores, whereas the mean transition value remained constant irrespective of the baseline 5L index score. **Conclusions**: Compared to

the 3L, value sets developed using the 5L exhibit a lower ceiling effect and improved measurement properties. There was a larger range of scale for the 3L compared to 5L; however, this difference was driven by differences in preference for the most severe level of problems in Mobility for the 3L (“confined to bed”) and 5L (“unable to walk about”).

2.2 Introduction

The EQ-5D descriptive system is the most widely used preference-based measure in economic evaluations (Richardson et al., 2011; Wisløff et al., 2014). It is a relatively brief and simple measure to complete, with a descriptive system consisting of five dimensions (Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression) (Rabin & Charro, 2001). The EQ-5D-3L ("3L"), which has three levels for each of the five dimensions, was expanded to the 5-level version, the EQ-5D-5L ("5L) in response to criticisms that the 3L lacked sensitivity to small changes in health (M. Herdman et al., 2011). In the process, the most severe level of the Mobility dimension of the 3L version ("confined to bed") was changed to "unable to walk about" **(Appendix A**). Value sets for the 5L are now published or forthcoming for countries throughout the world, including England (N. J. Devlin, Shah, Feng, Mulhern, & van Hout, 2018; Mulhern et al., 2018), South Korea (Kim et al., 2016), Uruguay (Augustovski et al., 2016), Canada (F. Xie et al., 2016), Japan (Shiroiwa et al., 2016), Indonesia (Purba et al., 2017), and Germany (Ludwig, Graf von der Schulenburg, & Greiner, 2018) which may be used in health technology assessment (HTA).

HTA agencies desire consistency in decision-making and seek to avoid potential "gaming" of the results of cost-utility analyses (NICE, 2017). Thus, given the central role of the EQ-5D in HTA, it is important to understand differences in societal value sets derived from the 3L and 5L descriptive systems. As a measure of self-reported health status, several advantages of the 5L over the 3L have been noted: a reduction in ceiling effects, an increase in unique self-reported health states, and an improved ability to discriminate between patient groups (N. Devlin, Brazier, Pickard, & Stolk, 2018; Feng et al., 2015; M. Janssen et al., 2013). However, less is known as to how differences in the measurement properties impact health valuations using the 5L compared to the 3L. Recent literature indicates that index scores and quality-adjusted life years generated from the 3L and 5L versions of the EQ-5D are different (Alava et al., 2017; Hernández-Alava & Pudney, 2017; M. F. Janssen, Bonsel, & Luo, 2018; Mulhern et al., 2018). However, the vast majority of these studies relied on existing 3L and 5L value sets, which typically differed in several ways: different respondents at different time periods (sometimes decades apart) and geographic locations, using dissimilar study protocols (which included different quality control processes and preference elicitation techniques). Such differences represent major sources of variation that are known to influence the values obtained. Therefore, it is challenging to isolate any differences between value sets due to changes in the descriptive system.

This study aims to compare and contrast 5L and 3L societal value sets derived from the same respondents and employing the same protocol, preference elicitation technique, and statistical modelling approach to determine how even small changes in a descriptive system may affect the preferences elicited.

2.3 Methods

2.3.1 Data source

This study utilizes data collected during the United States (US) EQ-5D-5L health state valuation study conducted in 2017 (A. Pickard et al., 2018) and was a methodological sub-aim incorporated into the main study design to allow a direct comparison of 3L and 5L value sets developed using time trade-off (TTO) values. The study included 1134 non-institutionalized adults (≥ 18 years of age) recruited from the US adult general public. Respondents were excluded from the present analysis if they failed to comprehend the TTO preference elicitation task per interviewer assessment.

2.3.2 Preference elicitation

The composite time trade-off (cTTO) was used to elicit health state values for the present analysis. Discrete choice data were collected for 5L health states in the main study, but not for 3L states. The cTTO method uses the conventional TTO to elicit "better than dead" (BTD) values and the lead-time TTO (LT-TTO) to elicit worse-than-dead (WTD) values. Additional details regarding the use of the cTTO in valuing EQ-5D health states have been reported previously (Attema, Edelaar-Peeters, Versteegh, & Stolk, 2013; B. M. Janssen, Oppe, Versteegh, & Stolk, 2013). Briefly, all cTTO tasks begin with the conventional TTO with a 10-year time horizon in the state being valued (Life B) and 10 years in full health that can be traded (Life A). For health states in Life B that are considered severe enough, respondents can trade all 10 years in Life A (full health). Therefore, the value for that state is at best equal to 0, which means it cannot be BTD and suggests it may be WTD. In this situation, if the respondent states that the health state being valued is WTD, then the LT-TTO is introduced to elicit values < 0. In the LT-TTO, the respondent is only given 10 additional years in Life A once, restricting the minimum TTO value to -1 or higher. Depending on the answer the participant provides, the amount of time in full health in Life A will change using an automated iteration process. This process continues until the participant indicates that they are indifferent between Life A and Life B, resulting in a cTTO value ranging from -1 to 1.

2.3.3 EQ-5D descriptive systems

Health states were described using the 5L and 3L descriptive systems. In the 5L, the five levels for each dimension are "no problems," "slight problems," "moderate problems," "severe problems" and "unable to" for the functional dimensions (Mobility, Self-Care, Usual Activities) and "extreme problems" for the affective dimensions (Pain/Discomfort, Anxiety/Depression) (M.

Herdman et al., 2011). A dimension in which there are “no problems” is designated level 1, while a dimension where there are extreme problems is designated level 5. Every distinct health state described by the 5L system is assigned a five-digit descriptor that ranges from 11111 (absence of problems) to 55555 (worst possible or “PITS” state), with each digit representing one dimension-level of health, resulting in a descriptive system that defines 3,125 (5^5) health states.

For the EQ-5D-3L, the three levels are “no problems” (level 1) and “some problems” for the functional dimensions and “moderate problems” for the affective dimensions (level 2) (van Reenen & Oppe, 2015). For level 3, Mobility is assigned “confined to bed”, “unable to” for Self-Care and Usual Activities, and “extreme problems” for Pain/Discomfort and Anxiety/Depression. Each unique health state for the EQ-5D-3L is also represented by a five-digit descriptor but ranges from 11111 to 33333 (worst possible or PITS state). The 3L descriptive system results in a total of 243 (3^5) possible health states.

A “misery score” can also be calculated as a proxy indicator for health state severity by summing the five digits of the health state descriptor. For example, the misery score for “13231” is equal to 10. The highest misery score for the 3L and 5L is 15 (“33333”) and 25 (“55555”), respectively.

2.3.4 Health state selection

For the 5L, 86 health states were included and selected for valuation (**Table 1)** (Oppe et al., 2014). 5L states were grouped into 10 blocks with 10 health states per block. All 10 blocks contained the PITS state (“55555”), one mild state, and eight health states that are unique to the block (Oppe et al., 2014). A total of 30 EQ-5D-3L states were selected for valuation: 18 states were chosen using the orthogonal design developed by Yang et al. (Yang, Luo, Bonsel, Busschbach, & Stolk, 2017), the 5 mildest states (11112, 11121, 11211, 12111, 21111), and 7 common health states used in the 2005 US EQ-5D-3L health valuation study, while taking into account dimension-level balance for the overall experimental design (**Table 2** (J. W. Shaw et al., 2005). The 3L health states were divided into 10 blocks of three states (one mild, one moderate, and one severe). Blocks were randomly assigned to each respondent, and the order or presentation within each block was also randomized. All respondents completed the 10 5L cTTO tasks first followed immediately by the three 3L tasks.

2.3.5 <u>Statistical analysis</u>

To compare and contrast 3L and 5L value sets, two separate random effects linear regression models were developed based on cTTO values for 3L and 5L health states. A random intercept at the respondent level was specified to account for respondents valuing multiple health states. Since all states were directly valued using the same cTTO method, all values were bounded by -1 and 1, obviating the need for any transformations (Patrick et al., 1994). Each model assumed changes between each dimension-level are independent. For the 3L, a dummy variable represents the difference measured between level 1 and level 2 and another for the difference between level 1 and level 3 for a total of 10 parameters plus the model intercept. Similarly, there were four dummy variables for each 5L dimension resulting in a total of 20 parameters plus the intercept. The constant term was interpreted as the average TTO score for 11111, where all dummy variables are equal to zero (i.e., level 1). Model parameters are interpreted as the decrement from 11111 (intercept) to 0 (the value for dead), though values may extend to negative space and be considered WTD.

The adequacy of each model's performance was evaluated on the basis of several criteria: construct validity of internal response structure (i.e., statistically significant and logically ordered coefficients), goodness-of-fit, and prediction accuracy between observed and predicted health state values. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were calculated for each model to evaluate goodness-of-fit. Prediction accuracy (i.e., predicted vs. observed values) was assessed by calculating Pearson's r, mean absolute error (MAE), and the proportion of absolute prediction errors that were greater than 0.05 and 0.1 for the 30 and 86 health states for the 3L and 5L, respectively.

The characteristics of the 3L and 5L value sets were described based on: point estimates for each model coefficient, value set range (i.e., difference between 11111 and the PITS state), range for each dimension (i.e., coefficient size of the most severe levels), number and proportion of health states valued as WTD, and estimates for the moderate (i.e., 22222 [3L] vs. 33333 [5L]) and PITS health states (i.e., largest utility decrement from 11111). Ceiling effects for both value sets were assessed by comparing the intercepts and their distance from 1, where the value 1 represents perfect health (distinct from no problems in all dimensions, i.e., 11111) (King Jr, Styn, Tsevat, & Roberts, 2003). In particular, we compared the 3L intercept with the five mildest health states in the 5L (11112, 11121, 11211, 12111, 21111) given that one of the main motivations for developing the 5L was to reduce the ceiling effect of the EQ-5D and increase its sensitivity for milder health states (M. Herdman et al., 2011). The modality of overall distributions was evaluated using a kernel density histogram.

Several analyses were conducted to understand the how the magnitude of gains and losses in health values with changes in health state severity, which in turn may impact QALY calculations. First, differences in changes in single-level transitions between adjacent states were compared at both ends of the utility scale: the mildest health states (i.e., movements away from 11111) and the most severe health states (i.e., movements away from the PITS health state). In comparing the mildest and most severe single-level transitions, 3L changes were calculated and compared to two 5L health state transition calculations: an unmatched health state (i.e., the adjacent health state to 11111 and 55555) and a “matched” health state (i.e., where level 2 on the 3L is matched to level 3 on the 5L).

Second, transition values between adjacent 3L and 5L corner health states were calculated to compare specific differences in transition values for a change in a single dimension by a single level (e.g., 21111 and 31111), holding all other dimensions constant. Changes in predicted values were calculated between level 2, 3, 4 and 5, with all other dimensions fixed at level 1 (no problems). For example, with Mobility, changes were compared between 5L health states 51111, 41111, 31111, 21111, and 11111. For the 3L health states, changes were compared between levels 2 and 3 (e.g., 31111 and 21111). Similar comparisons were made for the remaining health dimensions. To illustrate differences in values for comparable health states, “matched” differences were also calculated for all scenarios, where level 3 for the 5L was used in calculating the health state values when comparing to a 3L health state with level 2. For example, the size of the difference between the 3L health states 21111 and 31111 was compared to the difference between the 5L health states, 31111 and 51111.

Third, to understand how the overall distributional properties of the index scores obtained using each descriptive system, mean transition values for all 3,125 and 243 predicted health state values were calculated for the 5L and 3L, respectively, and compared. Mean transition values represent the average change in health utility for possible movements from a given health state. This approach has been reported previously to estimate “instrument-defined minimally important differences” or an overall single-level mean transition value (McClure, Al Sayah, Xie, Luo, & Johnson, 2017). A single-level transition was defined as the predicted difference in index score from a given “baseline health state” to an adjacent level for a single dimension (worse or better) without any changes in the other four levels. For example, a transition for Usual Activities in the health state “22222” results in health states 22122 (better) and 22322 (worse). Thus, all possible single-level transitions from 22222 results in health states: 12222, 21222, 22122, 22212, 22221, 32222, 23222, 22322, 22232, 22223. To calculate single mean transition values, the index score for each of these health states is obtained from a value set. Then the absolute difference in index scores between the baseline state and every single-level transition is calculated and averaged to determine a single mean transition value for that health state. All single mean transition values were plotted by baseline 3L or 5L index score to visually assess consistency of transition values across health state severity. For the 5L, this equates to 25,000 possible transitions and 1,620 for the 3L. A slope of the line-of-best fit was calculated by simple (ordinary least squares) linear regression for mean transition values by baseline index score. A calculated slope that is closer to zero indicates greater consistency of mean transition value across baseline health state values. Overall mean transition values were calculated for the 3L and 5L by aggregating all single mean transitions for the baseline health states in both value sets. Two additional analyses using the methods described above were performed to evaluate whether the 5L distributional properties would change if larger transitions were allowed. In the first analysis, all two-level transitions for every 5L health state (i.e., transitions between levels 1 and 2, 2 and 4, and 3 and 5) were calculated. The second analysis restricted the 5L health states to those that "matched" the 3L health state (i.e., states containing levels 1, 3, and 5 only).

All analyses were conducted using SAS 9.4 (Chapel Hill, NC). Because data were collected using the web-based EuroQol Valuation Technology platform (EQ-VT) and interviewer facilitated, no missing data was noted in any of the included interviews; therefore, missing data imputation procedures were not required.

2.4 Results

Of 1134 respondents who completed cTTO tasks for 3L and 5L health states, 72 respondents were excluded due to their inability to comprehend the tasks (per interviewer assessment), forming 1062 respondents in the analytic sample (**Table 3**).

2.4.1 Comparison of model estimates and value set characteristics

In comparing model estimates, all coefficients were negative, logically ordered, and were statistically significantly different from the preceding level in both value sets except Usual Activities (level 3 from level 2) in the 3L model and Self-Care (level 3 from 2 + level 5 from level 4) and (Usual Activities level 5 from level 4) in the 5L model (**Table 4**). The range of scale was for the 3L was 1.583 and 1.357 for the 5L (**Table 4**). The mean (and standard deviation) predicted value for the 3L (0.277 [0.332]) and 5L (0.294 [0.234]) were similar. Dimension ranges were also similar in both value sets, with the exception of the Mobility, where the dimension range for the 3L (-0.525) was approximately twice the size of the 5L (-0.262) coefficient (**Table 4**). Kernel densities for the predicted values for all 3,125 and 243 health state values for the 5L and 3L scoring algorithms are given in **Figure 5**. The model intercept for the 5L had a 95% confidence interval (CI) that overlapped with 1 (0.973; 95%CI 0.934-1.013). The 3L was statistically different from 1 (0.922; 95%CI 0.869-0.967). The mildest health states for the 3L were ranged from 0.820 (11112) and 0.895 (11211). For the 5L, the mildest states ranged from 0.888 (21111) to 0.924 (11112).

2.4.2 Comparison of transitions to mildest and most severe health states

Single-level transitions from 11111 to the mildest 3L ranged from -0.034 to 0.055 for the mildest (unmatched) 5L health states (**Table 5A**). For matched 5L health states, matched differences ranged from -0.066 to -0.016 (**Table 5A**). For single-level health state transitions from the PITS state, the differences between 3L and 5L health states transitions ranged from -0.411 to -0.083 (**Table 5B**). Differences in single-level transition values from PITS between 3L and matched 5L health states were smaller in Self-Care (0.011), Usual Activities (-0.049), Pain/Discomfort (-0.035), and Anxiety/Depression (-0.043), except with Mobility (-0.327) (**Table 5B**).

2.4.3 Comparison of transitions for adjacent corner states

A comparison of single-level transitions between adjacent states for the 3L and 5L value sets found the largest difference in values for the 3L occurred between level 2 and level 3 for all dimensions: Mobility (0.467), Self-Care (0.108), Usual Activities (0.158), Pain/Discomfort (0.280), and Anxiety/Depression (0.218) (**Table 6**). In the 5L value set, the largest change occurs between level 3 and level 4 for Self-Care (0.094), Usual Activities (0.139), Pain/Discomfort (0.182), and Anxiety/Depression (0.163). The change between level 1 and level 2 produced the largest difference for Mobility (0.085). Overall, adjacent health state differences between matched 3L and 5L health were similar (|<0.07|) across all matched dimension-levels, with the exception of the most severe level of Mobility (0.323).

2.4.4 Comparison of all mean transitions values

All single- and two-level mean transitions as a function of baseline 3L and 5L index scores were plotted to illustrate differences in transitions values between value sets (**Figure 6)**. The overall mean (SD) single-level transition value was 0.16 (0.03) and 0.07 (0.01) for the 3L and 5L, respectively. A significant negative relationship was found between single-level mean transition values and baseline index scores (slope = -0.081; standard error [SE] = 0.004; p<0.001) for the 3L, but not for the 5L (slope=-0.001; SE=0.001; p=0.07) (**Figures 6A & 6B**). For two-level transitions calculated for the 5L, the overall mean was 0.15 (0.01). A significant relationship was found between two-level mean transition values and baseline 5L index scores (slope=-0.030; SE=0.001; p<0.001) (**Figure 6C**). Finally, for single-level transitions calculated for the 243 matched 5L health states, the overall mean was 0.13 (0.01) and a significant relationship was found between two-level mean transition values and baseline 5L index scores (slope=-0.020; SE=0.002; p<0.001) (**Figure 6D**).

2.5 Discussion

The present study represents a key addition to the literature by examining the differences in societal value sets obtained from EQ-5D-5L and -3L health state valuations while holding many major sources of variations constant. Preferences were obtained from the same respondents, using the same preference elicitation technique (cTTO), implementing the same study protocol and interviewers, and employing the same statistical modelling approach. The vast majority of recent studies comparing the 5L and 3L have explored the comparative performance of *existing* 3L and 5L value sets, which were generated by studies that differed in many respects.

In comparing 3L and 5L value sets, we found several similarities and differences. In general, parameter coefficients and predicted health state values were comparable, especially when middle categories for dimension severity were matched. Further, differences in changes for single-level transitions between adjacent health states were generally larger with the 3L compared to the 5L; however, most differences were reduced when middle categories (i.e., level 2 in the 3L to level 3 in the 5L) were matched, with the exception of the most severe level of Mobility. The value sets also differed in three major ways. First, we observed a larger ceiling effect with the 3L compared to the 5L. Second, the range of scale for the 3L was larger than the 5L. And third, mean transition values were consistent across 5L health states whereas 3L mean transition values were related to the severity of the baseline health state.

This study found an increased ceiling effect in the 3L value set compared to the 5L. This was evidenced by the observation that the value for 11111 (i.e., the intercept) for the 3L was similar to values for the mildest 5L health states. This finding is conditioned on interpreting the intercept as equal to 11111 or the absence of problems in all dimensions. Others have argued that the intercept (or constant term) may be interpreted as the disutility associated with any transition away from full health and that 11111 is equivalent to “full or perfect health” and therefore can be assigned a value of 1 (Dolan, 1997; Shah, Mulhern, Longworth, & Janssen, 2016). However, this is problematic in several respects. First, the assertion that the intercept represents an additional disutility applied for any suboptimal health state is not grounded in any statistical basis. Second, the finding by Shah et al. that values obtained using “Full Health” or “11111” as the anchor did not differ may be explained by a framing effect, whereby respondents were biased to view “Full Health” in the exercise as “no problems” in all dimensions rather than viewing it as perfect health. Others have noted systematic differences in defining a utility value of 1 as “no problems” in all dimensions or perfect health (King Jr et al., 2003). To our knowledge, this is the first study that empirically quantifies the difference in ceiling effects between the 3L and 5L in the context of health state valuations.

Our observation of an increased range of scale for the 3L was likely driven by a large difference in the size of the coefficients for the most severe level of problems in Mobility between the value sets, given that all other dimension weights were similar. Indeed, the difference between the range of scales for the 3L and 5L approximates the size the difference in parameter estimates. In developing the EQ-5D-5L, the 5L replaced the 3L’s level 3 descriptor - ‘confined to bed’ - with “unable to walk about” (M Herdman et al., 2011). Therefore, the increased range may only be relevant to respondents able to endorse “confined to bed”. However, in practice, the rarity of respondents reporting Mobility level 3 has been noted (N. Devlin et al., 2018). For instance, in a large (n >3900) multinational cohort of patients with chronic conditions, only 1 out of the 8 patient groups (stroke) reported being “confined to bed” with a prevalence greater than 1% (M. Janssen et al., 2013). Similarly, no respondents in the present analysis reported being “confined to bed”. Thus, for many applications, the 3L Mobility dimension only has two levels: “no” and “some problems” which may lead to an underestimation of benefit (i.e., QALY gains) for interventions that improve severe problems with mobility in many situations (M. F. Janssen et al., 2018). However, in other settings, “confined to bed” is an outcome that is reasonably expected to occur (e.g., stroke) and the 3L may better capture utility gains/losses associated with these specific changes in Mobility. Therefore, in the broadest sense, the appropriateness of the selection of the 3L or 5L may depend on the application, but for most contexts where bed confinement does not occur, the 5L is the better measure.

We found that the 5L value set produced consistent mean transition values across all predicted index scores whereas the 3L exhibited larger gains/losses with more severe (lower) index scores has several implications. However, gains/losses are similar when comparing single-level transitions for the 3L with two-level or matched 5L transitions. Further, the consistency in 5L transition values was relatively maintained upon visual inspection for the two-level and matched 5L transitions. These results suggest that while the 3L is demonstrates larger gains/losses for single-level transitions, the 5L is also capable of producing similar values for matched transitions. Moreover, the 5L transitions are consistent and appear to exhibit interval-scale properties, which is required to fulfill a key assumption of the QALY framework: any gain or loss occurring on one location of the health utility scale is the same on another location of the scale (Drummond et al., 2015; Weinstein et al., 2009). This assumption is reflected in US guideline recommendations for selecting preference-based weights (ICER, 2017; Sanders et al., 2016).

We are aware of only one other head-to-head study that compared 3L and 5L health state values (Selivanova, Buskens, & Krabbe, 2018). Preferences for 3L and 5L health states were collected from an online panel of respondents among the Dutch general population who were randomly assigned to value either 3L or 5L health states using the discrete choice tasks. Similar to the present study, the authors observed a strong negative preference for “confined to bed” in the Mobility dimension with the 3L. However, Selivanova and colleagues observed only modest differences in the dimension ranges for Mobility between the 3L and 5L (-1.550 vs. -1.267) whereas we observed a 3L range that was almost twice that of the 5L. Further, there were several differences that limit the interpretability of their results with respect to potential QALY differences. First, health state values obtained by Selivanova et al. were on a latent utility scale (i.e., not anchored between Dead and Full health) while our study obtained values on the QALY scale using the TTO method. Second, their analysis compared model estimates and overall range and distribution of health states values between descriptive systems but did not report analyses comparing changes in utility between adjacent health states, as we were able to show. Finally, the DC study randomized respondents to valuing 3L or 5L health states, but the study design does not account for inter-subject variability by having the same respondents valuing health states from both version of the EQ-5D. While randomization should have theoretically accounted for differing respondent characteristics, the authors noted an imbalance in the study arms with respect to age.

The present study should be viewed in the context of several limitations. First, the 5L health state blocks randomly assigned to each respondent were balanced with respect to health state severity. In contrast, each of the 3L blocks were not necessarily so. However, the value sets represent average health state values taken from observations from all respondents and any block imbalance is unlikely to impact the model estimates. Second, each respondent valued few 3L health states, resulting in small sample of observations from which the 3L value set was based. Nonetheless, standard errors remained small (<0.03) and the resulting model contained logically ordered and statistically significant coefficients. Third, each respondent valued 5L health states first before valuing the 3L health states rather than randomizing health state orders. This may have resulted in systematic differences in how 5L and 3L health states were valued. Augestad et al. showed that health states valued later in a series of TTO tasks tended to provide more extreme values (i.e., more positive and negative BTD and WTD values, respectively) (Liv Ariane Augestad, Rand-Hendriksen, Kristiansen, & Stavem, 2012). However, if a similar learning effect was observed in our study, one would expect to systematically larger 3L model estimates relative to the 5L estimates. Instead, we observed small differences between matched coefficients (with the exception of Mobility) which indicates a specific response to specific descriptors rather than a broader framing effect. Finally, interviewers were not instructed to explicitly orient respondents to the 3L descriptive system prior to valuing 3L health states. However, explicitly introducing the 3L descriptive system after the 5L may have produced artificial similarities in how respondents interpreted and valued 3L health states. Respondents would likely have focused on how the 3L was intended to “fit” into the 5L rather than interpreting and valuing states on the basis of the health descriptors (i.e., labels)., which was the primary goal of this study.

2.6 <u>Conclusion</u>

The results of this study help to understand the underlying elements that serve as sources of differences observed between the 3L and 5L value sets. In addition to comparing model estimates, this study provides a novel comparison by way of calculating and plotting mean transition values for all 3L and 5L predicted health states. With these, we were able to provide insight into how individuals differentially respond to stimuli (i.e., labels of the descriptive system) when valuing health and how that impacts the distributional properties of the index scores. We found a larger range of scale for the 3L compared to 5L, driven largely by the difference in weight placed on the most severe level of problems in Mobility. The 5L produced a value set with better measurement properties and a reduced ceiling effect. For most applications, the 5L appears to be a superior descriptive system than the 3L. Future research should assess the extent to which choice of value set may directly impact the results of cost-effectiveness studies and subsequent health care decision-making.

Acknowledgements

We are grateful to the following individuals for their guidance and assistance in the study preparation, analysis, and/or manuscript preparation: Elly Stolk (EuroQol Research Foundation); Zhihao Yang (Erasmus University); Mark Oppe (EuroQol Research Foundation); Ruixuan Jiang (University of Illinois at Chicago).

TABLES

Table 1 Observed summary statistics for the 86 EQ-5D-5L health states

Health state	n	Mean	SD	Median	25th percentile	75th percentile	Health state	n	Mean	SD	Median	25th percentile	75th percentile
11112	210	0.940	0.214	1.000	0.950	1.000	31524	106	0.313	0.643	0.500	0.050	0.800
11121	221	0.964	0.095	1.000	0.950	1.000	31525	108	0.204	0.677	0.500	-0.475	0.700
11122	116	0.891	0.249	1.000	0.900	1.000	32314	104	0.358	0.683	0.600	0.100	0.875
11211	202	0.943	0.130	1.000	0.950	1.000	32443	103	0.132	0.667	0.300	-0.500	0.600
11212	98	0.898	0.258	1.000	0.900	1.000	33253	104	0.227	0.636	0.425	0.000	0.625
11221	107	0.910	0.223	1.000	0.900	1.000	34155	103	0.021	0.672	0.200	-0.700	0.500
11235	107	0.498	0.594	0.700	0.300	0.900	34232	116	0.390	0.628	0.600	0.250	0.825
11414	117	0.393	0.616	0.600	0.200	0.800	34244	98	-0.001	0.681	0.100	-0.700	0.500
11421	108	0.678	0.490	0.800	0.600	1.000	34515	107	0.185	0.658	0.400	-0.050	0.700
11425	99	0.373	0.637	0.500	0.100	0.900	35143	117	0.242	0.646	0.500	0.000	0.700
12111	223	0.933	0.181	1.000	0.950	1.000	35245	107	0.100	0.670	0.300	-0.500	0.600
12112	98	0.892	0.307	1.000	0.900	1.000	35311	116	0.488	0.619	0.700	0.400	0.900
12121	103	0.883	0.243	0.950	0.900	1.000	35332	99	0.494	0.551	0.600	0.300	0.950
12244	108	0.220	0.667	0.425	-0.100	0.700	42115	99	0.395	0.627	0.600	0.200	0.900
12334	104	0.428	0.620	0.600	0.275	0.900	42321	116	0.557	0.576	0.775	0.500	0.950
12344	104	0.260	0.669	0.500	0.000	0.700	43315	106	0.250	0.665	0.500	0.000	0.700
12513	104	0.605	0.496	0.700	0.500	0.950	43514	98	0.170	0.687	0.300	-0.100	0.700
12514	107	0.426	0.580	0.600	0.300	0.800	43542	103	0.090	0.673	0.300	-0.500	0.600
12543	103	0.191	0.656	0.400	-0.500	0.700	43555	116	-0.142	0.655	0.000	-0.900	0.400
13122	99	0.822	0.372	0.950	0.800	1.000	44125	104	0.197	0.660	0.400	-0.250	0.700
13224	116	0.424	0.625	0.625	0.325	0.850	44345	104	0.011	0.666	0.200	-0.550	0.500
13313	108	0.717	0.403	0.800	0.600	1.000	44553	98	-0.194	0.652	0.000	-1.000	0.300
14113	106	0.563	0.548	0.800	0.500	0.900	45133	103	0.303	0.614	0.500	0.150	0.700
14554	104	0.033	0.679	0.200	-0.650	0.575	45144	107	-0.003	0.703	0.100	-0.700	0.600
15151	106	0.214	0.699	0.500	-0.450	0.800	45233	108	0.304	0.682	0.500	0.000	0.775

21111	206	0.922	0.268	1.000	0.950	1.000	45413	99	0.191	0.714	0.400	-0.500	0.800
21112	104	0.850	0.358	1.000	0.875	1.000	51152	99	0.270	0.674	0.450	0.100	0.800
21315	106	0.485	0.556	0.700	0.400	0.900	51451	107	0.166	0.691	0.400	-0.400	0.700
21334	104	0.407	0.648	0.625	0.200	0.875	52215	103	0.216	0.678	0.400	-0.300	0.800
21345	98	0.063	0.728	0.300	-0.800	0.700	52335	116	0.164	0.638	0.400	0.000	0.600
21444	117	0.124	0.660	0.300	-0.500	0.700	52431	106	0.221	0.677	0.500	0.000	0.700
22434	99	0.323	0.647	0.500	0.100	0.800	52455	108	-0.159	0.703	0.000	-1.000	0.400
23152	98	0.217	0.745	0.500	-0.500	0.850	53221	104	0.506	0.622	0.700	0.450	0.925
23242	104	0.376	0.628	0.550	0.200	0.800	53243	117	0.158	0.669	0.400	0.000	0.700
23514	103	0.255	0.640	0.400	0.000	0.800	53244	117	0.026	0.692	0.100	-0.700	0.500
24342	104	0.197	0.666	0.425	0.000	0.700	53412	104	0.281	0.672	0.500	0.000	0.800
24443	106	0.061	0.670	0.250	-0.550	0.500	54153	106	0.013	0.692	0.250	-0.700	0.500
24445	116	-0.112	0.630	0.000	-0.750	0.350	54231	107	0.356	0.682	0.600	0.200	0.850
24553	99	0.120	0.693	0.300	-0.500	0.600	54342	104	0.132	0.677	0.325	-0.500	0.650
25122	108	0.596	0.521	0.700	0.500	1.000	55225	104	0.014	0.679	0.200	-0.650	0.500
25222	117	0.488	0.611	0.700	0.400	0.900	55233	108	0.213	0.689	0.450	-0.400	0.700
25331	117	0.433	0.597	0.600	0.300	0.850	55424	98	-0.114	0.689	0.000	-0.950	0.400
31514	117	0.320	0.645	0.500	0.000	0.800	55555	1062	-0.343	0.637	-0.500	-1.000	0.100

SD=standard deviation

Table 2 Summary of observed composite time trade-off values obtained for 30 EQ-5D-3L health states

Health state	n	Mean	SD	Median	25th percentile	75th percentile	Health state	n	Mean	SD	Median	25th percentile	75th percentile
11112	106	0.794	0.453	1.000	0.800	1.000	21323	110	0.240	0.670	0.500	0.000	0.700
11113	106	0.513	0.677	0.800	0.300	1.000	21332	112	0.139	0.704	0.400	-0.600	0.700
11121	111	0.910	0.157	1.000	0.850	1.000	22121	105	0.736	0.357	0.800	0.700	1.000
11122	99	0.805	0.355	0.900	0.800	1.000	22222	111	0.612	0.491	0.750	0.500	0.950
11211	115	0.919	0.281	1.000	0.950	1.000	22233	100	0.078	0.744	0.300	-1.000	0.700
11313	115	0.432	0.648	0.600	0.300	0.900	23112	112	0.498	0.634	0.700	0.500	0.950
12111	100	0.928	0.132	1.000	0.900	1.000	23323	100	0.107	0.702	0.300	-0.550	0.625
12212	100	0.719	0.421	0.850	0.600	1.000	31131	115	-0.036	0.732	0.100	-0.900	0.600
12222	110	0.672	0.436	0.800	0.500	0.950	31223	104	-0.063	0.671	0.100	-0.875	0.500
12331	105	0.182	0.679	0.400	-0.300	0.700	32113	99	-0.089	0.715	0.050	-1.000	0.500
13133	100	0.070	0.732	0.300	-0.800	0.600	32232	110	-0.150	0.662	0.000	-1.000	0.400
13221	105	0.488	0.596	0.700	0.300	0.900	32322	100	-0.023	0.719	0.125	-1.000	0.500
21111	112	0.919	0.189	1.000	0.950	1.000	33232	99	-0.254	0.659	0.000	-1.000	0.300
21133	106	0.140	0.742	0.400	-0.900	0.700	33311	104	-0.048	0.693	0.075	-0.900	0.500
21211	104	0.839	0.343	0.950	0.800	1.000	33333	111	-0.423	0.631	-0.600	-1.000	0.050

SD=standard deviation

Table 3 Respondent characteristics

Characteristic	Analytic sample (n=1062)
Age, mean (SD),	46.5 (18.0)
18-34, n (%)	347 (32.7)
35-54, n (%)	365 (34.4)
55+, n (%)	350 (33.0)
Range	18-99
Gender, n (%)	
Male	515 (48.5)
Female	542 (51.0)
Other	5 (0.5)
Race, n (%)	
White	661 (62.2)
Black	128 (12.1)
Other	174 (25.7)
Hispanic, n (%)	191 (18.0)
Education, n (%)	
Secondary or less	354 (34.0)
History of illness, n (%)	
Hypertension	244 (23.0)
Stroke	20 (1.9)
Depression	270 (25.4)
Arthritis	244 (23.0)
Asthma	125 (11.8)
Diabetes	95 (9.0)
Cancer	59 (5.6)
Bronchitis	23 (2.2)
None	356 (33.5)
EQ-VAS	
Mean (SD)	80.5 (15.5)
Median (IQR)	85 (15)
Health status, n (%)	
Excellent	219 (20.6)
Very good	392 (36.9)
Good	312 (29.4)
Fair	114 (10.7)
Poor	24 (2.3)

IQR=interquartile range; SD=standard deviation; VAS=visual analogue scale

Table 4 Comparison of EQ-5D-3L and -5L model parameter estimates and value set characteristics

Parameter		3L β	3L SE	3L p-value	5L β	5L SE	5L p-value
	Intercept	**0.922**	**0.025**	**<0.001**	**0.973**	**0.020**	**<0.001**
Mobility	Slight				**-0.085**	**0.014**	**<0.001**
	Some/Moderate	**-0.058**	**0.022**	**0.009**	**-0.122**	**0.014**	**0.012**
	Severe				**-0.206**	**0.016**	**<0.001**
	Confined to bed/Unable to	**-0.525**	**0.025**	**<0.001**	**-0.262**	**0.014**	**<0.001**
Self-care	Slight				**-0.069**	**0.014**	**<0.001**
	Some/Moderate	**-0.080**	**0.026**	**0.002**	**-0.102**	**0.015**	**0.0425**
	Severe				**-0.197**	**0.015**	**<0.001**
	Unable to	**-0.188**	**0.025**	**<0.001**	-0.221	0.014	0.0837
Usual Activities	Slight				**-0.061**	**0.014**	**0.0345**
	Some/Moderate	-0.027	0.024	0.257	**-0.093**	**0.015**	**<0.001**
	Severe				**-0.232**	**0.015**	**<0.001**
	Unable to	**-0.187**	**0.023**	**<0.001**	-0.204	0.014	0.0719
Pain/ Discomfort	Slight				**-0.062**	**0.013**	**0.0151**
	Moderate	**-0.081**	**0.024**	**0.001**	**-0.103**	**0.015**	**<0.001**
	Severe				**-0.285**	**0.014**	**<0.001**
	Extreme	**-0.361**	**0.023**	**<0.001**	**-0.347**	**0.015**	**<0.001**
Anxiety/ Depression	Slight				**-0.049**	**0.015**	**<0.001**
	Moderate	**-0.104**	**0.026**	**<0.001**	**-0.120**	**0.017**	**0.001**
	Severe				**-0.282**	**0.015**	**<0.001**
	Extreme	**-0.322**	**0.024**	**<0.001**	-0.295	0.014	0.375
Akaike Information Criterion				5073			13591
Bayes Information Criterion				5077			13601
# “illogical”				0			1
# non-significant (p <0.05)				1			3
Mean absolute error (predicted vs. observed)				0.055			0.036
Health states |>0.05|				13 of 30 (43%)			11 of 86 (37%)
Health states |>0.10|				3 of 30 (10%)			3 of 86 (10%)
Pearson’s r (predicted vs. observed)				0.984			0.981
Range (11111 to PITS)				0.922 to -0.662 (1.583)			0.973 to -0.356 (1.357)
Mean (SD)				0.277 (0.332)			0.29 (0.23)
Median (IQR)				0.30 (0.51)			0.30 (0.33)
Dimension importance				MO-PD-AD-SC-UA			PD-AD-MO-UA-SC
Number (%) of health states worse than dead				53 out of 243 (21.8%)			352 out of 3,125 (11.3%)
Moderate state (22222 / 33333)				0.570			0.433
PITS state (33333 / 55555)				-0.662			-0.356

IQR=interquartile range; SD=standard deviation; SE=standard error

Table 5 Comparison of single-level transitions from health state "11111" and the PITS state for each EQ-5D dimension

A) Mildest health states

$3L_{state}$	3L value	$11111\text{-}3L_{value}$ [A]	$5L_{state}$	5L value	$11111\text{-}5L_{value}$ [B]	[A] - [B]	Comparable $5L_{state}$	5L value	$11111\text{-}5L_{value}$ [C]	[A] - [C]
11111	0.922	--	**11111**	0.973	--	--	**11111**	0.973	--	--
11112	0.818	0.104	**11112**	0.924	0.049	0.055	**11113**	0.853	0.120	-0.016
11121	0.840	0.081	**11121**	0.911	0.062	0.019	**11131**	0.870	0.103	-0.022
11211	0.894	0.027	**11211**	0.912	0.062	-0.034	**11311**	0.880	0.094	-0.066
12111	0.842	0.080	**12111**	0.904	0.070	0.011	**13111**	0.871	0.103	-0.023
21111	0.864	0.058	**21111**	0.888	0.085	-0.027	**31111**	0.851	0.123	-0.065

B) Most severe health states

$3L_{state}$	3L value	$33333\text{-}3L_{value}$ [A]	$5L_{state}$	5L value	$55555\text{-}5L_{value}$ [B]	[A] - [B]	Comparable $5L_{state}$	5L value	$55555\text{-}5L_{value}$ [C]	[A] - [C]
33333	-0.662	--	**55555**	-0.356	--	--	**55555**	-0.356	--	--
33332	-0.444	-0.218	**55554**	-0.344	-0.013	-0.205	**55553**	-0.181	-0.175	-0.043
33323	-0.382	-0.280	**55545**	-0.293	-0.063	-0.217	**55535**	-0.112	-0.244	-0.035
33233	-0.502	-0.159	**55455**	-0.385	0.029	-0.188	**55355**	-0.246	-0.110	-0.049
32333	-0.554	-0.108	**54555**	-0.331	-0.025	-0.083	**53555**	-0.237	-0.119	0.011
23333	-0.194	-0.467	**45555**	-0.300	-0.056	-0.411	**35555**	-0.216	-0.140	-0.327

Table 6 Comparison of adjacent health states for changes within each EQ-5D dimension

Dimension	3L health state	Value	∆ from less severe adjacent 3L state [A]	5L health state	Value	∆ from less severe adjacent 5L state	Matched difference[a][B]	[A] - [B]
Mobility	11111	0.922		11111	0.973		--	
				21111	0.888	0.085	--	
	21111	0.864	0.058	31111	0.851	0.038	0.123	-0.065
				41111	0.767	0.084		
	31111	0.397	0.467	51111	0.711	0.056	0.140	0.327
Self-care	11111	0.922		11111	0.973			
				12111	0.904	0.070		
	12111	0.842	0.080	13111	0.871	0.033	0.103	-0.023
				14111	0.777	0.094		
	13111	0.734	0.108	15111	0.752	0.025	0.119	-0.011
Usual Activities	11111	0.922		11111	0.973			
				11211	0.912	0.062		
	11211	0.894	0.027	11311	0.880	0.032	0.094	-0.066
				11411	0.741	0.139		
	11311	0.736	0.158	11511	0.769	-0.029	0.110	0.048
Pain/ Discomfort	11111	0.922		11111	0.973			
				11121	0.911	0.062		
	11121	0.840	0.081	11131	0.870	0.041	0.103	-0.022
				11141	0.689	0.182		
	11131	0.561	0.280	11151	0.626	0.063	0.244	0.035
Anxiety/ Depression	11111	0.922		11111	0.973			
				11112	0.924	0.049		
	11112	0.818	0.104	11113	0.853	0.071	0.120	-0.016
				11114	0.691	0.163		
	11113	0.600	0.218	11115	0.678	0.013	0.175	0.043

[a]Matched difference is the calculated difference between changes from comparable 3L and 5L health states (e.g., 21111 on the 3L is comparable to 31111 on the 5L)

Figure 5 Kernel densities for all predicted 3L and 5L health state values

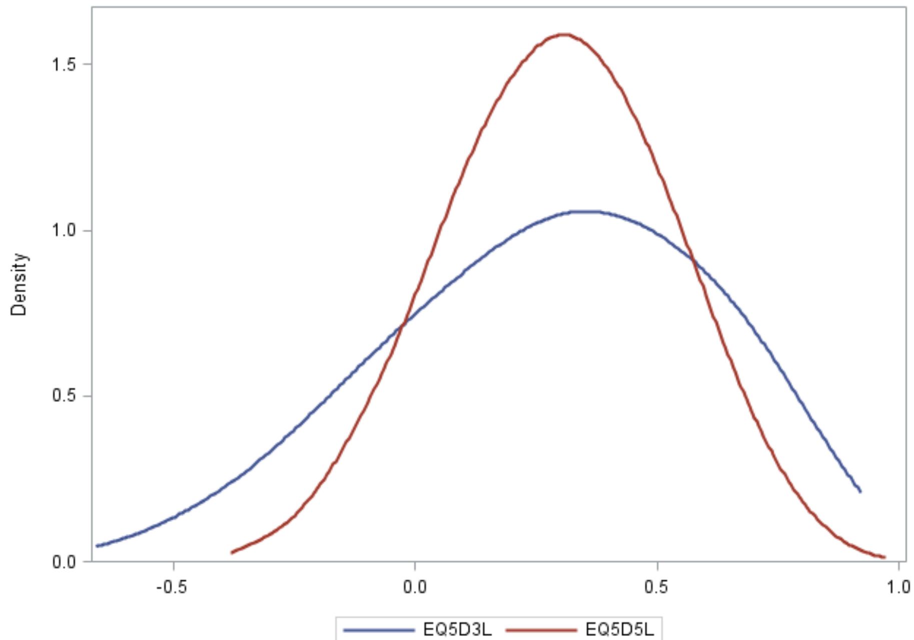
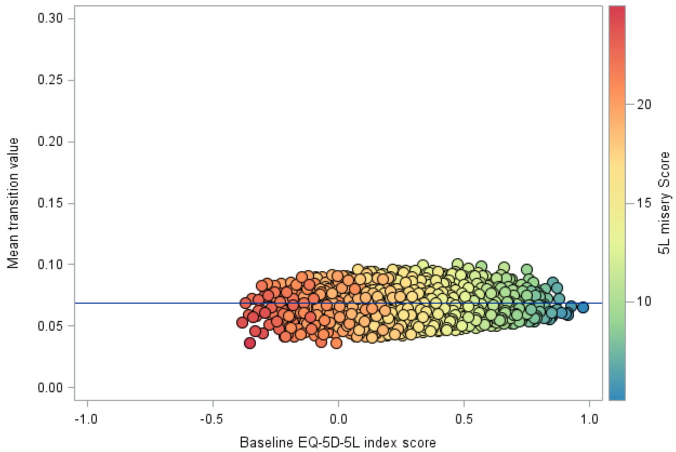
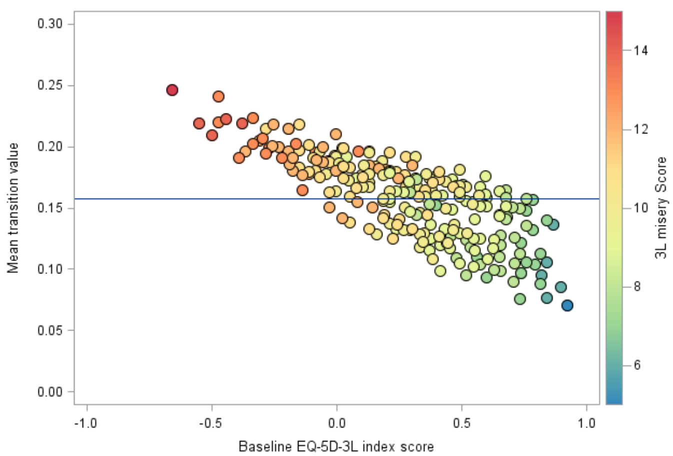


Figure 6 Mean transition values for all EQ-5D-3L and EQ-5D-5L index scores

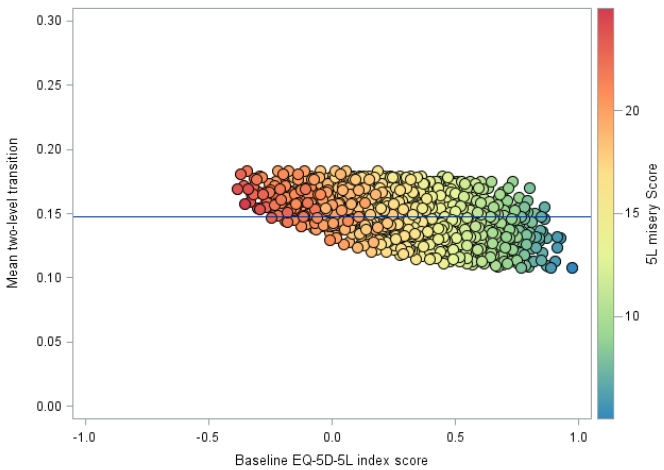
A) Single-level transitions for all EQ-5D-5L health states



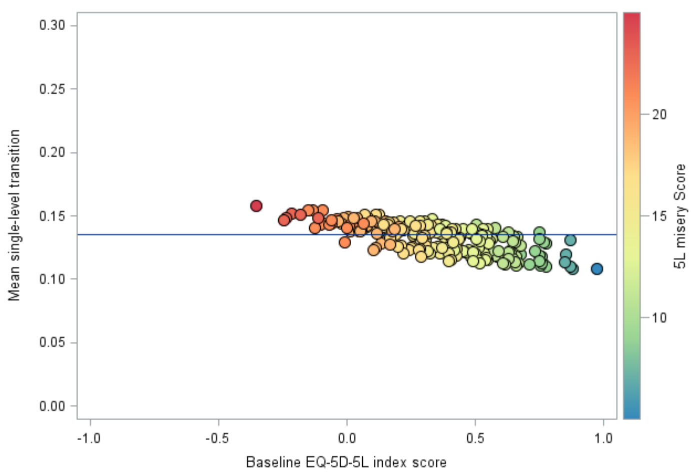
B) Single-level transitions for all EQ-5D-3L health states



C) Two-level transitions for all EQ-5D-5L health states



D) Single-level transitions for matched[a] EQ-5D-5L health states



Note: color gradient represents misery score for each EQ-5D health state vector, where blue indicates lower misery score and red indicates highest

[a]EQ-5D-5L health states (n=243) containing levels 1, 3 and 5 only

3. ADVANCE DIRECTIVES AND HEALTH VALUATION IN THE UNITED STATES

Authors: Ernest H. Law, PharmD; A. Simon Pickard, PhD; Surrey M. Walton, PhD; Todd A. Lee, PharmD, PhD; Feng Xie, PhD; Alan Schwartz, PhD (**s*ubmitted to Medical Decision Making)***

3.1 Abstract

Objectives: To better understand the implications of informed preferences, this study aimed to examine the association between advance directives for health care and health preferences.
Methods: Data from the 2017 US EQ-5D-5L valuation study was analyzed. Using advance directive status and values obtained from 10 time trade-off (TTO) exercises, regression models fitted health values to estimate the impact of advance directives, adjusting for respondent characteristics. Logistic regression models examined the association between advance directive status and likelihood of generally valuing health states as worse-than-dead (WTD) and at least one of the 10 health states as WTD during the TTO exercises.

Results: Of 1061 respondents, 27.4% had an advance directive. Advance directives were associated with significantly lower mean values assigned to overall health states (difference=-0.101; 95%CI -0.175 to -0.028) and specific dimension-levels: Mobility levels 4 (difference =-0.146) and 5 (difference=-0.138), Usual Activities level 2 (difference=-0.141), Pain/Discomfort level 4 (difference=-0.138), and Anxiety/Depression level 4 (difference=-0.158; all p-values <0.05). Advance directive status was not associated with an increased likelihood of valuing health states as WTD (OR=1.62; 95%CI 0.99-2.60; p=0.053) or at least one of the 10 health states as WTD (OR=1.29; 95%CI 0.91-1.84; p=0.146).

Conclusions: Advance directives were associated with a greater willingness to trade years of life for better health. If advance directives are an indicator of a substantial proportion of the population with “informed” preferences, it has implications not only for intensive care decision making at the individual level, but also resource allocation based on cost-utility analysis at the group level.

3.2 Introduction

In striving to maximize health in the context of limited resources, decision makers in health systems often utilize economic evaluations to inform decisions involving the adoption of health technologies (P. J. Neumann et al., 2015). Performing economic evaluations requires quantifying the costs and benefits of adopting an intervention and comparing it to the costs and benefits of an appropriate alternative. To this end, generic preference-based instruments facilitate the valuation of health outcomes, providing an index-based score used in calculating quality-adjusted life years (QALYs). These measures typically include a multi-attribute descriptive system and scoring algorithms (or societal value sets), estimated using preferences obtained from the general public.

The use of societal value sets in informing cost-utility analyses (CUA) has been endorsed by several health technology assessment (HTA) agencies around the world (CADTH, 2017; NICE, 2013; PBAC, 2016) , including the United States (US) despite the stance of the US federal government (Gold, 1996; ICER, 2017; Sanders et al., 2016). The Second Panel of Cost-Effectiveness in Health and Medicine continued to endorse “the view that the best articulation of a society’s preferences for particular health states would be gathered from a representative sample of *fully informed* members of the community” (Peter J Neumann et al., 2016). However, it is not clear from the current literature what defines *informedness*, and therefore which subgroups within the general population are most apt to provide informed preferences.

Individuals who have completed an advance directive constitute a group within the general population who are likely to have reflected substantially on the experience and consequences of a range of health states, and therefore, may have constructed informed preferences for health and care. Advance directives are meant to increase the autonomy of patients and communicate their wishes when unable to do so themselves. The process of completing an advance directive involves discussion and documentation of an individual's preferences concerning their goals of care in the event they lose capacity or communication ability (Emanuel, 1995). In the US, all 50 states and the District of Columbia regulate their advance directive legalities individually, but the process typically involves completion of a living will and durable power of attorney.

This study aimed to characterize individuals with and without advance directives and their preferences for health states of varying severity. We hypothesized that individuals who completed an advance directive would be more likely to value health states as WTD, as well as trade more time for better health resulting in lower mean health values, because they have reflected on health, and provided their preference about health care decision making issues surrounding poor health. We also aimed to illustrate the potential impact on the estimation of QALYs by calculating and comparing transition values for all estimated EQ-5D-5L health states.

3.3 Methods

3.3.1 Data source

Data from the US EQ-5D-5L health state valuation study was analyzed (A. Pickard et al., 2018). Briefly, quota sampling was used to obtain a sample of the US general population who were similar in terms of age, sex, race, and ethnicity. English- and Spanish-speaking participants were recruited from the four US census regions (Midwest, Northeast, South, and West) in 6 greater metropolitan areas of Chicago (Illinois), Philadelphia (Pennsylvania), Birmingham (Alabama), Phoenix (Arizona), Denver (Colorado), and Seattle (Washington). The dataset included 1134 respondents. Respondents were excluded from this analysis if they were not able to comprehend the preference elicitation tasks per interviewer assessment.

3.3.2 Advance directive status

The independent variable of interest was the completion of an advance directive as reported by the respondent. Respondents were asked about their advance directive status near the end of the study interview ("*A living will or an advance directive is a legal form that allows you to document your wishes for medical treatments at the end of life. Please choose one of the options below*") with the following response options: a) "*I have completed an advance directive*"; b) "*I have not completed an advance directive, but plan to*"; c) "*I have not completed an advance directive and do not plan to*"; and d) "*I have never thought about it*". Based on these responses, respondents were categorized as either having completed an advance directive (response a) or not (responses b, c, or d).

3.3.3 Health state descriptions

Health states were described using the EQ-5D-5L descriptive system. In order of increasing severity, the five levels for each dimension are "no problems," "slight problems," "moderate problems," "severe problems"; the most severe level is labelled "unable to" for the functional dimensions (Mobility, Self-Care, Usual Activities) and "extreme problems" for the affective dimensions (Pain/Discomfort, Anxiety/Depression) (M. Herdman et al., 2011). A dimension in which there are “no problems” is designated level 1, while a dimension where there are “extreme problems” is assigned level 5. Each unique health state described by the 5L system is represented by a five-digit descriptor that ranges from 11111 (absence of problems) to 55555 (worst possible health state), with each digit representing one dimension-level of health, resulting in a descriptive system that defines 3,125 (5^5) health states. A “misery score” can also be calculated as a proxy indicator for health state severity by summing the five digits of the health state descriptor. For example, the misery score for “51342” is equal to 15. The highest possible misery score for the EQ-5D-5L is 25 (for health state “55555”).

3.3.4 <u>Time trade-off</u>

Health states were described using the EQ-5D-5L descriptive system. In order of increasing severity, the five levels for each dimension are “no problems,” “slight problems,” “moderate problems,” “severe problems”; the most severe level is labelled ”unable to” for the functional dimensions (Mobility, Self-Care, Usual Activities) and “extreme problems” for the affective dimensions (Pain/Discomfort, Anxiety/Depression) (M. Herdman et al., 2011). A dimension in which there are “no problems” is designated level 1, while a dimension where there are “extreme problems” is assigned level 5. Each unique health state described by the 5L system is represented by a five-digit descriptor that ranges from 11111 (absence of problems) to 55555 (worst possible health state), with each digit representing one dimension-level of health, resulting in a descriptive system that defines 3,125 (5^5) health states. A “misery score” can also be calculated as a proxy indicator for health state severity by summing the five digits of the health state descriptor. For example, the misery score for “51342” is equal to 15. The highest possible misery score for the EQ-5D-5L is 25 (for health state “55555”).

3.3.5 Statistical analysis

3.3.5.1 Factors associated with advance directive completion

A multivariable logistic regression was performed to identify the factors associated with completing an advance directive and information collected on respondent characteristics: age (18-35, 36-54, and ≥55 years old), sex (female or not), race (White, Black, or Other), ethnicity (Hispanic or not), education level (received higher education [more than high school] or not), marital status (married or not), any children <18 years old (yes or no), insurance coverage (Medicare/Medicaid, private, or none), financial difficulty (any reported difficulty paying monthly bills or not), experience with serious illness (personal, family, or caregiver [yes or no for each]), favorable view of passive euthanasia (yes or no), favorable view of active euthanasia (yes or no), religiosity (informs approach to life or not), self-reported presence of chronic illness (any vs. none), and self-rated health (EQ-VAS).

3.3.5.2 Mean differences in health values by advance directive status

Independent two-sample t-tests were performed to compare mean health values for the 86 common health states by advance directive status. Due to the large number of comparisons, a Bonferroni correction for multiple testing was applied to reduce type I error (significance level = 0.05 after applying a Bonferroni correction).

Mean differences in health preferences between advance directive status was further investigated using three random effects linear regression models with random intercept between two statistical approaches. A random intercept at the respondent level was specified to account for respondents valuing multiple health states from several randomly assigned blocks. The first approach assessed mean differences in health values between those with and without advance directives by fitting the cTTO values for a given EQ-5D-5L health state using random effects linear regression models. Crude (*Model 1*) and adjusted differences (*Model 2*) in mean cTTO values between respondents with advance directives and those without were estimated with 95% confidence intervals and p-values (alpha = 0.05) while accounting for EQ-5D-5L health state descriptors. *Model 2* further adjusted for covariates of cTTO values that were available within the dataset. *Model 3* included additional interaction terms for a given health state descriptor and advance directive status (i.e., advance*[MO/SC/UA/PD/AD]). The purpose of this model was to assess whether differences in health preferences between respondents with and without advance directives varied across dimension-levels, while adjusting for respondent characteristics. Mean differences for the 20 dimension-levels were then estimated with a significance level of 0.05 after applying Bonferroni correction.

3.3.5.3 <u>Association between advance directive status and valuing health states WTD</u>

To evaluate whether differences in health preference were associated with a differential likelihood of valuing health states as WTD by advance directive status, two sets of logistic regression models were developed. The first set of logistic regression models focused on the crude (*Model 4*) and adjusted (*Model 5*) likelihood of respondents with and without advance directive valuing at least one of the 10 EQ-5D-5L health states in an assigned block as WTD (i.e., < 0). Both models included covariates for EQ-5D-5L health state descriptors as well as health state block assignment. The second set of logistic regression models, which estimated the crude (*Model 6*) and adjusted (*Model 7*) likelihood of valuing a given EQ-5D-5L health state as WTD, included a random intercept to account for multiple observations per respondent and included EQ-5D-5L descriptors as covariates to adjust for health state dimension-level severity. All models estimated odds ratios (OR) with 95% CI and p-values (alpha = 0.05).

Adjusted analyses included the following covariates: age (18-35, 36-54, and ≥55 years old), sex (female or not), race (White, Black, or Other), ethnicity (Hispanic or not), education level (received higher education [more than high school] or not), marital status (married or not), any children <18 years old (yes or no), insurance coverage (Medicare/Medicaid, private, or none), financial difficulty (any reported difficulty paying monthly bills or not), experience with serious illness (personal, family, or caregiver [yes or no for each]), favorable view of passive euthanasia (yes or no), favorable view of active euthanasia (yes or no), religiosity (informs approach to life or not), self-reported presence of chronic illness (any vs. none), and self-rated health (EQ-VAS).

3.3.5.4 Mean transition values for all possible health state values

To understand the potential impact of eliciting health values from individuals with and without advance directives on QALY gains/losses, mean transition values for all 3,125 EQ-5D-5L health states were calculated for each group. Mean transition values are defined as the mean change in health value for all possible transitions from a baseline health state (also referred to as "instrument-defined minimally important differences" (McClure et al., 2017). A movement from a baseline health state to an adjacent health state in which only a single level for one dimension changes (for worse or better) without any changes in the other four levels is defined as a single-level transition value. For example, a transition in the Anxiety/Depression dimension for health state "44444" results in health states 44443 (better) and 44445 (worse). Therefore, all possible single-level transitions from 44444 results in health states: 34444, 43444, 44344, 44434, 44443, 54444, 45444, 44544, 44454, 44445.

To calculate single mean transition values, index scores for both groups with and without advance directives were estimated using two separate random linear regression models in which each model was estimated using samples restricted to individuals who completed an advance directive and those without. Each model was specified similar to *Model 2* except without the variable for advance directive status. Next, the absolute difference in index scores between the baseline health state and all single-level transitions was calculated and averaged to determine a single mean transition value (and standard deviation) for that health state. The difference between overall mean transition values for individuals with and without advance directives was evaluated using an independent t-test. There are 25,000 possible transition values for the EQ-5D-5L descriptive system. All single mean transition values were then plotted by baseline EQ-5D-5L index score to visually assess consistency of transition values across health state severity. Analyses were conducted using, SAS 9.4 (Chapel Hill, NC).

3.4 Results

3.4.1.1 Factors associated with advance directive completion

After excluding respondents who did not comprehend the cTTO task (n=72) or were missing advance directive status (n=1), 1061 respondents were included in the present study with 291 (27.4%) individuals who completed an advance directive and 770 (72.6%) with no advance directives (**Table 7**). Compared with those who have not completed an advance directive, respondents with advance directives were similar in sex and reporting having children < 18 years old but were generally older, non-Hispanic, White, and reported higher rates of Medicare insurance coverage (compared to Medicaid) and chronic conditions such as hypertension, diabetes, and depression (**Table 7**).

In the multivariable logistic regression analysis, advance directives were significantly associated with older age (compared to those age 18-34 years: 35-54 years [OR=2.7; 95%CI 1.5-

4.6] and 55+ years [OR=6.3; 95%CI 3.4-11.8]), having private (OR=2.9; 95%CI 1.1-7.2) or Medicare insurance (OR=5.0; 95%CI 1.9-13.0), being married (OR=2.0; 1.3-3.1), and holding favorable views towards passive euthanasia (OR=2.1; 95%CI 1.0-4.3) (**Table 8**). Individuals who were Black (OR=0.3; 95%CI 0.2-0.6) or other non-White race (OR=0.5; 95%CI 0.3-0.8) and expressed at least some financial difficulty (OR=0.5; 95%CI 0.3-0.9) were significantly less likely to complete an advance directive. Sex, Hispanic/Latino ethnicity, presence of chronic illness, and religiosity were not independently associated with advance directive status.

3.4.1.2 Mean differences in health values by advance directive status

Frequency histograms for observed cTTO values were similar except for values at -1, in which respondents with advance directives provided more than those without (19.9% vs. 12.1%) (**Figure 7**). After applying Bonferroni adjustment, unadjusted mean values were significantly different between those with and without advance directives for only 1 out of 86 health states (55555) (**Table 9**). In grouping the 86 health states into 17 misery scores, unadjusted mean values were lower among six groups by misery score (13, 15, 16, 17, 18, and 25) for those with advance directives relative to those without (**Table 10**) after applying Bonferroni adjustment.

Estimates from the random effects linear regression models showed that crude (*Model 1*; $\beta_{Advance}$=-0.149; 95%CI -0.210 to -0.087; p<0.001) and adjusted (*Model 2*; $\beta_{Advance}$=-0.101; 95%CI -0.175 to -0.028; p=0.007) mean values were significantly lower among respondents with advance directives compared to those without (**Table 11**). The addition of interaction terms (*Model 3*) found points estimates for all dimension-levels were lower among respondents with advance directives (range: -0.061 to -0.158) but statistically significant for Mobility levels 4 ($\beta_{Advance}$=-0.146; adjusted p=0.045) and 5 ($\beta_{Advance}$=-0.138; adjusted p=0.04), Usual Activities level 2 ($\beta_{Advance}$=-0.141; adjusted=0.035), Pain/Discomfort level 4 ($\beta_{Advance}$=-0.138; adjusted p=0.048), and

Anxiety/Depression level 4 ($\beta_{Advance}$=-0.158; adjusted p=0.008) (**Table 12**). Thus, the average individual with an advance directive valued health more negatively overall, and for specific dimension-levels, relative to the average individual without an advance directive.

3.4.1.3 Association between advance directive status and valuing health as worse-than-dead

The likelihood of valuing at least one of the assigned 10 EQ-5D-5L health states in a block as WTD was higher for respondents with advance directives after accounting for EQ-5D-5L descriptors and assigned health state block (*Model 4*; OR=1.54; 95%CI 1.35-1.65; p<0.001). However, after adjusting for covariates (*Model 5*), there was no association with valuing any health state as WTD and advance directive status (**Table 13**). Similarly, there was a significant crude association with valuing a given EQ-5D-5L health state as WTD (*Model 6;* OR=1.08; 95%CI 1.04-1.12; p<0.001) but did not remain significant after adjusting for potential confounding (*Model 7*).

3.4.1.4 Mean transition values for all possible health state values

Separate models were developed for individuals with and without advance directives and mean transition values calculated. All coefficients were negative, logically ordered, and were statistically significantly different from the preceding level in both value sets except Usual Activities (level 2 from level 1) in the model using values from those with advance directives (**Table 14)**. The mean (and standard deviation) predicted value for those with advance directives (0.198 [0.263) was lower than those without (0.363 [0.0.225]). The overall mean (SD) single-level transition value was 0.081 (0.013) and 0.066 (0.011) for individuals with and without advance directives, respectively, and all transition values appeared consistent across all EQ-5D-5L index scores. (**Figure 9**). Thus, those with advance directives are expected to provide a greater average gain or loss in health utility, relative to those without an advance directive, by 0.015 points (95%CI 0.011-0.012) for a better or worse transition from a given health state.

3.5 Discussion

The present study identified factors associated with advance directive status and evaluated differences in preferences between individuals with and without advance directives. The likelihood of having advance directives increased with age, being married, and having private or Medicare health insurance coverage, which is consistent with previous studies (Judy Campbell, Jo Edwards, Ward, & Weatherby, 2007; Rao et al., 2014). Race and socioeconomic status were negatively associated with having an advance directive, similar to other studies which found Blacks were more than three times less likely to complete an advance directive than Whites (Huang, Neuhaus, & Chiong, 2016; Koss & Baker, 2017). These results highlight that decisions related to health care, and not just access to care, are related to race and socioeconomic factors.

In examining differences in health preferences, individuals with advance directives valued EQ-5D-5L health states 0.10 points lower on average than those without, after adjusting for respondent characteristics. Based on the TTO task used to elicit these values, this difference represents one additional year over a 10-year time horizon that the average individual with an advance directive is willing to forgo, compared to those without an advance directive. A small but significant increase was observed in the overall mean transition value for the advance directive group. While this difference is modest, even small differences in the denominator (i.e., change in QALYs) can have a meaningful impact on cost per QALY estimates depending on the decision contexts. These findings suggest cost per QALY estimates based on preferences obtained from individuals with advance directives are likely to be lower than those without advance directives, resulting in more support for the adoption of health technologies that produce QALY gains.

In addition to finding that respondents with advance directives were willing to trade years of life for quality of life (i.e., shorter life in exchange for better health), we also examined whether they were more likely to value health states as WTD. The increased likelihoods of valuing health states as WTD with advance directive completion was not statistically significant, but the odds ratios for both analyses suggest an important magnitude of effect in the expected direction. Thus, the study may have been underpowered to detect an important difference between those with and without advance directives in terms of the likelihood of considering health states WTD.

An underlying explanation for the differences in health values among those with and without advance directives is the notion of informed preferences (M. Karimi, J. Brazier, & S. Paisley, 2017; Rowen et al., 2017). Informed preferences are defined as preferences for a good (e.g., health) that an individual would have if they "had all the relevant information and made full use of this information" (Harsanyi, 1996). By completing an advance directive, an individual has made a legally-binding and actionable decision based on their health preferences as it pertains to a meaningful life. Past studies have shown high rates of concordance between treatment preferences and severe health states often discussed in advance care planning (Brinkman-Stoppelenburg et al., 2014; Patrick et al., 1997). Further, preferences for end-of-life treatments were demonstrated to be stable over time across several illness contexts and sufficient in guiding treatment when the patient's preferences cannot be directly confirmed, and a medical decision is required (Auriemma et al., 2014; Barrio-Cantalejo et al., 2013). Thus, advance directive completion appears to be a valid indicator of clearly stated and informed preferences surrounding end-of-life care that remains stable over time. Future research should examine informedness in the context of advance directive status and other respondent characteristics to conceptualize the relationship between causal and effect indicators with respect to informedness about health preferences.

We are not aware of previous studies examining the relationship between advance directive status and stated preferences for health. Most studies have sought to identify characteristics associated with differences in health preferences but have typically focused on factors commonly collected in community surveys, such as age, sex, race/ethnicity, and geographical setting (Kind & Dolan, 1995; Sayah et al., 2016; James W Shaw et al., 2007) and not advance directive status. Other studies have examined various beliefs and attitudes towards death and their potential influence on health state preferences. Jakubczyk et al. investigated the association of self-reported belief in life after death on health state preferences among the general adult population in Poland. Stronger beliefs in life after death were associated with a decreased likelihood of considering a given health state as WTD and an increased likelihood of refusing to give up any years of life in TTO tasks (Jakubczyk, Golicki, & Niewada, 2016). Augestad et al. evaluated the relationship between attitudes towards euthanasia on TTO-based health values of EQ-5D-3L health states and the likelihood of valuing health states as WTD (L. A. Augestad, Rand-Hendriksen, Stavem, & Kristiansen, 2013). The authors observed lower mean values and an increased number of health states valued as WTD with increasing level of agreement with euthanasia practices. Advance directive status may be conceptually related to belief in life after death and attitudes towards euthanasia but agnostic with respect to predicting either view. That is, an individual with an advance directive is not necessarily required to have strong beliefs in life after death or be in favor of euthanasia. For instance, we observed a high prevalence of support (>75%) for both passive (e.g., withdrawal of care if there is no hope of meaningful recovery) and active (e.g., physician-assisted suicide) euthanasia for both advance directive statuses.

This study should also be viewed in the context of its limitations. First, all data was based on self-reported information, and could not be confirmed, but there was no reason to suspect respondents would provide a certain type of response to the question about advance directives, e.g. due to social desirability. There may also be additional variables that may confound the relationship between advance directive status and health values that was not captured, such as income/financial assets, belief in life after death, or numeracy. Furthermore, we were unable to discern whether individuals with advance directives requested less aggressive end-of-life care (e.g., do not resuscitate) under certain circumstances, or the opposite (e.g., “full code”). Therefore, we were unable to identify possible preference heterogeneity among individuals with advance directives. Future research could examine whether the association between advance directives and health preferences is modified by the presence of specific diseases or conditions, such as cancer and advanced chronic diseases like diabetes and heart failure.

3.6 Conclusions

This study found that advance directives were associated with a greater willingness to trade years of life for better health and a (non-significant) trend towards a greater likelihood to value poor health states as WTD. Overall mean transition values across all baseline health state index scores were significantly higher among individuals with advance directives, suggesting potentially larger utility gains or losses compared to individuals without advance care planning, which may result in lower (and therefore more favorable) cost per QALY estimates that support greater adoption of health technologies that improve health-related quality of life. Future research should examine informedness in the context of advance directive status and other respondent characteristics to conceptualize the relationship between causal and effect indicators with respect to informedness of health preferences.

TABLES

Table 7 Respondent characteristics by advance directive completion status

Characteristic	Advanced Directive completed n=291 (27.4%)	No advanced directive n=769 (72.6%)	p-value
Age, mean (SD)	60.4 (17.2)	41.2 (15.3)	**<0.001**
18-34, n (%)	24 (8.3)	323 (42.0)	
35-54, n (%)	87 (29.9)	278 (36.1)	
55+, n (%)	180 (61.9)	169 (21.9)	
Gender, n (%)			0.102
Male	130 (44.7)	385 (50.0)	
Female	158 (54.3)	383 (49.7)	
Other	3 (1.0)	2 (0.3)	
Race, n (%)			**<0.001**
White	236 (81.1)	425 (55.2)	
Black	20 (6.9)	107 (13.9)	
Other	35 (12.0)	238 (30.9)	
Hispanic, n (%)	24 (8.3)	167 (21.7)	**<0.001**
Education, post-secondary n (%)	217 (74.6)	486 (63.1)	**0.002**
History of illness, n (%)			
Hypertension	103 (35.4)	142 (18.4)	**<0.001**
Stroke	13 (4.5)	7 (0.9)	**<0.001**
Depression	60 (20.6)	210 (27.3)	**0.030**
Arthritis	103 (35.4)	141 (18.3)	**<0.001**
Asthma	34 (11.7)	91 (11.8)	0.975
Diabetes	41 (14.1)	54 (7.0)	**<0.001**
Cancer	38 (13.1)	21 (2.7)	**<0.001**
None	65 (22.2)	291 (37.8)	**<0.001**
Married, n (%)	224 (77.0)	319 (41.4)	**<0.001**
Living alone, n (%)	116 (39.9)	225 (29.2)	**<0.001**
With children <18 years old	54 (18.6)	150 (19.5)	0.733
Financial difficulty (somewhat to extremely), n (%)	37 (12.7)	179 (23.3)	**<0.001**
Insurance			**<0.001**
Medicaid	23 (7.9)	181 (23.5)	
Medicare	127 (43.6)	103 (13.4)	
Private (i.e., employer or self-coverage)	135 (46.4)	403 (52.3)	
None	6 (2.1)	83 (10.8)	
Experience with serious illness			
Personal	159 (54.6)	275 (35.7)	**<0.001**
Family member	260 (89.4)	637 (82.7)	**0.008**
Caring for others	198 (68.0)	396 (51.4)	**<0.001**
Views on euthanasia			
Generally agree with practice of active euthanasia	237 (81.4)	587 (76.2)	**0.021**
Generally agree with practice of passive euthanasia	275 (94.5)	653 (84.8)	**<0.001**
Religion tends to play role in approach to life	160 (55.0)	357 (46.4)	**0.012**

Table 8 Factors associated with advance directive completion

Factor	Odds ratio	95% CI
Age, (vs. 18-34 years)		
35-54	**2.7**[a]	**1.5 - 4.6**
55+	**6.3**[a]	**3.4 - 11.8**
Female (vs. male)	1.0	0.7 - 1.4
Race (vs. White)		
Black	**0.3**[a]	**0.2 - 0.6**
Other	**0.5**[a]	**0.3 - 0.8**
Hispanic (vs. non-Hispanic/Latino)	0.6	0.3 - 1.1
Post-secondary education (vs. high school or less)	**1.7**[a]	**1.2 - 2.7**
Presence of chronic illness (vs. none)	1.0	0.7 - 1.5
Married (vs. not)	**2.0**[a]	**1.3 - 3.1**
Living alone (vs. not)	**1.9**[a]	**1.3 - 2.8**
With children <18 years old (vs. none)	**1.8**[a]	**1.1 - 2.9**
Financial difficulty (somewhat to extremely), n (%)	**0.5**[a]	**0.3 - 0.9**
Insurance (vs. none)		
Private insurance	**2.9**[a]	**1.1 - 7.2**
Medicare	**5.0**[a]	**1.9 - 13.0**
Medicaid	1.0	0.4 - 2.9
Experience with serious illness (vs. none)		
Personally	**1.5**[a]	**1.1 - 2.2**
Immediate family member	1.4	0.8 - 2.3
Caring for others	**1.5**[a]	**1.0 - 2.1**
Views on euthanasia		
Agree with practice of active euthanasia (vs. disagree)	1.1	0.7 - 1.7
Agree with practice of passive euthanasia (vs. disagree)	**2.1**[a]	**1.0 - 4.3**
Religion tends to play role in approach to life (vs. unsure or does not play a role)	1.0	0.7 - 1.4

[a] Statistically significant association

Table 9 Observed mean health values for 86 EQ-5D-5L health states by advance directive completion status

Health state	Misery score	Advance directive holder (n=289) n	Mean [A]	SD	No advance directive (n=769) n	mean [B]	SD	Difference [A] - [B]
11112	6	51	0.912	0.295	158	0.952	0.176	-0.040
11121	6	72	0.955	0.113	149	0.969	0.085	-0.014
11122	7	31	0.903	0.197	85	0.887	0.267	0.016
11211	6	53	0.938	0.144	149	0.945	0.125	-0.007
11212	7	24	0.944	0.143	74	0.884	0.285	0.060
11221	7	31	0.873	0.369	76	0.926	0.122	-0.053
11235	12	31	0.258	0.776	76	0.596	0.473	-0.338
11414	11	37	0.377	0.572	80	0.400	0.638	-0.023
11421	9	29	0.690	0.490	79	0.673	0.493	0.016
11425	13	30	0.215	0.703	69	0.442	0.598	-0.227
12111	6	62	0.915	0.283	161	0.939	0.122	-0.024
12112	7	24	0.954	0.121	74	0.872	0.345	0.082
12121	7	23	0.878	0.184	80	0.884	0.259	-0.006
12244	13	29	0.131	0.716	79	0.253	0.649	-0.122
12334	13	24	0.219	0.741	79	0.492	0.573	-0.273
12344	14	35	0.166	0.720	69	0.308	0.641	-0.142
12513	12	35	0.550	0.542	69	0.633	0.472	-0.083
12514	13	31	0.316	0.683	76	0.470	0.531	-0.154
12543	15	23	0.059	0.719	80	0.229	0.637	-0.171
13122	9	30	0.723	0.520	69	0.864	0.279	-0.141
13224	12	31	0.373	0.657	85	0.443	0.615	-0.070
13313	11	29	0.781	0.204	79	0.694	0.453	0.087
14113	10	27	0.446	0.646	79	0.603	0.509	-0.156
14554	19	35	-0.056	0.744	69	0.078	0.644	-0.133
15151	13	27	-0.178	0.767	79	0.348	0.625	-0.526
21111	6	53	0.965	0.097	153	0.907	0.304	0.059
21112	7	35	0.894	0.171	69	0.828	0.422	0.066
21315	12	27	0.454	0.566	79	0.496	0.556	-0.042
21334	13	24	0.231	0.743	79	0.459	0.616	-0.228
21345	15	24	-0.088	0.720	74	0.111	0.728	-0.199
21444	15	37	0.026	0.631	80	0.170	0.673	-0.144
22434	15	30	0.082	0.780	69	0.428	0.554	-0.346
23152	13	24	0.248	0.741	74	0.207	0.751	0.040
23242	13	24	0.104	0.794	79	0.463	0.551	-0.358
23514	15	23	0.087	0.727	80	0.304	0.609	-0.217
24342	15	24	-0.165	0.736	79	0.308	0.611	-0.472
24443	17	27	-0.219	0.723	79	0.156	0.628	-0.375
24445	19	31	-0.027	0.666	85	-0.142	0.618	0.115
24553	19	30	-0.133	0.719	69	0.230	0.656	-0.364
25122	12	29	0.597	0.485	79	0.596	0.537	0.001

25222	13	37	0.385	0.592	80	0.535	0.618	-0.150
25331	14	37	0.366	0.539	80	0.464	0.623	-0.098
31514	14	37	0.228	0.614	80	0.362	0.659	-0.133
31524	15	27	0.133	0.740	79	0.375	0.599	-0.241
31525	16	29	0.212	0.661	79	0.201	0.686	0.011
32314	13	24	0.169	0.777	79	0.413	0.650	-0.244
32443	16	23	-0.091	0.725	80	0.196	0.640	-0.288
33253	16	24	-0.046	0.724	79	0.311	0.591	-0.357
34155	18	23	-0.080	0.699	80	0.051	0.665	-0.131
34232	14	31	0.435	0.583	85	0.373	0.646	0.063
34244	17	24	-0.227	0.701	74	0.073	0.662	-0.300
34515	18	31	-0.068	0.755	76	0.288	0.589	-0.356
35143	16	37	0.239	0.559	80	0.243	0.686	-0.004
35245	19	31	-0.098	0.730	76	0.182	0.632	-0.280
35311	13	31	0.394	0.691	85	0.522	0.591	-0.129
35332	16	30	0.278	0.638	69	0.588	0.484	-0.309
42115	13	30	0.187	0.733	69	0.486	0.556	-0.300
42321	12	31	0.542	0.598	85	0.563	0.571	-0.021
43315	16	27	0.070	0.718	79	0.312	0.639	-0.242
43514	17	24	0.035	0.664	74	0.214	0.693	-0.178
43542	18	23	-0.124	0.732	80	0.151	0.647	-0.275
43555	22	31	-0.116	0.700	85	-0.152	0.641	0.036
44125	16	35	0.110	0.745	69	0.241	0.613	-0.131
44345	20	35	-0.059	0.733	69	0.046	0.633	-0.104
44553	21	24	-0.290	0.630	74	-0.163	0.660	-0.127
45133	16	23	0.133	0.677	80	0.352	0.590	-0.219
45144	18	31	-0.216	0.717	76	0.084	0.684	-0.300
45233	17	29	0.171	0.718	79	0.353	0.666	-0.182
45413	17	30	-0.033	0.776	69	0.289	0.667	-0.322
51152	14	30	0.055	0.737	69	0.363	0.627	-0.308
51451	16	31	-0.074	0.744	76	0.264	0.649	-0.338
52215	15	23	-0.020	0.794	80	0.283	0.630	-0.303
52335	18	31	0.106	0.703	85	0.185	0.616	-0.078
52431	15	27	0.020	0.709	79	0.290	0.656	-0.270
52455	21	29	-0.207	0.770	79	-0.142	0.682	-0.065
53221	13	35	0.389	0.728	69	0.566	0.557	-0.177
53243	17	37	0.081	0.647	80	0.194	0.679	-0.113
53244	18	37	-0.142	0.647	80	0.104	0.702	-0.246
53412	15	24	0.013	0.748	79	0.364	0.635	-0.351
54153	18	27	-0.298	0.756	79	0.120	0.640	-0.418
54231	15	31	0.123	0.765	76	0.451	0.626	-0.328
54342	18	35	0.061	0.722	69	0.167	0.655	-0.106
55225	19	24	-0.215	0.718	79	0.092	0.654	-0.307
55233	18	29	0.121	0.759	79	0.247	0.664	-0.126
55424	20	24	-0.194	0.628	74	-0.088	0.710	-0.106
55555	**25**	**291**	**-0.475**	**0.610**	**770**	**-0.294**	**0.641**	**-0.181**[a]

[a] Indicates adjusted $p < 0.05$ (Bonferroni correction for multiple comparisons); SD = standard deviation; misery score = sum of each digit of health state profile

Table 10 Comparison of mean health values for EQ-5D-5L health state grouped by misery score and advance directive status

	Advance directive holder (n=291)			No advance directive (n=770)			Difference		
Misery score	n	Mean [A]	SD	n	mean [B]	SD	[A] - [B]	P-value	Bonferroni correction
6	291	0.938	0.202	770	0.942	0.180	-0.005	0.808	1
7	168	0.905	0.218	458	0.881	0.294	0.024	0.266	1
9	59	0.707	0.501	148	0.763	0.417	-0.056	0.447	1
10	27	0.446	0.646	79	0.603	0.509	-0.156	0.265	1
11	66	0.555	0.490	159	0.546	0.572	0.009	0.909	1
12	184	0.463	0.615	473	0.551	0.544	-0.088	0.090	1
13	**370**	**0.228**	**0.723**	**996**	**0.435**	**0.613**	**-0.207**[a]	**<0.001**	**<0.001**
14	170	0.253	0.646	383	0.376	0.639	-0.124	0.037	0.6307
15	**293**	**0.029**	**0.724**	**855**	**0.300**	**0.638**	**-0.271**[a]	**<0.001**	**<0.001**
16	**259**	**0.103**	**0.688**	**691**	**0.298**	**0.630**	**-0.195**[a]	**<0.001**	**<0.001**
17	**171**	**-0.021**	**0.710**	**455**	**0.213**	**0.668**	**-0.234**[a]	**<0.001**	**0.003**
18	**267**	**-0.067**	**0.722**	**704**	**0.155**	**0.652**	**-0.222**[a]	**<0.001**	**<0.001**
19	151	-0.099	0.710	378	0.080	0.650	-0.179	0.008	0.1123
20	59	-0.114	0.690	143	-0.023	0.675	-0.090	0.393	1
21	53	-0.244	0.704	153	-0.152	0.669	-0.092	0.402	1
22	31	-0.116	0.700	85	-0.152	0.641	0.036	0.802	1
25	**291**	**-0.475**	**0.610**	**770**	**-0.294**	**0.641**	**-0.181**[a]	**<0.001**	**<0.001**

[a] Indicates adjusted $p < 0.05$ (Bonferroni correction for multiple comparisons); SD = standard deviation; misery score = sum of each digit of health state profile

Table 11 Crude and adjusted mean differences for health state values between respondents with and without advance directives

	Crude model			Adjusted model		
	Estimate[a]	95% CI	p-value	Estimate[a,b]	95% CI	p-value
Mean difference in health values	-0.149	-0.210 to -0.087	<0.001	-0.101	-0.175 to -0.028	0.007

[a] Adjusted for EQ-5D-5L health state.

[b] Adjusted for age, sex, race, ethnicity, education, marital status, having child(ren) <18, insurance coverage, financial coverage, living alone, presence of chronic conditions, experience with serious illness (personal, family member, caregiving), views towards euthanasia (active or passive), religiosity, EQ-VAS.

Table 12 Adjusted mean differences for EQ-5D-5L dimension-levels for respondents with and without advance directives

Dimension-level	Estimated difference	Standard Error	Unadjusted p-value	Adjusted p-value
$MO_{level-2}$	-0.097	0.044	0.029	0.715
$MO_{level-3}$	-0.126	0.046	0.006	0.153
$MO_{level-4}$	**-0.146**[a]	**0.047**	**0.002**	**0.045**
$MO_{level-5}$	**-0.138**[a]	**0.044**	**0.002**	**0.040**
$SC_{level-2}$	-0.124	0.045	0.006	0.158
$SC_{level-3}$	-0.108	0.046	0.020	0.498
$SC_{level-4}$	-0.132	0.045	0.003	0.085
$SC_{level-5}$	-0.130	0.043	0.003	0.063
$UA_{level-2}$	**-0.141**[a]	**0.044**	**0.001**	**0.035**
$UA_{level-3}$	-0.106	0.046	0.021	0.515
$UA_{level-4}$	-0.123	0.046	0.008	0.190
$UA_{level-5}$	-0.113	0.045	0.012	0.300
$PD_{level-2}$	-0.104	0.045	0.020	0.503
$PD_{level-3}$	-0.119	0.047	0.012	0.290
$PD_{level-4}$	**-0.138**[a]	**0.044**	**0.002**	**0.048**
$PD_{level-5}$	-0.137	0.047	0.004	0.095
$AD_{level-2}$	-0.096	0.045	0.034	0.843
$AD_{level-3}$	-0.111	0.046	0.015	0.385
$AD_{level-4}$	**-0.158**[a]	**0.044**	**0.000**	**0.008**
$AD_{level-5}$	-0.089	0.043	0.039	0.983

[a] Indicates adjusted $p < 0.05$.

[b] Adjusted for age, sex, race, ethnicity, education, marital status, having child(ren) <18, insurance coverage, financial coverage, living alone, presence of chronic conditions, experience with serious illness (personal, family member, caregiving), views towards euthanasia (active or passive), religiosity, EQ-VAS

[b] Bonferroni correction applied for multiple comparisons.

AD=Anxiety/Depression; MO=Mobility; PD=Pain/Discomfort; SC=Self-Care; UA=Usual Activities

Table 13 Crude and adjusted odds ratios for valuing health states as worse-than-dead between respondents with and without advance directives

	Crude model			Adjusted model		
	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Odds ratio for valuing at least 1 health state as worse-than-dead	1.54	1.35 - 1.65	<0.001	1.29[a]	0.91– 1.84	0.146
Odds ratio for valuing a specific health state as worse-than-dead	1.08[b]	1.04 - 1.12	<0.001	1.62[a,b]	0.99 – 2.60	0.053

[a] Adjusted for age, sex, race, ethnicity, education, marital status, having child(ren) <18, insurance coverage, financial coverage, living alone, presence of chronic conditions, experience with serious illness (personal, family member, caregiving), views towards euthanasia (active or passive), religiosity, EQ-VAS

[b] Adjusted for EQ-5D-5L dimension-level severity

Table 14 Comparison of model parameter estimates and value set characteristics by advance directive status

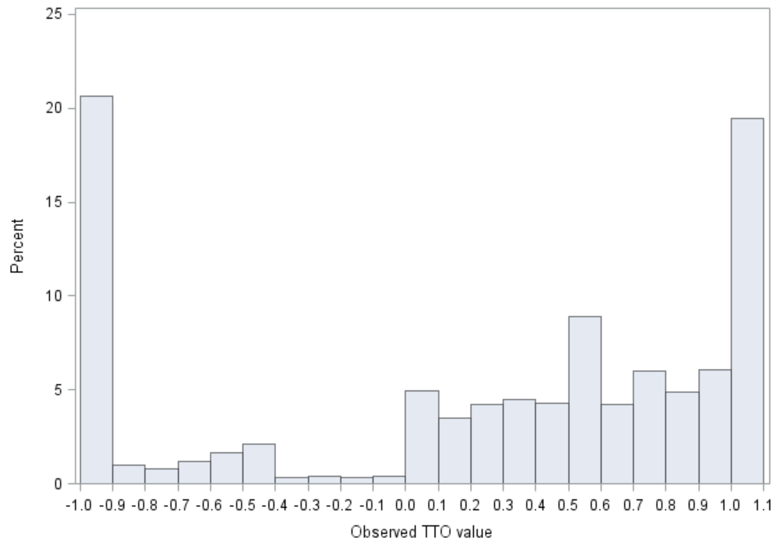
Parameter[a]		Advance Directive			No Advance Directives		
		β	SE	p-value	β	SE	p-value
Mobility	Slight	**-0.111**	**0.028**	**0.001**	**-0.075**	**0.016**	**<0.001**
	Moderate	**-0.171**	**0.029**	**<0.001**	**-0.107**	**0.016**	**<0.001**
	Severe	**-0.268**	**0.031**	**<0.001**	**-0.184**	**0.018**	**<0.001**
	Unable to	**-0.317**	**0.029**	**<0.001**	**-0.240**	**0.017**	**<0.001**
Self-care	Slight	**-0.106**	**0.028**	**0.006**	**-0.058**	**0.016**	**0.0003**
	Moderate	**-0.128**	**0.031**	**<0.001**	**-0.094**	**0.018**	**<0.001**
	Severe	**-0.239**	**0.031**	**<0.001**	**-0.181**	**0.018**	**<0.001**
	Unable to	**-0.262**	**0.028**	**<0.001**	**-0.206**	**0.016**	**<0.001**
Usual Activities	Slight	-0.102	0.029	0.066	**-0.046**	**0.016**	**0.005**
	Moderate	**-0.110**	**0.031**	**0.001**	**-0.087**	**0.018**	**<0.001**
	Severe	**-0.262**	**0.030**	**<0.001**	**-0.223**	**0.017**	**<0.001**
	Unable to	**-0.224**	**0.029**	**<0.001**	**-0.196**	**0.016**	**<0.001**
Pain/ Discomfort	Slight	**-0.088**	**0.026**	**0.005**	**-0.054**	**0.015**	**0.0003**
	Moderate	**-0.139**	**0.031**	**<0.001**	**-0.091**	**0.018**	**<0.001**
	Severe	**-0.332**	**0.028**	**<0.001**	**-0.265**	**0.016**	**<0.001**
	Extreme	**-0.398**	**0.030**	**<0.001**	**-0.331**	**0.017**	**<0.001**
Anxiety/ Depression	Slight	**-0.037**	**0.030**	**0.048**	**-0.054**	**0.017**	**0.0017**
	Moderate	**-0.119**	**0.034**	**<0.001**	**-0.122**	**0.019**	**<0.001**
	Severe	**-0.316**	**0.030**	**<0.001**	**-0.270**	**0.017**	**<0.001**
	Extreme	**-0.278**	**0.028**	**<0.001**	**-0.302**	**0.016**	**<0.001**
Range (11111 to 55555)				1.479			1.276
Mean (SD)				0.198 (0.263)			0.363 (0.225)
Median (IQR)				0.201 (0.364)			0.366 (0.312)
Dimension importance				PD-MO-AD-SC-UA			PD-AD-MO-SC-UA
Number (%) of health states WTD				724 out of 3,125 (23.2%)			185 out of 3,125 (5.9%)
33333				0.0.334			0.499
55555				-0.479			-0.276

WTD=worse-than-dead; SD=standard deviation; SE=standard error; IQR=interquartile range

[a] All coefficients adjusted for age, sex, race, ethnicity, education, marital status, having child(ren) <18, insurance coverage, financial coverage, living alone, presence of chronic conditions, experience with serious illness (personal, family member, caregiving), views towards euthanasia (active or passive), religiosity, EQ-VAS.

Figure 7 Frequency distributions of observed time trade-off based health values for individuals a) with an advance directive and b) without an advance directive

a) With advance directives



b) Without advance directive

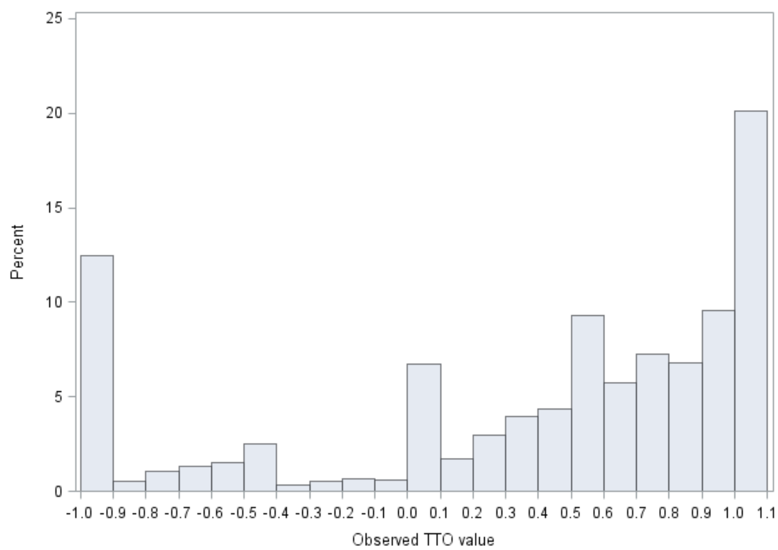
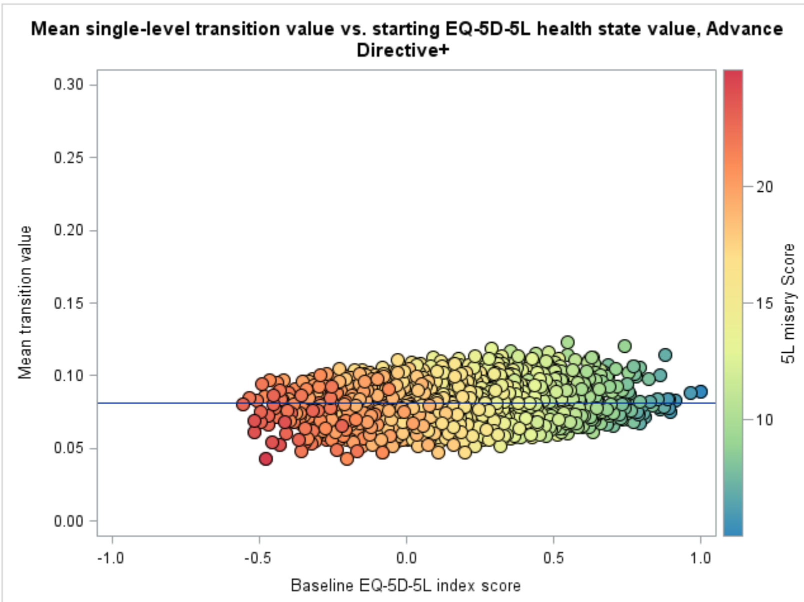
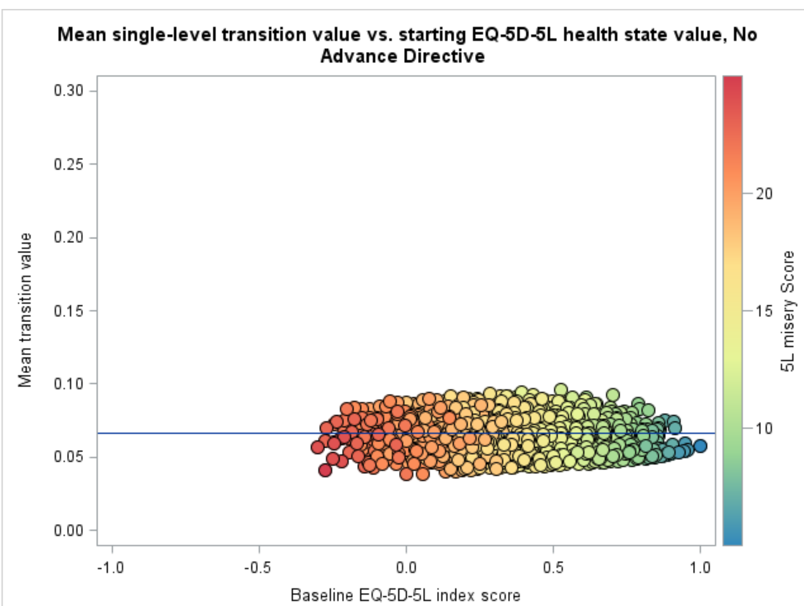


Figure 8 Mean transition values for all EQ-5D-5L index scores by advance directive

A) Individuals with advance directives



B) Individuals without advance directives



4. TIME-SPECIFIC DIFFERENCES IN STATED PREFERENCE FOR HEALTH IN THE UNITED STATES

Authors: Ernest H. Law, PharmD; A. Simon Pickard, PhD; Surrey M. Walton, PhD; Feng Xie, PhD; Todd A. Lee, PharmD, PhD; Alan Schwartz, PhD (**s*ubmitted to PharmacoEconomics***)

4.1 Abstract

Background: Changes in societal values over time may be an important reason to consider updating societal value sets for preference-based measures of health. **Objective**: To examine whether health preferences are different between 2002 and 2017, controlling for demographic changes in the United States. **Methods:** Data from 2002 and 2017 US EQ-5D-3L valuation studies were combined. The primary analysis compared valuations of better-than-dead (BTD) states only, as both studies used the same time trade-off (TTO) method for these states. For worse-than-dead (WTD) states, the 2017 study used the lead-time TTO and the 2002 study used the conventional TTO, which necessitated transformation. Regression models were fitted to BTD values to estimate time-specific differences, adjusting for respondent characteristics. Secondary analyses examined models that fitted WTD values (using linear and non-linear transformations of the 2002 data) and all values. **Results:** The adjusted BTD-only model showed mean values were higher for 2017 compared to 2002 (β_{2017}=0.05, $p<0.001$). WTD-only models showed negative changes over time but that were dependent on the transformation method (linear β_{2017}=-0.72; non-linear β_{2017}=-0.35; both $p<0.001$). Using all values, 2017 mean valuations were lower using a linear transformation (β_{2017}=-0.11; $p<0.001$) but did not differ with the non-linear transformation. **Conclusions:** Results suggest that societal preferences for health states changed over time. For

the most methodologically comparable data (BTD-only), values in 2017 were modestly higher, implying individuals in 2017 were generally less willing to trade quantity for quality of life compared to 2002.

4.2 Introduction

The values of a generic preference-based measure of health are used to inform resource allocation in healthcare and should represent the preferences of the relevant population (CADTH, 2017; ICER, 2017; PBAC, 2016; Sanders et al., 2016). Several prominent societal value sets were developed several decades ago (Dolan, 1997; David Feeny et al., 2002; J. W. Shaw et al., 2005). With time, demographics shift, values change, and science progresses (A. S. Pickard, 2015). Therefore, the issue of whether societal value sets should be re-visited is a pertinent question as value sets should represent the current values/preferences of the population of interest (A. S. Pickard, 2015).

Interest in developing a new societal value set in a country or jurisdiction is often catalyzed by innovations related to a preference-based measure rather than a concern about time-specific differences in health values. For instance, with the introduction of several EQ-5D-based measures,(N. J. Devlin et al., 2018; Wille et al., 2010) alternative preference elicitation techniques were implemented in the EuroQol's international standardized protocol for health state valuation studies (Krabbe et al., 2014; NICE, 2017; Oppe et al., 2014; Feng Xie et al., 2015). Moreover, studies comparing differences between societal value sets have focused on cross-country comparisons rather than evaluating the magnitude of differences in health preferences over time (Greiner et al., 2003; J. A. Johnson et al., 2005; Olsen et al., 2017). A possible reason for the absence of such comparisons is the lack of methodological comparability in valuation studies conducted in different time periods. If similar study designs for the same preference-based measure could be identified, the issue of time-related preferences could be examined.

Fifteen years have passed since the US societal valuation of the EQ-5D-3L, a three-level version of the EQ-5D (J. W. Shaw et al., 2005). Since then, the US has undergone several changes beyond demographic shifts (Hobbs & Stoops, 2002; Humes, Jones, & Ramirez, 2011) that may influence how members of the general public value health. Most notably, the Patient Protection and Affordable Care Act was enacted in 2010 and represented the most significant restructuring of the health insurance landscape in the US in decades (Blumenthal, Abrams, & Nuzum, 2015). There is also an increasing recognition among the American public that the US economy is moving away from physical skills and manual labor and towards knowledge- and technology-focused jobs (Pew, 2016). Moreover, Americans' personal views on health, particularly in the context of end-of-life treatment, have changed (Lugo, Cooperman, Funk, & O'Connell, 2013). Taken together, many factors may have led to changes in societal preferences for health in the US. The purpose of the present study is to examine time-specific differences in health preferences between 2002 and 2017.

4.3 Methods

4.3.1 Data sources

Individual-level data from two US EQ-5D-3L valuation studies conducted in 2002 and 2017 were pooled. The study designs for both studies have been previously described [7, 23]. Briefly, both data sets were comprised of responses collected through face-to-face interviews conducted by teams of trained interviewers, targeting English and Spanish-speaking non-institutionalized adults (≥ 18 years old) living in the United States. The 2002 study utilized a multistage probability sampling strategy based on residential mailing lists and oversampled Hispanic/Latino and non-Hispanic Blacks to investigate racial and ethnic differences in health valuation [7, 24]. The 2017 study utilized a quota sampling, where quotas were met based on US census demographic information [23].

4.3.2 **Health state selection**

All health states valued were defined using the EQ-5D-3L descriptive system with five dimensions (Mobility [MO], Self-Care [SC], Usual Activities [UA], Pain/Discomfort [PD], and Anxiety/Depression [AD]) [25]. Each dimension contains three levels: no problems (level 1), some or moderate problems (level 2), and unable to/extreme problems (or confined to bed for the Mobility dimension) (level 3). Health states described with the EQ-5D-3L are defined using a five-digit profile representing the level for each dimension of health. For example, “33333” describes the most severe health state, with extreme problems (i.e., level 3) in all dimensions. A “misery score” can also be calculated as a proxy indicator for health state severity by summing the five digits of the health state descriptor. For example, the misery score for “13231” is equal to 10. The highest and lowest misery score for the 3L is 15 (“33333”) and 5 (“11111”), respectively.

In the 2002 study, a total of 42 EQ-5D-3L health states were included in the experimental design and each respondent valued 12 health states. The 2017 study included 30 health states and each respondent was asked to value three health states [26]. The 2002 and 2017 studies shared a total of 16 common health states directly valued using the TTO method.

4.3.3 **Preference elicitation: time trade-off methods**

Values for better-than-dead (BTD) health states were elicited using the same conventional TTO in both studies, except that the 2002 study used pen & paper and a “wheel prop” to assist the respondent with their selection and the 2017 study used a computer-based platform called the EuroQol Valuation Technology (EQ-VT) [27]. The conventional TTO task presented respondents with two alternatives, Life A or Life B, and were asked to select one. In Life B, respondents were asked to imagine that they had 10 years in a suboptimal health state. Life A involved living in full health for x years, where $x \leq 10$. The value of x was varied until the respondent was indifferent between Life A and Life B. The TTO value was then calculated as x/10, which meant values for states BTD ranged from 0 to 1.

The two studies differed with respect to the preference elicitation technique used to obtain values for health states that were considered by respondents to be worse-than-dead (WTD). The 2017 study used the composite TTO, a method that employs the lead-time TTO to obtain WTD values [28, 29]. The lead-time TTO, like the conventional TTO, presents respondents with a choice between Life A or Life B, except that Life B is modified so that a respondent hypothetically lives 10 years in full health followed by 10 years in the suboptimal (WTD) state. Like the conventional TTO, Life A involves living in full health for x years, where $x < 10$ and x is varied until an indifference point is identified. The value for states WTD using the lead-time TTO is calculated as (x-10)/10. In contrast, the 2002 study modified the conventional TTO task for health states considered WTD. With this modification, the respondent was asked to choose between: (a) living in the suboptimal health state for (10 - x) years, where $x < 10$, then living in full health for x years or choosing (b) immediate death [30, 31]. With this modification, the lowest possible raw TTO value was -39, calculated when death was considered equal to 3 months (0.25 years) in a given health state was followed by 9.75 years in full health i.e., 9.75/(9.75-10) = -39. In the literature, there are two approaches to handle the skewed nature of TTO values for WTD states obtained using the modified conventional TTO employed in the 2002. The first is to apply a linear transformation to all negative TTO values collected by dividing values by 39 [32]. The 2005 EQ-5D-3L US value set was developed based on models using linear transformed TTO values [7]. A second approach, proposed by Patrick et al., is to apply a non-linear transformation, where TTO value is divided by 1 minus the TTO value [33]. Both transformations results in WTD values ranging between -1 and 0 (**Table 14**).

4.3.4 Statistical analysis

Independent two-sample t-tests were performed to compare the mean valuations for the 16 common health states that restricted observations to TTO values ≥ 0 (i.e., BTD-only). An additional series of independent t-tests comparing the 16 common health states that included all TTO values. Mean values were weighted using sampling weights from the original 2002 study, which were developed to account for the oversampling of respondents who were non-Hispanic Black and Hispanic/Latino (J. W. Shaw et al., 2005). Due to the large number of comparisons with the 16 commonly valued EQ-5D-3L health states, a Bonferroni correction was applied to reduce type I error (alpha = 0.05 after Bonferroni correction). Correlations between mean valuations for 2002 and 2017 were examined using Spearman's rank correlation and the intra-class correlation coefficient (ICC) for one-way random effects for absolute agreement, including 95% confidence intervals (CI) (Koo & Li, 2016). Finally, the difference in the proportion of observations that indicated the health state was WTD were compared between 2017 and 2002 with Chi-square tests (alpha = 0.05 after applying a Bonferroni correction).

Differences in valuations between time-periods were investigated using several random-effects linear regression models with a random intercept. The independent variable of interest was the time period in which the EQ-5D health valuation interview was completed with the 2002 study was designated the referent category. The dependent variable was the TTO value for a given EQ-5D-3L health state. The primary analysis focused on values for health states considered BTD by respondents (i.e., positive values only). Two specifications were used to assess time-specific differences in health values. *Specification 1* was developed to estimate the mean difference in health state valuation between time periods while adjusting for differences in respondent characteristics:

$$TTO_{ij} = \alpha + \beta1(2017\text{-}2002)_i + \beta2(MO_{level\text{-}2})_i + \beta3(MO_{level\text{-}3})_i + \beta4(SC_{level\text{-}2})_i + \beta5(SC_{level\text{-}3})_i + \beta6(UA_{level\text{-}2})_i + \beta7(UA_{level\text{-}3})_i + \beta8(PD_{level\text{-}2})_i + \beta9(PD_{level\text{-}3})_i + \beta10(AD_{level\text{-}2})_i + \beta11(AD_{level\text{-}3})_i + \beta12(age_{18\text{-}34})_i + \beta13(age_{\geq 55})_i + \beta14(female)_i + \beta15(race_{Black})_i + \beta16(race_{Other})_i + \beta17(Hispanic)_i + \beta18(Education_{>high\ school})_i + \beta19(Condition_{1\ condition\ only})_i + \beta20(Condition_{\geq 2\ conditions})_i + \beta21(EQ\text{-}VAS)_i + \upsilon_i + \varepsilon_{ij}$$

Where TTO_{ij} is the value given by the *i*-th respondent for the *j*-th EQ-5D-3L health state; v_i is the random effects term representing the extent to which the intercept of the *i*-th respondent deviated from the overall intercept. The covariates were: age (18-35, 36-54, and ≥55 years old), sex (female or not), race (White, Black, or Other), ethnicity (Hispanic or not), education level (received higher education [more than high school] or not), and reported health conditions (none, 1 condition, or ≥2 conditions), and self-rated health (EQ-VAS) (**Table 15**).

Specification 2 included additional interaction terms for a given health state descriptor and time-period (i.e., 2017*[MO/SC/UA/PD/AD]). The purpose of this model was to assess whether differences in health preferences between time-periods varied across dimension-levels, while adjusting for respondent characteristics:

$$TTO_{ij} = \alpha + \beta1(2017\text{-}2002)_i + \beta2(MO_{level\text{-}2})_i + \beta3(MO_{level\text{-}3})_i + \beta4(SC_{level\text{-}2})_i + \beta5(SC_{level\text{-}3})_i + \beta6(UA_{level\text{-}2})_i + \beta7(UA_{level\text{-}3})_i + \beta8(PD_{level\text{-}2})_i + \beta9(PD_{level\text{-}3})_i + \beta10(AD_{level\text{-}2})_i + \beta11(AD_{level\text{-}3})_i + \beta12(age_{18\text{-}34})_i + \beta13(age_{\geq 55})_i + \beta14(female)_i + \beta15(race_{Black})_i + \beta16(race_{Other})_i + \beta17(Hispanic)_i + \beta18(Education_{>high\ school})_i +$$

$\beta 19(\text{Condition}_{1\text{ condition only}})_i + \beta 20(\text{Condition}_{\geq 2\text{ conditions}})_i + \beta 21(\text{EQ-VAS})_i + \beta 22(2017{*}\text{MO}_{\text{level-2}})_i + \beta 23(2017{*}\text{MO}_{\text{level-3}})_i + \beta 24(2017{*}\text{SC}_{\text{level-2}})_i + \beta 25(2017{*}\text{SC}_{\text{level-3}})_i + \beta 26(2017{*}\text{UA}_{\text{level-2}})_i + \beta 27(2017{*}\text{UA}_{\text{level-3}})_i + \beta 28(2017{*}\text{PD}_{\text{level-2}})_i + \beta 29(2017{*}\text{PD}_{\text{level-3}})_i + \beta 30(2017{*}\text{AD}_{\text{level-2}})_i + \beta 31(2017{*}\text{AD}_{\text{level-3}})_i + \upsilon_i + \varepsilon_{ij}$

Mean differences for all 15 dimension-levels were then estimated with a significance level of 0.05 Bonferroni correction.

In a secondary analysis, two additional sets of models were developed to examine differences between the two time-periods that may be attributable to the variations in preference elicitation techniques and transformation approach to WTD values (i.e., values < 0). The first set fitted WTD-only values and the second set fitted all TTO values (i.e., BTD and WTD values). *Specifications 1* and *2* were examined for both sets of models. Linear and non-linear transformations of the 2002 WTD data were applied. Coefficients were considered statistically significant at an alpha level of 0.05.

An analogous multivariable random effects logistic regression with random intercept at the respondent-level was performed where TTO values were analyzed as a dichotomous outcome: WTD (i.e., TTO value < 0) vs. BTD (i.e., TTO value ≥ 0). The purpose of this model was to assess the likelihood of respondents valuing a health state as WTD for 2017 vs 2002 after controlling for respondent characteristics and health state severity. Crude and adjusted odds ratios (OR), 95% CIs, and p-values were reported for the likelihood of 2017 respondents valuing a health state as WTD. The crude model only included variables for the time period and EQ-5D-3L descriptors (i.e., 2017 + 10 EQ-5D-3L dummy variables) and the adjusted model followed *Specification 1*. All statistical analyses were conducted using, SAS 9.4 (Chapel Hill, NC).

4.4 Results

After excluding respondents with incomplete or unusable data, 3773 and 1062 respondents were from the 2002 and 2017 studies, respectively, were included in the present study (**Table 16**). Compared with the 2002 weighted sample, more respondents in 2017 received education beyond the high school level, fewer had chronic conditions, and self-reported problems in EQ-5D-3L Mobility, Self-Care, and Usual Activities. Mean age, sex, and self-rated health were comparable between the two study samples (**Table 16**). Visual inspection of the frequency distributions for all observed TTO values in 2017 and 2002 showed that 2017 values tended to cluster at the ends of the scale (-1 and 1), with few values between -0.05 and 0 (**Figure 8**). The 2002_{linear} distributions appear to cluster 0 and 1, with few values below 0. Lastly, $2002_{non\text{-}linear}$ values also appeared to cluster around 0 and 1, but there were considerably more values between -1 and 0 compared to 2017 or 2002_{linear}.

Restricting to BTD-only values, mean values for two out of 16 health states (12111 and 11113) were significantly higher in 2017 than in 2002 (**Table 17**). For all values, estimated mean valuations for the 16 common EQ-5D-3L health states ranged from -0.42 (33333) to 0.93 (12111) for the 2017, -0.10 (33333) to 0.87 (21111) for the 2002 using linear transformation for WTD values, and -0.38 (33333) to 0.87 (11211 and 21111) for 2002 using the non-linear transformation (**Table 18**). Mean valuations in 2002 using the linear transformation were statistically significantly different than the 2017 for 5 out of 16 health states (adjusted p-values < 0.01) while the non-linear transformation resulted in no significant differences (**Table 18**). No differences were observed for the proportion of common health states that were considered WTD.

BTD-only valuations for 2017 and 2002 were also highly correlated (Spearman's rho=0.94) and in high agreement (ICC=0.98; 95%CI 0.94-0.99]) (**Figure 9**). The degree of the differences between the 2017 and 2002 US mean valuations appeared to increase with severity of health states for the linear transformed values (**Figure 10**). Differences in mean valuations using the non-linear transformation values remained consistent across health state severity (**Figure 10).** Mean valuation scores were highly correlated (Spearman's rho=0.97 using both linear and non-linear transformations) and in high agreement (ICC_{linear}=0.96; 95%CI 0.90-0.99; $ICC_{non\text{-}linear}$=0.99; 95%CI 0.99-1.00).

In the adjusted models (*Specification 1*), the BTD-only model showed a statistically significant increase (β_{2017}=0.05, 95%CI 0.03 to 0.07, p<0.001) in mean valuations for 2017 compared with 2002 (**Table 4**). The WTD-only models showed a decrease in mean valuations for 2017 respondents irrespective of the transformation approach (linear β_{2017}=-0.71, 95%CI -0.74 to -0.69; non-linear β_{2017}=-0.34, 95%CI -0.36 to -0.31; both p<0.001). The model that included all values showed that mean valuations between 2017 and 2002 differed depending on whether linear (β_{2017}=-0.11, 95%CI -0.13 to -0.08, p<0.001) or non-linear (β_{2017}=-0.01, 95%CI –0.01 to 0.02, p=0.3) transformations were applied (**Table 19**). Full model results with parameter estimates for all variables are shown in **Table 20**. In the random effects logistic regression models, respondents in 2017 were slightly less likely to value health as WTD in both crude (OR=0.98 [95%CI 0.96-0.99; p=0.008]) and adjusted (OR=0.97 [95%CI 0.95-0.98; p<0.001]) analyses.

The models with additional interaction terms (*Specification 2*) found significant dimension-level difference between time periods that included BTD values only for Self-Care level 3 (0.04; adjusted p=0.01) (**Table 21**). For WTD-only models, values demonstrated that 2017 respondents had systematically lower mean valuations across all dimension-levels when applying the linear transformation (coefficient range: -0.71 to -0.78; all adjusted p-values p<0.001). With the non-linear transformation, 2017 values were significantly lower for all dimension-levels (coefficient range: -0.335 to -0.410; all adjusted p-values <0.001). The models including all TTO values showed that application of linear or non-linear transformation produced different results, with linear transformations resulting 10 out of 10 dimension-level variables being statistically significant (all adjusted p-values <0.05). In contrast, the non-linear transformation resulted in only 2 out of 10 dimension-levels with statistical significance (**Table 21**).

4.5 Discussion

This study found stated health preferences were moderately different when collected from the US population 15 years apart, accounting for differences in study design and population demographics. By restricting the analysis to BTD values only, which were obtained using the same TTO method, we found that EQ-5D-3L health states were valued 0.05 points higher in 2017 than in 2002, after adjusting for sociodemographic differences. While a difference between time periods of 0.05 could be interpreted as modest, this TTO-based value translates into 6 months that the average American in 2017 is unwilling to trade, irrespective of health state severity, compared to those in 2002. The value of 0.05 was comparable to minimally important differences for the EQ-5D-3L derived from samples of patients with chronic conditions (0.074) (Walters & Brazier, 2005), cancer (0.06) (A. S. Pickard, Neary, & Cella, 2007), and post-traumatic stress disorder (0.05 to 0.08) (Le, Doctor, Zoellner, & Feeny, 2013). Furthermore, the increase in mean valuations among 2017 respondents was consistent with the slight decrease in likelihood of endorsing a health state as WTD, adjusting for health state severity and sociodemographic differences. Therefore, this study provides evidence of potentially meaningful time-specific differences in health state preferences, which may be an important reason to consider updating societal value sets.

Other studies have compared widely utilized country-specific value sets (M. F. Janssen et al., 2018; J. A. Johnson et al., 2005; Mulhern et al., 2018). For example, Mulhern et al. compared differences between two EQ-5D value sets (Mulhern et al., 2018): the EQ-5D-3L algorithm developed by Dolan based on the 1993 MVH study (Dolan, 1997) and the England EQ-5D-5L value set by Devlin et al (N. J. Devlin et al., 2018). However, important methodological differences may explain the differences observed beyond the fact that the studies were collected two decades apart, such as the different EQ-5D descriptive systems, study protocols, settings, and modeling approaches in value set development. The present study is strengthened by using the same descriptive system, modelling approach, and preference elicitation technique.

The estimation of health state values was sensitive, both in direction and magnitude, to the preference elicitation method and the choice of transformation approach for states WTD. Health states WTD continue to be controversial within the health valuation field (L. A. Augestad et al., 2013; Jakubczyk et al., 2016). We developed several models to understand the influence of using non-linear vs. linear transformations and lead-time vs. modified conventional TTO. Mean observed valuations between 2017 and were more comparable when using the non-linear transformation compared to linear transformation. After adjusting for respondent characteristics, we found 2017 values to be, on average, 0.11 points lower than those in 2002 using the linear transformation. However, no difference was observed when the non-linear transformation was applied. In evaluating the potential influence of the two TTO variants on elicited values, the WTD-only models showed a large difference in values obtained from 2017 respondents compared to those in 2002, irrespective of whether the linear or non-linear transformation was applied. This average difference was also seen across *all* 15 EQ-5D dimension-levels for both linear and non-linear transformations. This suggests that the lead-time results in systematically lower values compared to the conventional TTO rather than actual differences in health preferences.

As noted earlier, different preference elicitation methods were used to value WTD health states in 2017 and 2002 (i.e., lead-time vs. modified conventional TTOs), and more ideally, the same methods would have been used. Consequently, we separately examined health states valued as WTD and BTD, and analyzed the WTD data using several approaches and transformations. Other study design differences cannot be completed ruled out as potential sources of variation in valuations between the time periods. For instance, the 2017 study implemented routine quality control measures to mitigate respondent confusion and reduce interviewer effects, while the 2002 study did not report any regular quality control processes. The combined total influence of these differences in protocol are impossible to disentangle from the obtained valuations and the results should be interpreted cautiously.

4.6 Conclusion

This study provides evidence of time-specific differences in a society's preferences and suggests that the era in which values were elicited may be an important reason to consider updating societal value sets. For the most methodologically comparable data, values in 2017 were modestly higher, which implies people were less willing to trade time for quality of life than in 2002, i.e. 6 months over 10 years. The large differences between 2002 and 2017 when WTD data were included appeared to be driven by differences in methodology, and also illustrated the profound impact of the choice of transformation on estimated values. This study focused on time as a reason to update societal value sets independent of shifts in population subgroups and therefore removed the effect of factors such as age, race/ethnicity, and sex. However, the demographic composition of the general population has been associated with different values (e.g., Black assigning less disutility to extreme health problems (James W Shaw et al., 2007) and may represent an additional reason to revisit and update value set that warrants further investigation.

Table 15 List of health state values for all possible time trade-off scenarios in 2002 study protocol, including application of linear and non-linear transformations

Time horizon	Time in full health	Time in suboptimal health	Calculated raw TTO value	$TTO_{non\text{-}linear}$	TTO_{linear}
10	0	10	0	0	0.00
10	-0.25	9.75	-0.03	-0.03	0.00
10	-0.5	9.5	-0.05	-0.05	0.00
10	-0.75	9.25	-0.08	-0.08	0.00
10	-1	9	-0.11	-0.1	0.00
10	-1.25	8.75	-0.14	-0.13	0.00
10	-1.5	8.5	-0.18	-0.15	0.00
10	-1.75	8.25	-0.21	-0.18	-0.01
10	-2	8	-0.25	-0.2	-0.01
10	-2.25	7.75	-0.29	-0.23	-0.01
10	-2.5	7.5	-0.33	-0.25	-0.01
10	-2.75	7.25	-0.38	-0.28	-0.01
10	-3	7	-0.43	-0.3	-0.01
10	-3.25	6.75	-0.48	-0.33	-0.01
10	-3.5	6.5	-0.54	-0.35	-0.01
10	-3.75	6.25	-0.60	-0.38	-0.02
10	-4	6	-0.67	-0.4	-0.02
10	-4.25	5.75	-0.74	-0.43	-0.02
10	-4.5	5.5	-0.82	-0.45	-0.02
10	-4.75	5.25	-0.90	-0.48	-0.02
10	-5	5	-1.00	-0.5	-0.03
10	-5.25	4.75	-1.11	-0.53	-0.03
10	-5.5	4.5	-1.22	-0.55	-0.03
10	-5.75	4.25	-1.35	-0.58	-0.03
10	-6	4	-1.50	-0.6	-0.04
10	-6.25	3.75	-1.67	-0.63	-0.04
10	-6.5	3.5	-1.86	-0.65	-0.05
10	-6.75	3.25	-2.08	-0.68	-0.05
10	-7	3	-2.33	-0.7	-0.06
10	-7.25	2.75	-2.64	-0.73	-0.07
10	-7.5	2.5	-3.00	-0.75	-0.08
10	-7.75	2.25	-3.44	-0.78	-0.09
10	-8	2	-4.00	-0.8	-0.10
10	-8.25	1.75	-4.71	-0.83	-0.12
10	-8.5	1.5	-5.67	-0.85	-0.15
10	-8.75	1.25	-7.00	-0.88	-0.18
10	-9	1	-9.00	-0.9	-0.23
10	-9.25	0.75	-12.33	-0.93	-0.32
10	-9.5	0.5	-19.00	-0.95	-0.49

10	-9.75	0.25	-39.00	-0.98	-1.00

TTO=time trade-off value

Table 16 Definitions of explanatory variables used in regression analyses

Variable	Definition
2017	1 if 2017 respondent; 0 if 2002 respondent
Age1	1 if 18-35 years old; 0 otherwise
Age2	1 if ≥55 years old; 0 otherwise
Female	1 if female sex; 0 if male
Education	1 if greater than high school education; 0 otherwise
Condition1	1 if self-reported 1 chronic condition; 0 otherwise
Condition2	1 if self-reported ≥ 2 chronic conditions; 0 otherwise
Race1	1 if self-reported as Black or African-American
Race2	1 if self-reported as neither Black/African-American or White/Caucasian
Ethnicity	1 if Hispanic/Latino/Spanish origin; 0 otherwise
EQ-VAS	Self-rated health as continuous variable (0 to 100)
$\mathbf{MO}_{\text{level-2}}$	1 if level 2 for Mobility; 0 otherwise
$\mathbf{MO}_{\text{level-3}}$	1 if level 3 for Mobility; 0 otherwise
$\mathbf{SC}_{\text{level-2}}$	1 if level 2 for Self-care; 0 otherwise
$\mathbf{SC}_{\text{level-3}}$	1 if level 3 for Self-care; 0 otherwise
$\mathbf{UA}_{\text{level-2}}$	1 if level 2 for Usual Activities; 0 otherwise
$\mathbf{UA}_{\text{level-3}}$	1 if level 3 for Usual Activities; 0 otherwise
$\mathbf{PD}_{\text{level-2}}$	1 if level 2 for Pain/Discomfort; 0 otherwise
$\mathbf{PD}_{\text{level-3}}$	1 if level 3 for Pain/Discomfort; 0 otherwise
$\mathbf{AD}_{\text{level-2}}$	1 if level 2 for Anxiety/Depression; 0 otherwise
$\mathbf{AD}_{\text{level-3}}$	1 if level 3 for Anxiety/Depression; 0 otherwise

Table 17 Respondent characteristics by time period

	2002, n=3773			2017 (n=1062)	
Characteristic	**n**	**% (unweighted)**	**% (weighted)**	**n**	**%**
Age, mean (SE)		42.9 (0.27)	44.5 (0.28)		46.5 (0.55)
18-34 years old	1348	35.7	31.7	347	32.7
35-54 years old	1570	41.6	42.4	365	34.4
55+ years old	855	22.7	26.0	360	32.7
Female	2179	57.8	52	542	51.0
Race					
White	1457	38.6	71.9	762	62.2
Black	1055	28.0	11.0	125	12.1
Other	1261	33.4	17.1	174	25.7
Ethnicity, Hispanic	1115	29.6	11.9	191	18.0
Education, secondary or less	2153	57.4	52.6	354	34.0
History of illness					
None	1365	40.1	34.2	400	37.7
1 condition only	783	23.0	24.2	282	26.6
>/= 2 conditions	1257	36.9	41.5	379	35.7
Self-rated health (EQ-VAS)					
Mean (SE)	83.6	0.3	0.3	80.5	0.5
Median (IQR)	90.0	16.0	16.0	85.0	15.0

SE=standard error; IQR=interquartile range; EQ-VAS=EuroQol Visual Analogue Scale

Table 18 Observed mean valuations of 16 EQ-5D-3L health states in the 2002 and 2017 samples using better-than-dead observations only

		2017 (BTD-only)			2002 (BTD-only)			
Health state profile	Misery score	n	mean [A]	SE	n	mean [B]	SE	[A] - [B]
11112	6	101	0.88	0.02	1665	0.85	0.01	0.03
11121	6	111	0.91	0.01	1213	0.89	0.01	0.02
11211	6	113	0.95	0.01	1709	0.87	0.01	0.08
12111	6	100	0.93	0.01	1211	0.85	0.01	0.08
21111	6	112	0.92	0.02	1671	0.88	0.01	0.04
11113	**7**	**89**	**0.78**	**0.03**	**735**	**0.68**	**0.01**	**0.10**[a]
11122	7	96	0.86	0.02	805	0.79	0.01	0.07
22121	8	102	0.78	0.02	1175	0.78	0.01	0.00
12222	9	105	0.75	0.03	1131	0.72	0.01	0.02
21133	10	75	0.58	0.04	555	0.51	0.01	0.07
22222	10	104	0.72	0.03	755	0.68	0.01	0.04
21323	11	84	0.57	0.03	661	0.56	0.01	0.02
22233	12	67	0.56	0.04	520	0.48	0.02	0.08
32232	12	62	0.36	0.04	467	0.42	0.02	-0.06
33232	13	51	0.33	0.04	388	0.35	0.02	-0.03
33333	15	41	0.31	0.05	1209	0.18	0.01	0.13

[a] Adjusted $p < 0.05$

SE=standard error; WTD = worse-than-dead; misery score = sum of each digit of health state profile

Table 19 Observed mean valuations of 16 EQ-5D-3L health states in the 2002 (linear and non-linear transformations) and 2017 samples

		2017				2002 (linear)					2002 (non-linear)				
Health state profile	**Misery score**	**n**	**mean [A]**	**SE**	**% WTD**	**n**	**mean [B]**	**SE**	**% WTD**	**[A] - [B]**	**n**	**mean [C]**	**SE**	**% WTD**	**[A] - [C]**
11112	6	106	0.79	0.04	4.7	1694	0.83	0.01	1.7	-0.04	1694	0.83	0.01	1.7	-0.03
11121	6	111	0.91	0.01	0.0	1227	0.88	0.01	1.1	0.03	1227	0.88	0.01	1.1	0.03
11211	6	115	0.92	0.03	1.7	1717	0.87	0.01	0.5	0.05	1717	0.87	0.02	0.5	0.05
12111	6	100	0.93	0.01	0.8	1223	0.84	0.01	0.0	0.09	1223	0.84	0.02	0.0	0.09
21111	6	112	0.92	0.02	0.8	1685	0.87	0.01	0.0	0.05	1685	0.87	0.02	0.0	0.05
11113	7	106	0.51	0.07	16.0	852	0.56	0.01	13.7	-0.04	852	0.50	0.02	13.7	0.02
11122	7	99	0.81	0.04	3.0	833	0.76	0.01	3.4	0.04	833	0.75	0.01	3.4	0.05
22121	8	105	0.74	0.03	2.9	1232	0.74	0.01	4.6	-0.01	1232	0.73	0.01	4.6	0.01
12222	9	110	0.67	0.04	4.6	1223	0.66	0.01	7.5	0.01	1223	0.64	0.01	7.5	0.04
21133	10	106	0.14	0.07	29.3	852	0.28	0.02	34.9	-0.14	852	0.14	0.01	34.9	0.00
22222	10	111	0.61	0.05	6.3	833	0.60	0.01	9.4	0.02	833	0.57	0.01	9.4	0.04
21323	11	110	0.24	0.07	23.6	**856**	**0.39**	**0.01**	**22.8**	**-0.15**[a]	856	0.30	0.02	22.8	-0.06
22233	12	100	0.08	0.06	33.0	**833**	**0.20**	**0.01**	**37.6**	**-0.12**[a]	833	0.05	0.01	37.6	0.03
32232	12	110	-0.15	0.07	43.6	**856**	**0.15**	**0.02**	**45.4**	**-0.30**[a]	856	-0.06	0.02	45.4	-0.09
33232	13	99	-0.25	0.06	48.5	**856**	**0.06**	**0.01**	**54.7**	**-0.31**[a]	856	-0.18	0.02	54.7	-0.08
33333	15	111	-0.42	0.06	63.1	**3773**	**-0.10**	**0.01**	**68.0**	**-0.32**[a]	3773	-0.38	0.02	68.0	-0.05

[a] Adjusted $p < 0.05$

WTD = worse-than-dead; misery score = sum of each digit of health state profile

Table 20 Summary of adjusted main effects models of time trade-off values using better-than-dead values only, worse-than-dead values only, and all values, (linear and non-linear transformations)

	BTD-only		WTD-only				All values			
			Linear		Non-linear		Linear		Non-linear	
Parameter[a]	Estimate	SE	Estimate	S	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	0.90	0.02	-0.04	0.03	-0.34	0.03	0.91	0.03	0.91	0.04
2017	**0.05[b]**	**0.01**	**-0.71[b]**	**0.01**	**-0.34[b]**	**0.01**	**-0.11[b]**	**-0.01**	-0.01	0.01

[a] All estimates were adjusted for age, sex, race, ethnicity, education-level, self-reported number of chronic conditions, and EQ-VAS.

[b] Adjusted $p < 0.05$

BTD=better-than-dead; WTD = worse-than-dead; SE=standard error.

Table 21 Full model results for adjusted main effects models of time trade-off values using all values, better-than-dead values only, and worse-than-dead values only (linear and non-linear transformations)

	BTD-only		WTD-only				All values			
			Linear		Non-linear		Linear		Non-linear	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	**0.90[a]**	**0.02**	-0.04	0.03	**-0.34[a]**	**0.03**	**0.91[a]**	**0.03**	**0.91[a]**	**0.04**
2017	**0.05[a]**	**0.01**	**-0.71[a]**	**0.01**	**-0.34[a]**	**0.01**	**-0.11[a]**	**0.01**	-0.01	0.01
$MO_{level-2}$	**-0.03[a]**	**0.00**	0.01	0.01	0.01	0.01	**-0.04[a]**	**0.00**	**-0.05[a]**	**0.00**
$MO_{level-3}$	**-0.19[a]**	**0.00**	**-0.03[a]**	**0.01**	**-0.05[a]**	**0.01**	**-0.27[a]**	**0.01**	**-0.35[a]**	**0.01**
$SC_{level-2}$	**-0.04[a]**	**0.00**	0.00	0.01	0.01	0.01	**-0.06[a]**	**0.00**	**-0.06[a]**	**0.01**
$SC_{level-3}$	**-0.14[a]**	**0.00**	**-0.03[a]**	**0.01**	**-0.03[a]**	**0.01**	**-0.18[a]**	**0.01**	**-0.22[a]**	**0.01**
$UA_{level-2}$	**-0.03[a]**	**0.00**	-0.01	0.01	**-0.02[a]**	**0.01**	**-0.06[a]**	**0.00**	**-0.07[a]**	**0.01**
$UA_{level-3}$	**-0.11[a]**	**0.00**	**-0.03[a]**	**0.01**	**-0.02[a]**	**0.01**	**-0.15[a]**	**0.00**	**-0.18[a]**	**0.01**
$PD_{level-2}$	**-0.04[a]**	**0.00**	0.00	0.01	0.00	0.01	**-0.05[a]**	**0.00**	**-0.06[a]**	**0.01**
$PD_{level-3}$	**-0.18[a]**	**0.00**	**-0.05[a]**	**0.01**	**-0.06[a]**	**0.01**	**-0.28[a]**	**0.00**	**-0.35[a]**	**0.01**
$AD_{level-2}$	**-0.04[a]**	**0.00**	0.00	0.01	0.01	0.01	**-0.05[a]**	**0.00**	**-0.05[a]**	**0.01**
$AD_{level-3}$	**-0.14[a]**	**0.00**	**-0.04[a]**	**0.01**	**-0.04[a]**	**0.01**	**-0.19[a]**	**0.00**	**-0.23[a]**	**0.01**
Age, 18-34	**-0.04[a]**	**0.01**	**0.06[a]**	**0.01**	**-0.04[a]**	**0.01**	0.00	0.01	0.00	0.01
Age, >/= 55	**-0.02[a]**	**0.01**	**-0.08[a]**	**0.01**	**-0.11[a]**	**0.01**	**-0.09[a]**	**0.01**	**-0.09[a]**	**0.02**
FEMALE	-0.01	0.01	**-0.02[a]**	**0.01**	**-0.03[a]**	**0.01**	-0.02	0.01	**-0.02[a]**	**0.01**
Black (vs. White)	**0.08[a]**	**0.01**	**-0.04[a]**	**0.01**	0.01	0.01	**0.07[a]**	**0.01**	**0.10[a]**	**0.01**
Other (vs. White)	-0.01	0.01	-0.02	0.02	-0.01	0.02	-0.02	0.02	-0.03	0.02
Education, more than high school	**-0.02[a]**	**0.01**	**0.02[a]**	**0.01**	0.00	0.01	-0.01	0.01	-0.01	0.01
Ethnicity, Hispanic	**0.05[a]**	**0.01**	-0.01	0.02	-0.02	0.02	0.03	0.02	0.03	0.02
1 condition vs. no conditions	0.01	0.01	0.00	0.01	-0.01	0.01	0.01	0.01	0.01	0.01
2 or more conditions vs. no conditions	0.01	0.01	0.02	0.01	0.00	0.01	0.02	0.01	0.02	0.01
VAS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[a] Indicates Bonferroni corrected $p < 0.05$

SE=standard error; BTD=better-than-dead; WTD = worse-than-dead

Table 22 Differences in dimension-level valuations between 2017 and 2002 based on random effects linear regression models that include interactions terms for time-period and EQ-5D-3L health state descriptor.

	All values				**BTD-only**		**WTD-only**			
	Linear		**Non-linear**				**Linear**		**Non-linear**	
Dimension-level comparison[a]	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
$MO_{level\text{-}2}$	**-0.06**[b]	**0.02**	0.03	0.02	0.03	0.01	**-0.77**[b]	**0.03**	**-0.41**[b]	**0.03**
$MO_{level\text{-}3}$	**-0.32**[b]	**0.02**	**-0.15**[b]	**0.02**	0.02	0.01	**-0.71**[b]	**0.02**	**-0.35**[b]	**0.02**
$SC_{level\text{-}2}$	**-0.16**[b]	**0.02**	-0.06	0.02	0.01	0.01	**-0.77**[b]	**0.02**	**-0.40**[b]	**0.02**
$SC_{level\text{-}3}$	**-0.13**[b]	**0.02**	0.01	0.02	**0.04**[b]	**0.01**	**-0.73**[b]	**0.02**	**-0.36**[b]	**0.02**
$UA_{level\text{-}2}$	**-0.10**[b]	**0.02**	0.00	0.02	0.00	0.01	**-0.73**[b]	**0.02**	**-0.38**[b]	**0.03**
$UA_{level\text{-}3}$	**-0.17**[b]	**0.02**	-0.05	0.02	0.00	0.01	**-0.78**[b]	**0.03**	**-0.37**[b]	**0.02**
$PD_{level\text{-}2}$	**-0.13**[b]	**0.02**	-0.04	0.02	0.00	0.01	**-0.76**[b]	**0.03**	**-0.40**[b]	**0.03**
$PD_{level\text{-}3}$	**-0.19**[b]	**0.02**	-0.03	0.02	0.01	0.01	**-0.74**[b]	**0.02**	**-0.34**[b]	**0.02**
$AD_{level\text{-}2}$	**-0.13**[b]	**0.02**	-0.03	0.02	0.02	0.01	**-0.78**[b]	**0.03**	**-0.41**[b]	**0.03**
$AD_{level\text{-}3}$	**-0.22**[b]	**0.02**	**-0.08**[b]	**0.02**	0.00	0.01	**-0.74**[b]	**0.02**	**-0.36**[b]	**0.02**

[a] All estimates were adjusted for age, sex, race, ethnicity, education-level, self-reported number of chronic conditions, and EQ-VAS.

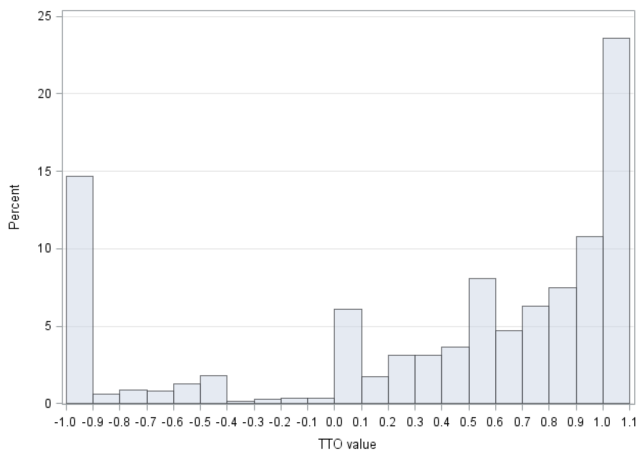
[b] Indicates adjusted $p < 0.05$ (Bonferroni method).

AD=Anxiety/Depression; BTD=better-than-dead; MO=Mobility; PD=Pain/Discomfort; SC=Self-Care; SE = standard error; UA=Usual Activities; WTD =worse-than-dead.

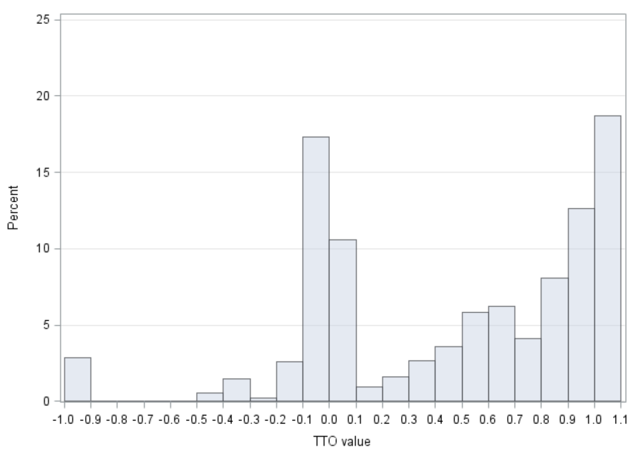
FIGURES

Figures 9 Distribution of all observed time trade-off values for a) 2017 b) 2002, applying a linear transformation for WTD values, and c) 2002, applying non-linear transformations)

a) Observed 2017 time trade-off values



b) Observed 2002 time trade-off values with linear transformation



c) Observed 2002 time trade-off values with non-linear transformation

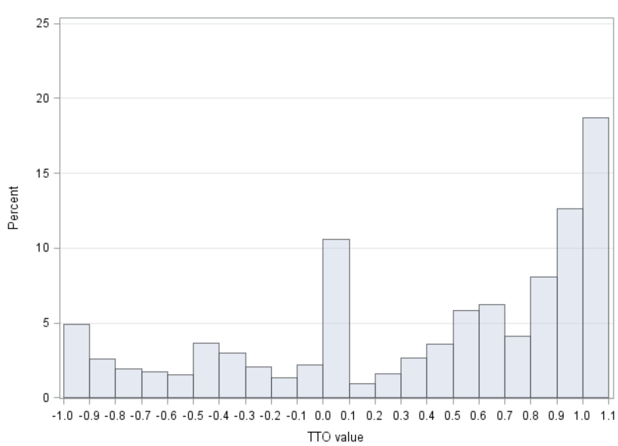
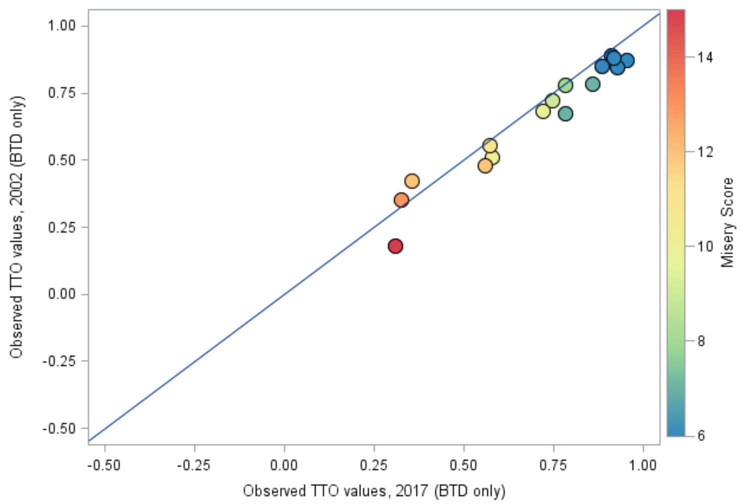


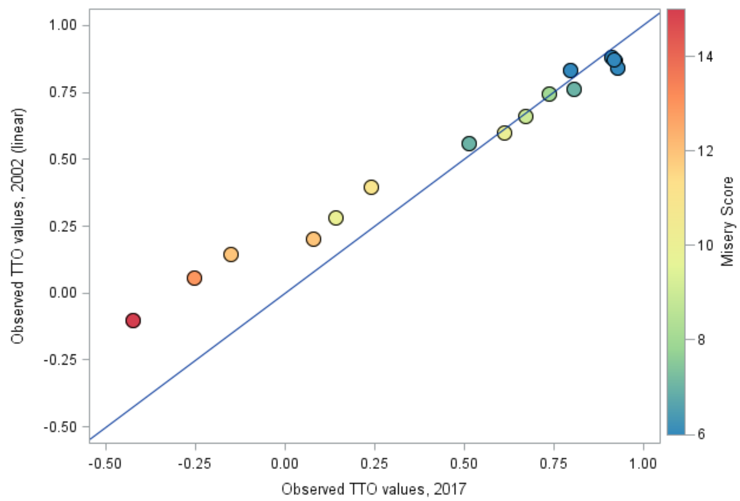
Figure 10 Correlation between the 2017 versus 2002 mean valuations for 16 EQ-5D-3L health states for better-than-dead values only (Spearman's rho=0.94)



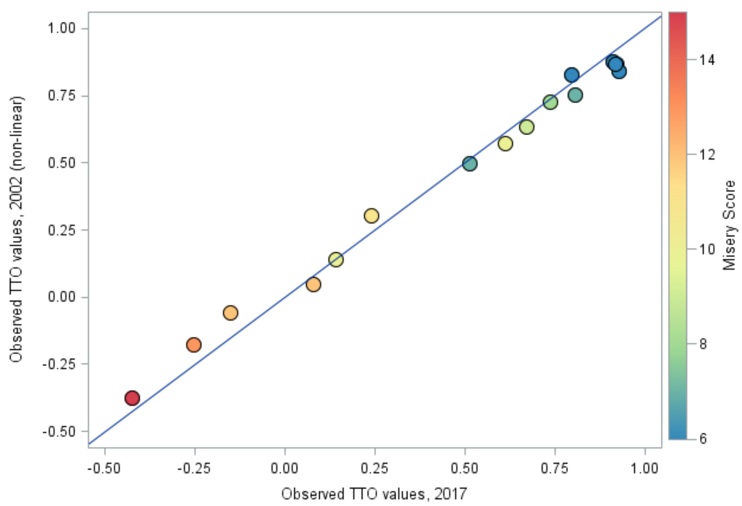
(Note: color gradient represents misery score for each EQ-5D health state vector, where blue indicates lower misery score and red indicates highest; solid diagonal line represents the line of unity)

Figure 11 Correlation between the 2017 versus 2002 for a) linear and b) non-linear transformed mean valuations for 16 EQ-5D-3L health states

a) Linear transformation (Spearman's rho=0.97)



b) Non-linear transformation (Spearman's rho=0.97)



(Note: color gradient represents misery score for each EQ-5D health state vector, where blue indicates lower misery score and red indicates highest; solid diagonal line represents the line of unity)

5. CONCLUSION

5.1 Summary of Research

The overall objective of this dissertation was to examine specific sources of variation in developing societal value sets. To do so, we modified an existing framework that detailed the process that individuals undergo when forming health preferences (**Figure 3**) to reflect the process of capturing societal preferences and developing a health state value set. Sources of variation that were evaluated included the descriptive system, completion of advance directives as a potential conversion factor for health preferences, and the time in which a valuation study was conducted. The major themes that arose from each study and how it relates to the modified conceptual framework (**Figure 12**) are summarized below.

Minor modifications to the wording of an existing descriptive system can result in differences in health state values that have potential implications for decisions around the cost-effectiveness of new treatments. In the first paper (**Chapter 2**), value sets developed for the 5L and 3L versions of the EQ-5D were compared. In our conceptual framework (**Figure 12**), the health state (and its description) is the very first set of stimuli internalized by the respondent before continuing through the process for valuing health states. The most striking difference between the descriptive systems is the labelling for the most severe level of problems in Mobility. For the 3L, it is “confined to bed” and for the 5L it is “unable to walk about”. As expected, our study found a larger range of scale for the 3L compared to 5L, driven largely by the difference in weight placed on the most severe level of problems in Mobility. Furthermore, we found that the 5L produced a value set a reduced ceiling effect relative to the 3L providing evidence that the addition of “slight problems” in the 5L (level 2) provides additional sensitivity to milder health states not observed with the 3L.

Individuals who have prepared an advance directive have different characteristics and tend to value health differently than those without an advance directive. In the second paper (**Chapter 3**), we identified factors associated with advance directive status and examined differences in stated preferences for health between individuals with and without advance directives. We hypothesized that advance directives status would influence health values obtained among individuals in the general population, independent of other respondent characteristics (**Figure 12**). We found that people who completed an advance directive were more likely to be older, White, and have health insurance coverage from Medicare or private sources, which may suggest that there are racial and socioeconomic barriers to accessing end-of-life planning. Advance directive completion was also associated with a greater willingness to trade years of life for quality of life but not necessarily more likely to view health states as worse-than-dead, independent of health state severity and differences in respondent characteristics. Therefore, advance directive status may represent an additional conversion factor that is associated with differences in weighing the consequences of a health state, independent of other respondent characteristics, including age, illness experience, and living arrangements. Further, if advance directives are an indicator of a substantial proportion of the population with “informed” preferences, it has implications not only for intensive care decision making at the individual level, but also resource allocation based on cost-utility analysis at the group level.

Societal values in terms of health preferences may change over time, requiring periodic updates. In the third study (Chapter 4), we assessed whether there were time-specific differences between two eras. Time as a factor that influences health state valuations is unique in that it can be thought of as an additional dimension outside that of the process for developing societal value sets (**Figure 12**) that represents an overall shift in health preferences, irrespective of sociodemographic changes in the population and methodological differences in surveying health state preferences. We found that values are systematically higher in 2017 compared to 2002 in the US. We also showed that methodological differences, such as choice of transformation and preference elicitation technique for states WTD, heavily influenced values obtained for states WTD. Overall, our findings suggest that the era in which values were elicited may be an important reason to consider updating societal value sets.

5.2 General discussion

From a broader perspective, this dissertation sought to inform policy makers in understanding issues that may impact the selection and development of societal value sets for the purpose of informing health care decision-making. The following section discusses the implications of each study for future research and policy.

5.2.1 Transitioning from the 3L to the 5L value set

The first study (**Chapter 2**) compared the 3L and 5L versions of the EQ-5D and is directly relevant to HTA bodies considering the transition from the 3L to the 5L. For instance, NICE has issued a position statement recommending that the UK MVH 3L value set (Dolan, 1997) be used for reference-case analyses over the recently published 5L value set for England for the time-being until additional research is available that evaluates “the impact of adopting the EQ-5D-5L valuation set in the NICE reference-case” (N. J. Devlin et al., 2018; NICE, 2017). Moreover, in situations where 5L data was collected, 3L values should be calculated by using the van Hout et al. scoring algorithm (van Hout et al., 2012). NICE is expected to review this decision in August 2018 (N. Devlin et al., 2018).

To date, comparative studies in the recent literature that may be used by HTA agencies to understand differences between the 3L and 5L are limited. These studies tend to focus on existing value sets which differ in many important ways (e.g., different respondents in different time periods, dissimilar study protocols/methods, and statistical approaches). In contrast, our work provides evidence of the impact of the descriptive system on health values by employing a study design that holds all other respondent- and protocol-related factors constant. In this study, the same respondents valued 3L and 5L health states using the same interviewer, study protocol, preference elicitation technique, and time point. Moreover, the same statistical models were used in the development of both value sets. Therefore, our study represents a powerful addition to the literature and will inform HTA agencies, such as NICE, on the implications of potentially adopting one preference-based measure over the other, particularly if new value sets are generated. The results of our study support the 5L over the 3L in most applications, given the reduced ceiling effects and improved measurement properties. The one exception may be for conditions/diseases where being confined to bed is an expected outcome. In these situations, the 3L, relative to the 5L, may actually be more sensitive to the disutility experienced by patients. Finally, future research should consider using a similar study design to facilitate internally valid investigations into how slight changes to a preference-based measure’s descriptive system may impact subsequent value sets.

5.2.2 Advance directives and representativeness of societal values

As discussed in **Chapter 3**, the second study is directly relevant to understanding how individuals in the general public who have completed an advance directive may be systematically different from those without a directive with respect to sociodemographic and attitudinal factors, as well as their health preferences. Future research is required to understand the relationship between advance directive status, level of informedness of individual health preferences, and other conversion factors, such as serious illness experience and social situation (e.g., marital status, living arrangements, and having children).

This study is also important from the perspective of a researcher undertaking the development of a societal value set or a user/policy-maker interpreting and selecting an appropriate societal value set. This study contributes to the body of literature identifying respondent characteristics that may be associated with differences in health values. Specifically, we found meaningful differences in health values by advance directive status, overall health and within specific dimension-levels, even after accounting for known factors associated with differences in health values (e.g., age, race, and ethnicity). Furthermore, these differences translated to small but statistically significant differences in transition values across all health states. This suggests larger QALY gains/loss when utilizing the preferences of individuals with advance directives, which in turn implies lower cost per QALY estimates, given the same incremental costs.

Therefore, given the sizeable prevalence of advance directive completion in the US (~25%) , the aging US population (Ortman et al., 2014) increasing rate of advance directive completion (Silveira, Wiitala, & Piette, 2014), debate over the ethics of end-of-life care (e.g., physician-assisted suicide) (Lugo et al., 2013), and the magnitude of the differences found in the present study, it may be important to consider advance directive status as a criterion for sampling and recruitment. At minimum, it appears useful for investigators to collect data on advance directive status in valuation studies to better aid in the interpretation of the results.

5.2.3 When to re-visit/update existing societal value sets

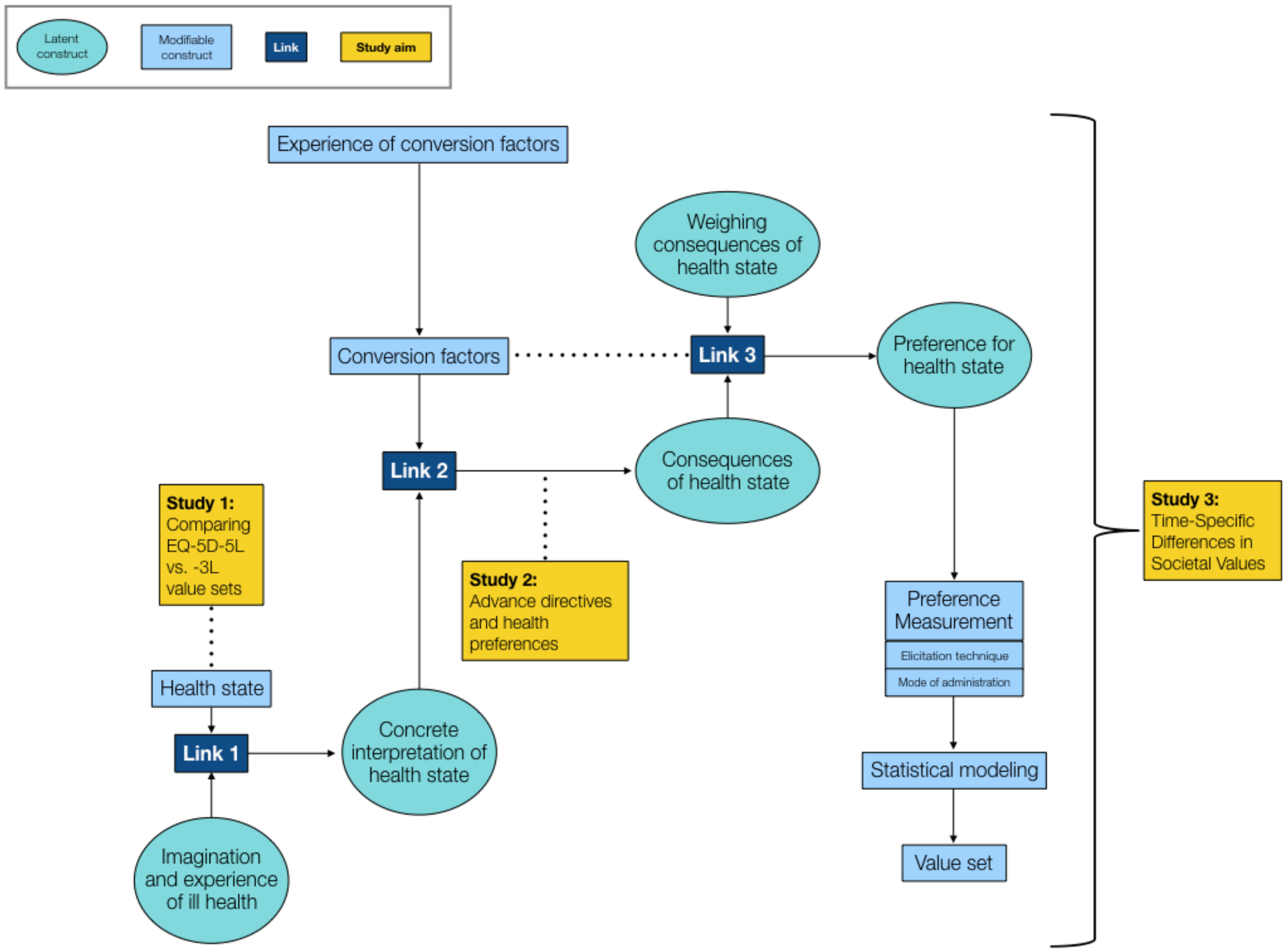
The third study (**Chapter 4**) examined whether value sets are dated to ensure adequate representation of current societal health preferences. This is a particularly relevant question as some value sets were derived decades ago. For example, in the United Kingdom, the UK MVH 3L value set required by NICE was developed based on preferences collected over two decades ago (Dolan, 1997). The value set for the HUI was developed from preferences collected among residents in Hamilton, Ontario over 25 years ago (David Feeny et al., 2002). In our study, we examined differences in health values obtained 15 years apart (A. Pickard et al., 2018; J. W. Shaw et al., 2005). While the differences we found were modest, they approximated several minimally important differences derived from previous work.

Our work has direct implications for jurisdictions that rely on metrics such as health utilities/QALYs to inform health care decision making. In several such jurisdictions, guidance recommends or mandates that such metrics should adequately reflect the preferences of the target population (CADTH, 2017; ICER, 2017; PBAC, 2016). Given our results showing evidence of time-specific differences in societal preferences, there should be a heightened awareness that societal values may change over time. Therefore, future research may explore strategies to account for such changes, such as an inflation or discount factor used to adjust for time-specific differences in health utilities. Additional studies are also needed to evaluate the precise amount of time that is required to pass before societal values may "expire" and require an update. Our findings have further implications for informing HTA guidance on selecting value sets for the reference-case analyses, in which some HTA bodies may consider recommendations that health value inputs in CUAs be no older than a certain number of years. Our study results provide evidence of meaningful shifts in societal values as early as 15 years in the US. However, the specific amount of time that HTA agencies may allow to pass should be driven by country- or jurisdiction-specific data, but our present work represents the first of its kind for the US. Future research may also explore whether time-specific differences exist for other preference-based measures and, if so, how comparable the differences are to the results of our study to help further refine recommendations on selecting value sets for CUAs.

5.3 Conclusion

This dissertation has highlighted several important sources of variation that may influence values obtained from individuals within a society. Small changes to the descriptive system of a preference-based measure can lead to dramatic differences in the health values obtained. Advance directive status is associated with differences in health preferences, independent of a health state's severity and conversion factors known to impact health preferences. Finally, the passing of time itself is associated with differences in societal preferences. This dissertation advances our understanding of how societal values can vary through the sources studied. This work will be important for informing health technology assessment, health care resource allocation decisions, and understanding how specific groups in society value health.

Figure 12 Modified framework outlining the process by which societal value sets are developed using individual health preferences



(Adapted from Karimi et al. 2017)

CITED LITERATURE

Ahlstrom, A., Behm, A., Blanchette, C. M., Brusig, E. L., Chavala, S., Daniel, G., . . . Francer, J. K. (2017). AMCP Partnership Forum: Enabling the Exchange of Clinical and Economic Information Pre-FDA Approval. *JOURNAL OF MANAGED CARE & SPECIALTY PHARMACY, 23*(1), 105-112.

Al Sayah, F., Mladenovic, A., Gaebel, K., Xie, F., & Johnson, J. A. (2016). How dead is dead? Qualitative findings from participants of combined traditional and lead-time time trade-off valuations. *Quality of life research, 25*(1), 35-43.

Alava, M. H., Wailoo, A., Grimm, S., Pudney, S., Gomes, M., Sadique, Z., . . . Irvine, L. (2017). EQ-5D-5L versus EQ-5D-3L: The Impact on Cost-Effectiveness in the United Kingdom. *Value in Health*.

AMCP. (2010). The AMCP format for formulary submissions version 3.0. *Journal of Managed Care Pharmacy, 16*(1 Supp A), 1-32.

Anderson, J. L., Heidenreich, P. A., Barnett, P. G., Creager, M. A., Fonarow, G. C., Gibbons, R. J., . . . Mark, D. B. (2014). ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures. *Circulation, 129*(22), 2329-2345.

Arnold, D., Girling, A., Stevens, A., & Lilford, R. (2009). Comparison of direct and indirect methods of estimating health state utilities for resource allocation: review and empirical analysis. *Bmj, 339*, b2688.

Attema, A. E., Edelaar-Peeters, Y., Versteegh, M. M., & Stolk, E. A. (2013). Time trade-off: one methodology, different methods. *Eur J Health Econ, 14 Suppl 1*, S53-64. doi:10.1007/s10198-013-0508-x

Augestad, L. A., Rand-Hendriksen, K., Kristiansen, I. S., & Stavem, K. (2012). Learning effects in time trade-off based valuation of EQ-5D health states. *Value in Health, 15*(2), 340-345.

Augestad, L. A., Rand-Hendriksen, K., Stavem, K., & Kristiansen, I. S. (2013). Time trade-off and attitudes toward euthanasia: implications of using 'death' as an anchor in health state valuation. *Qual Life Res, 22*(4), 705-714. doi:10.1007/s11136-012-0192-9

Augustovski, F., Rey-Ares, L., Irazola, V., Garay, O. U., Gianneo, O., Fernandez, G., . . . Ramos-Goni, J. M. (2016). An EQ-5D-5L value set based on Uruguayan population preferences. *Qual Life Res, 25*(2), 323-333. doi:10.1007/s11136-015-1086-4

Auriemma, C. L., Nguyen, C. A., Bronheim, R., Kent, S., Nadiger, S., Pardo, D., & Halpern, S. D. (2014). Stability of end-of-life preferences: a systematic review of the evidence. *JAMA internal medicine, 174*(7), 1085-1092.

Barrio-Cantalejo, I. M., Simón-Lorda, P., Molina-Ruiz, A., Herrera-Ramos, F., Martínez-Cruz, E., Bailon-Gómez, R. M., . . . Gorlat, P. P. (2013). Stability over time in the preferences of older persons for life-sustaining treatment. *Journal of bioethical inquiry, 10*(1), 103-114.

Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: care, health, and cost. *Health affairs, 27*(3), 759-769.

Bleichrodt, H. (2002). A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Economics, 11*(5), 447-456.

Blumenthal, D., Abrams, M., & Nuzum, R. (2015). The affordable care act at 5 years. In: Mass Medical Soc.

Brazier, J., Ara, R., Rowen, D., & Chevrou-Severac, H. (2017). A review of generic preference-based measures for use in cost-effectiveness models. *PharmacoEconomics, 35*(1), 21-31.

Brazier, J., Ratcliffe, J., Saloman, J., & Tsuchiya, A. (2017). *Measuring and valuing health*

benefits for economic evaluation: OXFORD university press.
Brinkman-Stoppelenburg, A., Rietjens, J. A., & van der Heide, A. (2014). The effects of advance care planning on end-of-life care: a systematic review. *Palliative medicine, 28*(8), 1000-1025.
CADTH. (2017). *Guidelines for the economic evaluation of health technologies: Canada*. Retrieved from Ottawa:
Craig, B. M., Reeve, B. B., Cella, D., Hays, R. D., Pickard, A. S., & Revicki, D. A. (2014). Demographic differences in health preferences in the United States. *Med Care, 52*(4), 307.
Devlin, N., Brazier, J., Pickard, A. S., & Stolk, E. (2018). 3L, 5L, What the L? A NICE Conundrum. *PharmacoEconomics*, 1-4.
Devlin, N. J., Shah, K. K., Feng, Y., Mulhern, B., & van Hout, B. (2018). Valuing health-related quality of life: An EQ-5 D-5 L value set for England. *Health Economics, 27*(1), 7-22.
Doctor, J. N., Bleichrodt, H., & Lin, H. J. (2010). Health utility bias: a systematic review and meta-analytic evaluation. *Medical Decision Making, 30*(1), 58-67.
Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35*(11), 1095-1108.
Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*: Oxford university press.
Emanuel, L. L. (1995). Advance directives: do they work? *Journal of the American College of Cardiology, 25*(1), 35-38.
EQ-5D Instruments | About EQ-5D. (28 April 2017). Retrieved from https://euroqol.org/eq-5d-instruments/
Feeny, D., Furlong, W., Saigal, S., & Sun, J. (2004). Comparing directly measured standard gamble scores to HUI2 and HUI3 utility scores: Group- and individual-level comparisons. *Social Science and Medicine, 58*(4), 799-809.
Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., . . . Boyle, M. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care, 40*(2), 113-128.
Feng, Y., Devlin, N., & Herdman, M. (2015). Assessing the health of the general population in England: how do the three-and five-level versions of EQ-5D compare? *Health and Quality of Life Outcomes, 13*(1), 171.
Feng, Y., Herdman, M., van Nooten, F., Cleeland, C., Parkin, D., Ikeda, S., . . . Devlin, N. J. (2017). 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*, 1-12.
Gold, M. R. (1996). *Cost-effectiveness in health and medicine*: Oxford university press.
Greiner, W., Weijnen, T., Nieuwenhuizen, M., Oppe, S., Badia, X., Busschbach, J., . . . Krabbe, P. (2003). A single European currency for EQ-5D health states. *The European Journal of Health Economics, formerly: HEPAC, 4*(3), 222-231.
Harsanyi, J. C. (1996). Utilities, preferences, and substantive goods. *Social choice and welfare, 14*(1), 129-145.
Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., . . . Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727-1736.
Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., . . . Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).

Qual Life Res, 20(10), 1727-1736. doi:10.1007/s11136-011-9903-x

Hernández-Alava, M., & Pudney, S. (2017). Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. *Journal of Health Economics*.

Hobbs, F., & Stoops, N. (2002). US Census Bureau, Census 2000 Special Reports, Series CENSR-4. *Demographic trends in the 20th century, 24*, 27-28.

Huang, I. A., Neuhaus, J. M., & Chiong, W. (2016). Racial and ethnic differences in advance directive possession: role of demographic factors, religious affiliation, and personal health values in a National Survey of older adults. *Journal of palliative medicine, 19*(2), 149-156.

Humes, K. R., Jones, N. A., & Ramirez, R. R. (2011). *Overview of race and Hispanic origin: 2010*: US Department of Commerce, Economics and Statistics Administration, US Census Bureau Washington, DC.

Husereau, D., Culyer, A. J., Neumann, P., & Jacobs, P. (2015). How do economic evaluations inform health policy decisions for treatment and prevention in Canada and the United States? *Applied Health Economics and Health Policy, 13*(3), 273-279.

ICER. (2017). *Overview of the ICER value framework and proposals for an update for 2017-2018*. Retrieved from

Jakubczyk, M., Golicki, D., & Niewada, M. (2016). The impact of a belief in life after death on health-state preferences: True difference or artifact? *Quality of Life Research, 25*(12), 2997-3008.

Janssen, B. M., Oppe, M., Versteegh, M. M., & Stolk, E. A. (2013). Introducing the composite time trade-off: a test of feasibility and face validity. *The European Journal of Health Economics, 14*(1), 5-13.

Janssen, M., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., . . . Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research, 22*(7), 1717-1727.

Janssen, M. F., Bonsel, G. J., & Luo, N. (2018). Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *PharmacoEconomics*, 1-23.

Johnson, F. R. (2009). Moving the QALY forward or just stuck in traffic? *Value in Health, 12*(s1).

Johnson, J. A., Luo, N., Shaw, J. W., Kind, P., & Coons, S. J. (2005). Valuations of EQ-5D health states: are the United States and United Kingdom different? *Medical Care, 43*(3), 221-228.

Judy Campbell, M., Jo Edwards, M., Ward, K. S., & Weatherby, N. (2007). Developing a parsimonious model for predicting completion of advance directives. *Journal of Nursing Scholarship, 39*(2), 165-171.

Karimi, M., Brazier, J., & Paisley, S. (2017). Are preferences over health states informed? *Health Qual Life Outcomes, 15*(1), 105. doi:10.1186/s12955-017-0678-9

Karimi, M., Brazier, J., & Paisley, S. (2017). How do individuals value health states? A qualitative investigation. *Social Science & Medicine, 172*, 80-88.

Kim, S. H., Ahn, J., Ock, M., Shin, S., Park, J., Luo, N., & Jo, M. W. (2016). The EQ-5D-5L valuation study in Korea. *Qual Life Res*. doi:10.1007/s11136-015-1205-2

Kind, P., & Dolan, P. (1995). The effect of past and present illness experience on the valuations of health states. *Medical Care*, AS255-AS263.

King Jr, J. T., Styn, M. A., Tsevat, J., & Roberts, M. S. (2003). “Perfect health” versus “disease free”: the impact of anchor point choice on the measurement of preferences and the calculation of disease-specific disutilities. *Medical Decision Making, 23*(3), 212-225.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine, 15*(2), 155-163.

Koss, C. S., & Baker, T. A. (2017). Race differences in advance directive completion: the narrowing gap between White and African American older adults. *Journal of aging and health, 29*(2), 324-342.

Krabbe, P. F., Devlin, N. J., Stolk, E. A., Shah, K. K., Oppe, M., van Hout, B., . . . Xie, F. (2014). Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-state values. *Medical Care, 52*(11), 935-943.

Lamers, L. M., McDonnell, J., Stalmeier, P. F., Krabbe, P. F., & Busschbach, J. J. (2006). The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics, 15*(10), 1121-1132.

Lancsar, E., & Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *PharmacoEconomics, 26*(8), 661-677.

Le, Q. A., Doctor, J. N., Zoellner, L. A., & Feeny, N. C. (2013). Minimal clinically important differences for the EQ-5D and QWB-SA in Post-traumatic Stress Disorder (PTSD): results from a Doubly Randomized Preference Trial (DRPT). *Health and Quality of Life Outcomes, 11*(1), 59.

Lipscomb, J., Drummond, M., Fryback, D., Gold, M., & Revicki, D. (2009). Retaining, and enhancing, the QALY. *Value in Health, 12*, S18-S26.

Ludwig, K., Graf von der Schulenburg, J. M., & Greiner, W. (2018). German Value Set for the EQ-5D-5L. *PharmacoEconomics*. doi:https://doi.org/10.1007/s40273-018-0615-8

Lugo, L., Cooperman, A., Funk, G., & O'Connell, E. (2013). *Views on end-of-life medical treatments*. Retrieved from Pew Research Center:

Luo, N., Johnson, J. A., Shaw, J. W., Feeny, D., & Coons, S. J. (2005). Self-reported health status of the general adult US population as assessed by the EQ-5D and Health Utilities Index. *Medical Care, 43*(11), 1078-1086.

McCabe, C., Claxton, K., & Culyer, A. J. (2008). The NICE cost-effectiveness threshold. *PharmacoEconomics, 26*(9), 733-744.

McClure, N. S., Al Sayah, F., Xie, F., Luo, N., & Johnson, J. A. (2017). Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. *Value in Health, 20*(4), 644-650.

McDonough, C. M., & Tosteson, A. N. A. (2007). Measuring Preferences for Cost-Utility Analysis: How Choice of Method May Influence Decision-Making. *PharmacoEconomics, 25*(2), 93-106.

Mulhern, B., Feng, Y., Shah, K., Janssen, M. F., Herdman, M., van Hout, B., & Devlin, N. (2018). Comparing the UK EQ-5D-3L and English EQ-5D-5L Value Sets. *PharmacoEconomics, 36*(6), 699-713.

Neumann, P., & Weinstein, M. (2010). Legislating against use of cost-effectiveness information. *New England Journal of Medicine, 363*(16), 1495-1497.

Neumann, P. J., Sanders, G. D., Russell, L. B., Siegel, J. E., & Ganiats, T. G. (2016). *Cost-effectiveness in health and medicine*: Oxford University Press.

Neumann, P. J., Thorat, T., Shi, J., Saret, C. J., & Cohen, J. T. (2015). The changing face of the cost-utility literature, 1990-2012. *Value Health, 18*(2), 271-277. doi:10.1016/j.jval.2014.12.002

NICE. (2013). *Guide to the methods of technology appraisal 2013*. Retrieved from

NICE. (2017). NICE position statement on the EQ-5D-5L. Retrieved from https://euroqol.org/nice-position-statement-on-the-eq-5d-5l/

Nord, E., Daniels, N., & Kamlet, M. (2009). QALYs: some challenges. *Value in Health, 12*, S10-

S15.
Olsen, J. A., Lamu, A. N., & Cairns, J. (2017). In search of a common currency: A comparison of seven EQ-5D-5L value sets. *Health Economics*.
Oppe, M., Devlin, N. J., van Hout, B., Krabbe, P. F., & de Charro, F. (2014). A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health, 17*(4), 445-453. doi:10.1016/j.jval.2014.04.002
Ortman, J. M., Velkoff, V. A., & Hogan, H. (2014). *An aging nation: the older population in the United States*: United States Census Bureau, Economics and Statistics Administration, US Department of Commerce.
Patrick, D. L., Pearlman, R. A., Starks, H. E., Cain, K. C., Cole, W. G., & Uhlmann, R. F. (1997). Validation of preferences for life-sustaining treatment: implications for advance care planning. *Annals of Internal Medicine, 127*(7), 509-517.
Patrick, D. L., Starks, H. E., Cain, K. C., Uhlmann, R. F., & Pearlman, R. A. (1994). Measuring preferences for health states worse than death. *Medical Decision Making, 14*(1), 9-18.
PBAC. (2016). *Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee (Version 5.0)*. Retrieved from
Pew. (2016). *The State of American Jobs*. Retrieved from Washington, District of Columbia:
Pickard, A., Law, E., Jiang, R., Oppe, M., Shaw, J., Xie, F., . . . Bussbach, J. (2018). A United States Valuation of EQ-5D-5L Health States Using an International Protocol. *Medical Decision Making*, (submitted).
Pickard, A. S. (2015). Is it Time to Update Societal Value Sets for Preference-Based Measures of Health? *PharmacoEconomics, 33*(3), 191-192.
Pickard, A. S., Kohlmann, T., Janssen, M. F., Bonsel, G., Rosenbloom, S., & Cella, D. (2007). Evaluating equivalency between response systems: application of the Rasch model to a 3-level and 5-level EQ-5D. *Medical Care, 45*(9), 812-819.
Pickard, A. S., Neary, M. P., & Cella, D. (2007). Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health and Quality of Life Outcomes, 5*(1), 70.
Prochaska, J. O. (2013). Transtheoretical model of behavior change. In *Encyclopedia of behavioral medicine* (pp. 1997-2000): Springer.
Purba, F. D., Hunfeld, J. A., Iskandarsyah, A., Fitriana, T. S., Sadarjoen, S. S., Ramos-Goñi, J. M., . . . Busschbach, J. J. (2017). The Indonesian EQ-5D-5L value set. *PharmacoEconomics, 35*(11), 1153-1165.
Rabin, R., & Charro, F. d. (2001). EQ-5D: a measure of health status from the EuroQol Group. *Annals of medicine, 33*(5), 337-343.
Rabin, R., Oemar, M., Oppe, M., Janssen, B., & Herdman, M. (2011). EQ-5D-3L user guide. *Basic information on how to use the EQ-5D-5L instrument. Rotterdam: EuroQol Group, 22*.
Rao, J. K., Anderson, L. A., Lin, F.-C., & Laux, J. P. (2014). Completion of advance directives among US consumers. *American journal of preventive medicine, 46*(1), 65-70.
Richardson, J., McKie, J., & Bariola, E. (2011). *Review and critique of health related multi attribute utility instruments*: Monash University, Business and Economics, Centre for Health Economics Melbourne.
Rowen, D., Brazier, J., Tsuchiya, A., Peasgood, T., Ratcliffe, J., & Karimi, M. (2017). Experience-based utility and own health state valuation: why do it and how to do it.
Sanders, G. D., Neumann, P. J., Basu, A., Brock, D. W., Feeny, D., Krahn, M., . . . Prosser, L. A. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. *JAMA, 316*(10), 1093-1103.

Sayah, F. A., Bansback, N., Bryan, S., Ohinmaa, A., Poissant, L., Pullenayegum, E., . . . Johnson, J. A. (2016). Determinants of time trade-off valuations for EQ-5D-5L health states: data from the Canadian EQ-5D-5L valuation study. *Quality of Life Research, 25*(7), 1679-1685.

Schnipper, L. E., Davidson, N. E., Wollins, D. S., Tyne, C., Blayney, D. W., Blum, D., . . . Langdon, R. (2015). American Society of Clinical Oncology statement: a conceptual framework to assess the value of cancer treatment options. *Journal of Clinical Oncology, 33*(23), 2563-2577.

Selivanova, A., Buskens, E., & Krabbe, P. F. (2018). Head-to-Head Comparison of EQ-5D-3L and EQ-5D-5L Health Values. *PharmacoEconomics*, 1-11.

Shah, K., Mulhern, B., Longworth, L., & Janssen, M. F. B. (2016). An empirical study of two alternative comparators for use in time trade-off studies. *Value in Health, 19*(1), 53-59.

Shaw, J. W., Johnson, J. A., Chen, S., Levin, J. R., & Coons, S. J. (2007). Racial/ethnic differences in preferences for the EQ-5D health states: results from the US valuation study. *Journal of clinical epidemiology, 60*(5), 479-490.

Shaw, J. W., Johnson, J. A., & Coons, S. J. (2005). US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care, 43*(3), 203-220.

Shiroiwa, T., Ikeda, S., Noto, S., Igarashi, A., Fukuda, T., Saito, S., & Shimozuma, K. (2016). Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. *Value in Health, 19*(5), 648-654.

Silveira, M. J., Wiitala, W., & Piette, J. (2014). Advance directive completion by elderly Americans: a decade of change. *Journal of the American Geriatrics Society, 62*(4), 706-710.

van Hout, B., Janssen, M., Feng, Y.-S., Kohlmann, T., Busschbach, J., Golicki, D., . . . Pickard, A. S. (2012). Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value in Health, 15*(5), 708-715.

van Reenen, M., & Oppe, M. (2015). EQ-5D-3L user Guide. *EuroQol Res Found, 22*.

Versteegh, M., & Brouwer, W. (2016). Patient and general public preferences for health states: a call to reconsider current guidelines. *Social Science & Medicine, 165*, 66-74.

Walters, S. J., & Brazier, J. E. (2005). Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of Life Research, 14*(6), 1523-1532.

Weinstein, M. C., Torrance, G., & McGuire, A. (2009). QALYs: the basics. *Value in Health, 12*, S5-S9.

Wilkinson, A., Wenger, N., & Shugarman, L. R. (2007). Literature review on advance directives. *US Department of Health and Human Services, Volume June*.

Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., . . . Herdman, M. (2010). Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research, 19*(6), 875-886.

Wisløff, T., Hagen, G., Hamidi, V., Movik, E., Klemp, M., & Olsen, J. A. (2014). Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *PharmacoEconomics, 32*(4), 367-375.

Xie, F., Pickard, A. S., Krabbe, P. F., Revicki, D., Viney, R., Devlin, N., & Feeny, D. (2015). A checklist for reporting valuation studies of multi-attribute utility-based instruments (CREATE). *PharmacoEconomics, 33*(8), 867-877.

Xie, F., Pullenayegum, E., Gaebel, K., Bansback, N., Bryan, S., Ohinmaa, A., . . . Johnson, J. A. (2016). A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada. *Medical Care, 54*(1), 98-105.

Yang, Z., Luo, N., Bonsel, G., Busschbach, J., & Stolk, E. (2017). Selecting Health States for EQ-5D-3L Valuation Studies: Statistical Considerations Matter. *Value in Health*.

Appendix

Appendix - Two versions of the EQ-5D instrument

a) EQ-5D-5L descriptive system

MOBILITY
I have no problems in walking about ☐
I have slight problems in walking about ☐
I have moderate problems in walking about ☐
I have severe problems in walking about ☐
I am unable to walk about ☐

SELF-CARE
I have no problems washing or dressing myself ☐
I have slight problems washing or dressing myself ☐
I have moderate problems washing or dressing myself ☐
I have severe problems washing or dressing myself ☐
I am unable to wash or dress myself ☐

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*
I have no problems doing my usual activities ☐
I have slight problems doing my usual activities ☐
I have moderate problems doing my usual activities ☐
I have severe problems doing my usual activities ☐
I am unable to do my usual activities ☐

PAIN / DISCOMFORT
I have no pain or discomfort ☐
I have slight pain or discomfort ☐
I have moderate pain or discomfort ☐
I have severe pain or discomfort ☐
I have extreme pain or discomfort ☐

ANXIETY / DEPRESSION
I am not anxious or depressed ☐
I am slightly anxious or depressed ☐
I am moderately anxious or depressed ☐
I am severely anxious or depressed ☐
I am extremely anxious or depressed ☐

b) EQ-5D-3L descriptive system

Mobility
I have no problems in walking about ☐
I have some problems in walking about ☐
I am confined to bed ☐

Self-Care
I have no problems with self-care ☐
I have some problems washing or dressing myself ☐
I am unable to wash or dress myself ☐

Usual Activities *(e.g. work, study, housework, family or leisure activities)*
I have no problems with performing my usual activities ☐
I have some problems with performing my usual activities ☐
I am unable to perform my usual activities ☐

Pain/Discomfort
I have no pain or discomfort ☐
I have moderate pain or discomfort ☐
I have extreme pain or discomfort ☐

Anxiety/Depression
I am not anxious or depressed ☐
I am moderately anxious or depressed ☐
I am extremely anxious or depressed ☐

Vita

Ernest H. Law | **B.Sc.(Pharm), Pharm.D., ACPR, BCPS**

CONTACT INFORMATION

Department of Pharmacy Systems, Outcomes and Policy
College of Pharmacy
University of Illinois at Chicago (UIC)
833 S Wood Street, MC 871
Chicago, IL 60612
Direct: 312-619-7707
Fax: 312-996-2954
E-mail: ernestlaw@me.com

DATE OF PREPARATION

June 20, 2018

EDUCATION

University of Illinois at Chicago — August 2014 - Present
Doctor of Philosophy, Pharmacy — Chicago, IL, USA
PhD candidate, specialization: Pharmacoeconomics and Outcomes Research

University of British Columbia — August 2012 - June 2014
Graduate Doctor of Pharmacy — Vancouver, BC, Canada

University of Alberta — September 2005 - June 2009
Bachelor's of Science in Pharmacy, First Class Standing — Edmonton, AB, Canada

POST-PROFESSIONAL TRAINING

Health Economics and Outcomes Research Fellowship — August 2014 – August 2016
University of Illinois / Takeda Pharmaceuticals — Chicago, IL, United States

Accredited Canadian Pharmacy Residency (**ACPR**) — July 2009 - July 2010
Alberta Health Services – Edmonton Zone — Edmonton, AB, Canada

ACADEMIC APPOINTMENTS

Graduate Research Assistant — January 2017 – Present

Department of Pharmacy Systems, Outcomes & Policy — Chicago, IL, United States
University of Illinois at Chicago College of Pharmacy

Graduate Teaching Assistant — August 2016 – December 2016
Department of Pharmacy Systems, Outcomes & Policy — Chicago, IL, United States
University of Illinois at Chicago College of Pharmacy

Postdoctoral Research Associate — August 2014 – August 2016
Department of Pharmacy Systems, Outcomes & Policy — Chicago, IL, United States
University of Illinois at Chicago College of Pharmacy

CLINICAL APPOINTMENTS

Clinical Pharmacist — May 2014 – August 2014
Pediatric Intensive Care Unit — Edmonton, AB, Canada
University of Alberta Hospital/Stollery Children's Hospital

Clinical Pharmacist — July 2010 – July 2012
Cardiac Intensive Care Unit — Edmonton, AB, Canada
University of Alberta Hospital/Mazankowski Heart

FIELDS OF RESEARCH

Health Economics & Outcomes, Health Preferences, Patient-Reported Outcomes, Comparative Effectiveness Research, Shared Decision-Making, Pharmacotherapy & Pharmaceutical Care

PUBLICATIONS

IN PREPARATION

1. **Law EH**, Lee TA, Schwartz A, Walton SM, Xie F. Pickard AS. Time-Specific Differences in Stated Preferences for Health in the United States. (PhD dissertation)

2. **Law EH**, Lee TA, Schwartz A, Walton SM, Xie F. Pickard AS. Health Values among US Adults with and without Advance Directives (PhD dissertation)

3. **Law EH**, Cha AS, Shaw JW, Pickard AS. Have American Perceptions of Health Changed? Comparing self-rated health among the United States adult general population in 2002 and 2017

PEER-REVIEWED JOURNAL ARTICLES (* indicates corresponding author)

1. Pickard AS*, **Law EH**, Jiang R, Oppe M, Shaw JW, Xie F, Boye KS, Gong CL, Chapman RH, Balch A. United States Valuation of EQ-5D-5L Health States Using an International Protocol (*submitted*)

2. **Law EH**, Pickard AS*, Xie F, Walton SM, Lee TA, Schwartz A. Parallel Valuation: A Direct Comparison of EQ-5D-5L and EQ-5D-3L Societal Value Sets (*submitted*).

3. **Law EH***, Jiang R, Kaczynski A, Mühlbacher AC, Pickard AS. The Role of Personality in Treatment-related Outcome Preferences among Future Pharmacists. American Journal of Pharmaceutical Education (in press)

4. Jiang R, **Law EH***, Zhou Z, Yang H, Wu EQ, Seifeldin R. Clinical Trajectories, Healthcare Resource Use, and Costs of Diabetic Nephropathy among Patients with Type 2 Diabetes: a Latent Class Analysis. Diabetes Therapy 2018; 9(3):1021-1036.

5. **Law EH***, Harrington R, Alexander GC, Saha S, Oerhlein E, Perfettto E. Stakeholder Uptake of Comparative Effectiveness & Patient-Centered Outcomes Research: Conference Insights. Journal of Comparative Effectiveness Research 2018 ;7(2):181-191

6. **Law EH**, Pickard AL, Kaczynski A, Pickard AS*. Choice Blindness and Health-State Choices among Adolescents and Adults. Medical Decision Making 2017; 37: 680-687.

7. Zhou Z, Chaudhari P, Yang H, Fang AP, Zhao J, **Law EH**, Wu EQ, Jiang R, Seifeldin R*. Healthcare Resource Use, Costs, and Disease Progression Associated with Diabetic Nephropathy in Adults with Type 2 Diabetes: A Retrospective Observational Study. Diabetes Therapy 2017; 8(3):555-571.

8. **Law EH***, Badowski M, Hung YT, Weems K, Sanchez A, Lee TA. Association Between Proton Pump Inhibitors and Microscopic Colitis: Implications for Practice and Future Research. Annals of Pharmacotherapy 2017; 51(3): 253 - 263

9. Chaudhari P, Vallarino C, **Law EH**, Seifeldin R*. Evaluation of patients with type 2 diabetes mellitus receiving treatment during the pre-diabetes period: Is early treatment associated with improved outcomes? Diabetes Research and Clinical Practice 2016; 122: 162-169.

10. Leung TS, **Law EH*.** Differential benefit risk assessment of DOACs in the treatment of venous thromboembolism: focus on dabigatran. Drug Design, Development and Therapy 2015; 2015(9): 3557–3569.

11. Calip GS*, **Law EH**, Ko NY. Racial and ethnic differences in risk of second primary cancers among breast cancer survivors. Breast Cancer Research and Treatment 2015; 151(3): 687-96.

12. Bong JL, **Law EH***, Koshman SL, Tymchak WJ, Moghrabi A, Ackman ML. Evaluating current practice and outcomes of therapeutic anticoagulation during intra-aortic balloon counterpulsation in a coronary care unit. Current Research: Cardiology 2015;2(1):15-18.

13. **Law EH***, Leung M. Corticosteroids in Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis: Current Evidence and Implications for Future Research. Annals of Pharmacotherapy 2015; 49(3) 335–342

14. **Law EH*** & Gordon W. Target-specific oral anticoagulants in patients undergoing cardioversion. American Journal of Health-System Pharmacists 2014; 71:1171-6.

15. Barry AR, Boswell R, **Law EH** et al. Review of the top 5 cardiology studies of 2011. Canadian Pharmacists Journal 2013;146(4): 213–217.

16. **Law EH** & Simpson SH*. Aspirin use rates in diabetes: A systematic review and cross-sectional study. Canadian Journal of Diabetes 2010; 34(3):211-217

17. Batchu SN, **Law EH**, Brocks DR et al. Epoxyeicosatrienoic acid prevents post-ischemic electrocardiogram abnormalities in an isolated heart model. Journal of Molecular and Cellular Cardiology 2009;46(1):67-74

COMMENTARIES, LETTERS, & DEBATES

1. **Law EH***, Harrington R. A Primer on Latent Class Analysis. Value & Outcomes Spotlight. 2016;2(6):18-19.

2. **Law EH***, Badowski M, Lee TA.
   Reply: Association Between Proton Pump Inhibitors and Microscopic Colitis: Implications for Practice and Future Research.

3. **Law EH***. “What does the p-value actually mean?“. Q&A: Clinical Research Column. Pharmacy Practice+, CanadianHealthcareNetwork.ca - October 2015 - Current

4. Egan G, **Law EH**, Mailman J et al. Should Accreditation Canada’s Required Organizational Practices and Standards Lead to Prioritization of Clinical Pharmacy Services over Distribution-Related Medication Safety Strategies? (Pro Side). Can J Hosp Pharm 2013; 66(3):372.

POLICY REPORTS

1. Pickard AS, **Law EH**, Crawford SY, Manasse HR Jr., Hopkins T, Lee TA, Sharp LK, Suda K, Cosel G. National Assessment of State Oversight of Sterile Drug Compounding. A report from the Pew Charitable Trusts. February 2016. (Available at: http://www.pewtrusts.org/en/research-and-analysis/reports/2016/02/national-assessment-of-state-oversight-of-sterile-drug-compounding

2. **Law EH**, Moulton K, Kamel C. Rapid Response: Copper Intra-uterine Devices as a Contraceptive for Adult Women: A Review of Clinical Effectiveness and Safety. Canadian Agency of Drugs and Technologies in Health. (Published September 16, 2013: http://www.cadth.ca/en/products/rapid-response/publication/3951)

CONFERENCE REPORTS

1. **Law EH**. Comparative Effectiveness and Patient-Centered Outcomes Research: Enhancing Uptake and Use by Patients, Clinicians, and Payers. April 2017. (Available at:

http://www.phrmafoundation.org/wp-content/uploads/2017/01/2017-04-02-CER-Conference-Report-Final.pdf)

BOOK CHAPTERS

1. Editors: David Hui, Ann Thompson, Peter Hamilton; Section Editor: Catherine Sych; **Assistant Editor: Ernest Law.** (2011). Chapter 18: Fluids & Supplements. Drugs and Drugs- A Practical Guide to the Safe Use of Common Drugs in Adults (Fourth Edition). Published, Department of Medicine, University of Alberta

2. Editors: David Hui, Ann Thompson, Peter Hamilton; Section Editor: Catherine Sych; **Assistant Editor: Ernest Law.** (2011). Chapter 9: Inotropes & Vasopressors. Drugs and Drugs: A Practical Guide to the Safe Use of Common Drugs in Adults (Fourth Edition). Published, Department of Medicine, University of Alberta

BOOK REVIEWS

1. Sadowski CA* & **Law E.** "Pathology and Therapeutics for Pharmacists: A Basis for Clinical Pharmacy Practice." Journal of Pharmacy & Pharmaceutical Sciences 11.3 (2008): 66-67.

PRESENTATIONS

ABSTRACTS & POSTERS

1. Cha AS, **Law EH**, Shaw JW, Pickard AS. Is Self-Rated Health of Americans Different in 2017 Compared to 2002? ISPOR Annual International Meeting 2018. Baltimore, MD, USA.

2. **Law EH**, Hopkins T, Jiang R, Pickard AS. Does exposure to health valuation tasks impact one's self-rated health? ISPOR Annual International Meeting 2018. Baltimore, MD, USA.

3. Cha AS, **Law EH**, Shaw JW, Pickard AS. American Perceptions of Health Then and Now: Comparing the United States General Adult Population in 2002 and 2017. University of Illinois at Chicago College of Pharmacy Research Day, Chicago, IL, USA - 2018

4. Zhou Z, Yang H, Jiang R, **Law EH,** Munsaka M, Seifeldin R. Diabetic Nephropathy Clinical Trajectories Among Groups of Patients with Type 2 Diabetes: A Latent Class Analysis. 77th American Diabetes Association Scientific Sessions. San Diego, USA – 2017.

5. **Law EH**, Jiang R, Kaczynski A, Mühlbacher AC, Pickard AS. Clinician preferences for cancer treatment outcomes: does personality influence choice? Value in Health 2017; 20 (5): A116.

6. Adamson BJ, Lyons N, Chou Y, Gangan N, Cannon-Dang E, Chen C, **Law EH**. Meaningful mentorship associated with rewarding short-term research: results from an evaluation of ISPOR student experiences. Value in Health 2017; 20 (5): A340.

7. Gangan N, Chou Y, Adamson BJ, Cannon-Dang E, Chen C, **Law EH**, Lyons N. Assessing trends of developing student interests within the ISPOR Student Network, 2011 to 2016. Value in Health 2017; 20 (5): A341.

8. **Law EH**, Jiang R, Kaczynski A, Mühlbacher AC, Pickard AS. Does clinician personality affect the choice of cancer treatment outcomes defined by the EQ-5D-Y? 2nd EuroQol Academy Meeting, Noordwijk, Netherlands – 2017

9. **Law EH**, Jiang R, Kaczynski A, Mühlbacher AC, Pickard AS. Does pharmacy student personality affect the choice of cancer treatment outcomes? University of Illinois at Chicago College of Pharmacy Research Day, Chicago, IL, USA - 2017

10. **Law EH**, Pickard A, Kaczynski A, Pickard AS. Validity of Health Preferences Elicited from Children: Comparing Choice Blindness and Stated Preferences between Children and Adults. Value in Health 2016; 19(3): A9.

11. **Law EH**, Joo MJ, Vinicky M, Pickard AS. An Approach To Evaluating The Usefulness Of Supplementing The Eq5d Descriptive System In Obstructive Pulmonary Disease. Value in Health 2016; 19(3): A117.

12. **Law EH**, Sharma D, Pickard AS. Models Used in Economic Analyses of Ticagrelor and Prasugrel For Acute Coronary Syndromes: A Structured Review. Value in Health 2015; 18(3): A14.

13. Jalundhwala YJ, Saraf SL, Rondelli D, **Law EH**, Pickard AS. Improvement in Health-related Quality of Life for Sickle Cell Disease Patients Treated with Chemotherapy-free Allogeneic Hematopoietic Stem-cell Transplantation: a pilot study. Value in Health 2015; 18(3): A302

14. Calip GS, **Law EH**, Winer EP, Ko NY. Racial and ethnic differences in risk of second primary cancers among breast cancer survivors. J Clin Oncol 33, 2015 (suppl; abstr e12592)

15. Calip GS, **Law EH**, Lee WJ, Schwartz S, Ko NY. Differences in comorbidity at diagnosis and receipt of adjuvant chemotherapy among older women by race and ethnicity. J Clin Oncol 33, 2015 (suppl; abstr e17599)

16. Bong JL, **Law EH**, Koshman SI, Tymchak WJ, Moghrabi A, Ackman ML. Evaluating current practice and outcomes of therapeutic anticoagulation during intra-aortic balloon counterpulsation in a coronary care unit. Canadian Cardiovascular Congress 2013

17. **Law EH**, Gray M, Koshman S, Fuller J, et al. Characterizing the selection and use of anti-microbials for management of bacterial urinary tract infections in a tertiary care facility. Abstract. Faculty of Pharmacy and Pharmaceutical Sciences Research Day, Edmonton, Alberta — 2010

18. **Law EH**, Mitchell SM, Carson M, Leung S. Characterisation of Antibiotic Use in a Long-term Care Setting in the Edmonton Region: Do Bugs Need Drugs? A Multidisciplinary Team Approach. Abstract & Poster. Association of Faculties of Pharmacy of Canada Annual Conference, Vancouver, British Columbia — 2010

PODIUM PRESENTATIONS

1. **Law EH**, Pickard A, Kaczynski A, Pickard AS. Podium presentation: “Validity of Health State Choices Elicited from Children: Comparing Choice Blindness and Health State Choices between Children and Adults.” Midwest Social & Administrative Pharmacy Conference, Ann Arbor, MI, USA — 2016

2. **Law EH**, Pickard A, Kaczynski A, Pickard AS. Podium presentation: “Validity of Health Preferences elicited from Children: Comparing Choice Blindness and Stated Preferences between Children and Adults" 21st ISPOR Annual Meeting, Washington, DC, USA — 2016

3. **Law EH** & Simpson SH. Podium presentation: ”Cross-sectional Survey of Aspirin Use in Diabetes: An Interim Report." Faculty of Pharmacy and Pharmaceutical Sciences Research Day, Edmonton, Alberta — 2007 - Awarded Best Podium Presentation

4. **Law EH**, Seubert J. Podium presentation: “Haloperidol-induced QT-prolongation in ex vivo mice hearts”. Faculty of Pharmacy and Pharmaceutical Sciences Research Day, Edmonton, Alberta — 2006 - Awarded Best Podium Presentation

INVITED SPEAKER

1. **Law EH**. “The US EQ-5D-5L Face-to-face Value Set Using an International Protocol” EuroQol Academy Meeting 2018. Budapest, Hungary.

2. **Law EH**. “What’s the difference? A comparison of 3L and 5L value sets using a standardized protocol.” EuroQol Academy Meeting 2018. Budapest, Hungary.

3. **Law EH**, Pickard A, Kaczynski A, Pickard AS. Conference paper: “Validity of Health State Choices Elicited from Children: Comparing Choice Blindness and Health State Choices between Children and Adults.” EuroQol Scientific Plenary 2016. Berlin, Germany.

4. **Law EH**. A Look at the Pre-Conference Survey Results. Comparative Effectiveness and Patient-Centered Outcomes Research: Enhancing Uptake and Use by Patients, Clinicians and Payers 2017. Washington, DC, USA.

5. **Law EH**, Pickard AS, Xie F. Workshop: Developments in preference-based measures of health: scoring approaches and guidance. ISOQOL 2016 Annual Meeting. Copenhagen, Denmark

6. **Law EH**, Pickard A, Kaczynski A, Pickard AS. Conference paper: “Validity of Health State Choices Elicited from Children: Comparing Choice Blindness and Health State Choices between Children and Adults.” EuroQol Scientific Plenary 2016. Berlin, Germany.

7. **Law EH**. (2014). Liver Dysfunction and Anticoagulation: Chalk full of Irony. Canadian Society of Hospital Pharmacists Western Branches. 40th Annual Banff Seminar, Banff, Canada

8. **Law EH**. Sodium Restriction & Systolic Heart Failure: Is it adding in-salt to injury? Canadian Society of Hospital Pharmacists - BC Branch. Annual General Meeting 2013, Vancouver, Canada

9. **Law EH**, Moghrabi, A. Novel Antiplatelet Agents. Mazankowski Heart Institute Heart Month Nursing Conference 2011, Edmonton, Canada

10. **Law EH**. Novel Anticoagulants. Mazankowski Heart Institute Heart Month Nursing Conference 2011, Edmonton, Canada

GRANT SUPPORT

CURRENT

US Valuation of EQ-5D-5L Health States
Role: Project Coordinator & Co-Investigator
Dates: January 1, 2017 – July 1, 2018
Funder: EuroQol Research Foundation, $293,000 USD
Description: Development of United States population-based EQ-5D-5L preference weights through a series of face-to-face interviews across the United States; as project coordinator, led training and oversaw site visits and data collection by interviewer team, statistical modeling efforts, publication management, and dissemination efforts.

COMPLETE

National Assessment of State Oversight of Compounding
Role: Study Coordinator & Co-Investigator
Dates: January 1, 2015 – December 1, 2015
Funder: The Pew Charitable Trusts, $137,213
Description: A nationwide cross-sectional survey of the United States describing state oversight of compounding practices

Enhancing Student Learning and Expanding Institutional Experiential Clerkship Opportunities Across UBC's Pharmacy Programs
Role: Co-applicant
Dates: February 22, 2015 – February 22, 2015
Funder: University of British Columbia Teaching and Learning Enhancement Fund, $96,550
Description: The first of a 3-year comprehensive project to address all stakeholder perspectives and system redesign issues required to develop and implement an enhanced student-learning module for institutional pharmacy clerkships.

TEACHING EXPERIENCE

UNIVERSITY OF ILLINOIS AT CHICAGO

- PSOP 390 Independent Study (Pharm.D. course, Lecturer) — Spring 2018
- PSOP 516 Comparative Effectiveness Research (MS & PhD course, Instructor) — Spring 2018

- PMPR 329 Dean's Leadership Forum (Pharm.D. course, Lecturer) Fall 2017
- PHAR 445 Pharmacy Law (Pharm.D. course, Lecturer) Fall 2016

UNIVERSITY OF BRITISH COLUMBIA

- PHAR 299 Cases in Pharmaceutical Sciences I (BSc.Pharm course, Lecturer) Winter 2014

UNIVERSITY OF ALBERTA

- PHARM 367 Cardiology (B.Sc.Pharmacy course, Instructor) Winter 2012
- PHARM 367 Cardiology (B.Sc.Pharmacy course, Instructor) Winter 2011

TRAINEES

Pharmacy students/residents

- Student: Basel Alsedi (Cardiology rotation) – 2011
- Resident: Joseph Blais (Residency Research Project) – 2012
- Resident: Joseph Blais (Cardiology rotation) – 2010

LEADERSHIP & SERVICE

EXTRAMURAL

- Organizer, Conference Planning Committee, "Comparative Effectiveness and Patient-Centered Outcomes Research: Enhancing Uptake and Use by Patients, Clinicians and Payers" - January 26 & 27, 2017 in Washington, DC (2016-2017)
- Co-Chair, Survey & Evaluation Committee, International Society of Pharmacoeconomics and Outcomes Research – Student Network (2016-2017)
- Member, Educational Services Committee, Canadian Society of Hospital Pharmacists – National (2012-2014)
- Chair, Education Committee, Canadian Society of Hospital Pharmacists - Alberta Branch (2010-2012)

UNIVERSITY & DEPARTMENTAL

- Past President, ISPOR-UIC Student Chapter, International Society of Pharmacoeconomics and Outcomes Research (2017-2018)
- President, ISPOR-UIC Student Chapter, International Society of Pharmacoeconomics and Outcomes Research (2016-2017)
- Graduate Student Representative, College of Pharmacy, University of Illinois at Chicago (2014-2017)
- Student Representative, UIC/Takeda HEOR Fellowship Applicant Search Committee (2014)

SERVICE

- Reviewer, Health and Quality of Life Outcomes (2018-Current)
- Reviewer, Journal of Patient-reported Outcomes (2017-Current)
- Reviewer, Value in Health (2016-Current)
- Reviewer, Quality of Life Research (2015-Current)
- Judge, AMCP UIC Student Chapter P&T Competition (2015-Current)
- Reviewer, Thrombosis (2015-Current)

- Reviewer, Pharmacotherapy (2015-Current)
- Reviewer, Canadian Journal of Hospital Pharmacists (2015-Current)
- Reviewer, University of British Columbia Pharmaceutical Sciences Student Journal (2012-2014)

AWARDS AND HONORS

- Lloyd Yale Memorial Scholarship (2018)
- ISPOR Student Network Outstanding Service Award (2017)
- Van Doren Scholarship, University of Illinois at Chicago (2016)
- Chicago Consular Corps Scholarship, University of Illinois at Chicago (2015)
- Best Publication Award, Canadian Society of Hospital Pharmacy British Columbia Branch (2015)
- APEX Future of Pharmacy Award, Alberta College of Pharmacists & Alberta Pharmacists' Association (2012)
- Professional Development Grant, Canadian Society of Hospital Pharmacists Alberta Branch (2012)
- Recognition Program, Canadian Society of Hospital Pharmacists Alberta Branch (2012)
- Future Professional Award, CSHP - AB Branch (2009)
- Best Podium Presentation, Faculty of Pharmacy & Pharm. Sci. Research Day, Edmonton, CAN (2008)
- Best Poster Presentation, Faculty of Pharmacy & Pharm. Sci. Research Day, Edmonton, CAN (2007)
- Best Podium Presentation, Faculty of Pharmacy & Pharm. Sci. Research Day, Edmonton, CAN (2006)

CERTIFICATIONS

- Board Certified Pharmacotherapy Specialist (**BCPS**), ASHP (2014-Present)
- Healthcare Modeling Course, TreeAge Software Inc. (2015)
- Standard Author Training, Canadian Cochrane Collaboration (2013)
- Instruction Skills Workshop, UBC Center for Teaching, Learning, and Technology (2012)
- Alberta College of Pharmacy Practice License, #7356 (2009-Present)
- Certificant of Pharmacy Licensing Board of Canada (2009)

PROFESSIONAL ORGANIZATIONS AND SOCIETIES

- International Society of Quality of Life Research (2015 - Present)
- International Society of Pharmacoeconomics and Outcomes Research (2014 - Present)
- International Society of Pharmacoepidemiology (2014 - 2016)
- Canadian Society of Hospital Pharmacists (2007 - Present)