KCE REPORT 342

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

AN EQ-5D-5L VALUE SET FOR BELGIUM

HOW TO VALUE HEALTH-RELATED QUALITY OF LIFE?





KCE REPORT 342
HEALTH SERVICES RESEARCH

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

AN EQ-5D-5L VALUE SET FOR BELGIUM

HOW TO VALUE HEALTH-RELATED QUALITY OF LIFE?

NICOLAS BOUCKAERT, SOPHIE GERKENS, STEPHAN DEVRIESE, IRINA CLEEMPUT

2021 www.kce.fgov.be

COLOPHON

Title:	An EQ-5D-5L value set for Belgium – How to value health-related quality of life?
Authors:	Nicolas Bouckaert (KCE), Sophie Gerkens (KCE), Stephan Devriese (KCE), Irina Cleemput (KCE)
Project facilitator:	Nathalie Swartenbroeckx (KCE)
External experts:	Rana Charafeddine (Sciensano), Brecht Devleesschauwer (Sciensano), Kristina Ludwig (EuroQol Research Foundation), Geert Molenberghs (Universiteit Hasselt, KU Leuven), Mark Oppe (Maths in Health), Bram Roudijk (EuroQol Research Foundation), Elly Stolk (EuroQol Research Foundation)
External validators:	Luiz Flavio Andrade (University of Birmingham), Delphine De Smedt (Universiteit Gent), Koonal Shah (PHMR Ltd, EuroQol Group's Executive Committee)
Acknowledgements:	We would like to thank the Profacts' team members and interviewers for their essential contribution to the study, the numerous KCE colleagues who have been involved throughout the project and Elly Stolk (EuroQol Research Foundation) for her valuable contribution as representative of EuroQol. We also thank the EuroQol Research Foundation for the financial support to carry out this study.
Reported interests:	The current study was co-financed by the EuroQol Research Foundation that covered half of the expected study costs. All experts and stakeholders consulted within this report were selected because of their involvement in research on the EQ-5D family of instruments or EQ-5D-5L valuation studies. Therefore, by definition, each of them might have a certain degree of conflict of interest to the main topic of this report.
Cover picture:	The image is taken from the EuroQol website (https://euroqol.org/support/tools/educational-material/) and reproduced by permission of EuroQol Research Foundation
Layout:	Ine Verhulst
Disclaimer:	• **The external experts were consulted about a (preliminary) version of the scientific report. Their comments were discussed during meetings. They did not co-author the scientific report and did not necessarily agree with its content.** • **Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. The validators did not co-author the scientific report and did not necessarily all three agree with its content.** • **Finally, this report has been approved by common assent by the Executive Board.** • **Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE.**

Publication date:	01 July 2021
Domain:	Health Services Research (HSR)
MeSH:	Quality of Life; Quality-Adjusted Life Years; Surveys and Questionnaires; Cost-Benefit Analysis; Models, Statistical
NLM Classification:	W84.3 Health Services - Research (General)
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2021/120210.273/19
ISSN:	2466-6459
Copyright:	The EQ-5D sample version in Figure 1 is reproduced by permission of EuroQol Research Foundation. Reproduction of this sample version is not allowed. For reproduction, use or modification of the EQ-5D (any version), please register your study by using the online EQ registration page: www.euroqol.org KCE reports are published under a “by/nc/nd” Creative Commons Licence http://kce.fgov.be/content/about-copyrights-for-kce-publications.



How to refer to this document?	Bouckaert N, Gerkens S, Devriese S, Cleemput I. An EQ-5D-5L value set for Belgium – How to value health-related quality of life? Health Services Research (HSR) Brussels: Belgian Health Care Knowledge Centre (KCE). 2021. KCE Reports 342. D/2021/10.273/19. This document is available on the website of the Belgian Health Care Knowledge Centre.

KEY POINTS

- The ***objective*** of this study is to produce an EQ-5D-5L value set for Belgium, based on the preferences of the general public. The new EQ-5D-5L value set updates the currently used EQ-5D-3L value set from 2003 with 243 potential health states to a more extensive value set with 3 125 potential health states in addition to the state of unconsciousness.
- The EQ-5D-5L value set developed in this study is based on a ***strong and valid method***.
  - The study design followed the most recent valuation protocol developed by EuroQol.
  - Respondents were randomly selected using a multistage, stratified, cluster sampling with unequal probability design. The final sample of 892 respondents was representative for the Belgian population in terms of age, sex, region, health status, health-related quality of life, education and employment status.
  - In face-to-face, computer-assisted interviews, respondents completed 10 composite time trade-off (cTTO) exercises and a discrete choice experiment (DCE) with 7 paired choice tasks. About 20% of the respondents also valued the state of unconsciousness by cTTO.
  - Data were collected over the period May 2018 to September 2020. One limitation of the study is that the desired target of 1 000 respondents was not reached. The COVID-19 crisis was one reason, but in general it proved very difficult to recruit respondents for face-to-face interviews at home.
  - Face validity of the data has been demonstrated, with more severe health states generally given lower values. Interviewer specific effects may impact the valuation of health states considered worse than dead.
  - A wide range of potential models were estimated and evaluated on logical consistency, goodness of fit, predictive accuracy and theoretical considerations. A parsimonious 8-parameter multiplicative hybrid model with random effects and correction for heteroscedasticity was selected to produce the value set for the 3 125 EQ-5D-5L health states. The utility of the unconscious state was not modelled but taken directly from survey results.
- Problems in the dimensions pain/discomfort and anxiety/depression have the highest negative impact on health-related quality of life.
- ***Compared to the existing EQ-5D-3L*** value set, the 5L value set has a higher fraction of health states valued worse than dead (15.0% in 5L and 6.6% in 3L) and covers a wider range of utility values (-0.532 to 0.939 in 5L and -0.158 to 0.817 in 3L, excluding state 11111). A wider range of utility values allows for a better discrimination between patients, both for mild and (very) severe health conditions.
- The new EQ-5D-5L value set will support the use of the EQ-5D-5L instrument in research – e.g. clinical trials, population norms –, outcome measurement – e.g. PROMs –, and decision-making – e.g. cost-effectiveness analysis for pricing decisions in the public health insurance, quality management.
  Analyses performed with the current 3L and the new 5L value set are expected to produce different results. For example, QALY gains for a specific intervention in cost-effectiveness analysis can be larger or smaller.

TABLE OF CONTENTS

KEY POINTS 1
LIST OF FIGURES 5
LIST OF TABLES 6
LIST OF ABBREVIATIONS 7
SCIENTIFIC REPORT 9
1 INTRODUCTION 9
1.1 HEALTH-RELATED QUALITY OF LIFE 9
1.2 APPLICATION OF HEALTH STATE UTILITY DATA IN CLINICAL AND ECONOMIC EVALUATIONS 12
1.3 THE EQ-5D INSTRUMENT 13
1.4 STUDY OBJECTIVE 15
2 METHODS 16
2.1 STUDY PROTOCOL 16
2.1.1 Valuation techniques 16
2.1.2 EQ-VT Software 19
2.1.3 Structure of the interview 19
2.1.4 Target sample 20
2.1.5 Number of interviewers 20
2.1.6 Training and pilot tests 21
2.1.7 Quality control 21
2.2 SAMPLING PROCESS 21
2.2.1 Inclusion and exclusion criteria 21
2.2.2 Sampling procedure 22

2.3 STUDY APPROVAL AND PRIVACY PROTECTION ASPECTS 31
2.4 DATA COLLECTION 32
2.4.1 Initial recruitment of interviewers 32
2.4.2 Recruitment of respondents 32
2.4.3 Data collection time frame and recruitment of new interviewers 32
3 DATA ANALYSIS 37
3.1 DATA PREPARATION AND REPRESENTATIVENESS 37
3.1.1 Creation modelling dataset 37
3.1.2 Representativeness 38
3.2 MODEL CONSTRUCTION 39
3.2.1 Modelling cTTO valuation 39
3.2.2 Modelling DCE valuation 41
3.2.3 Hybrid model 42
3.2.4 Fitting of the models 43
3.3 MODEL SELECTION 43
3.3.1 Logical consistency 43
3.3.2 Goodness of fit 43
3.3.3 Predictive accuracy 44
3.3.4 Theoretical considerations 45
3.4 PREFERRED MODEL AND VALUE SET 46
4 RESULTS: THE BELGIAN EQ-5D-5L VALUE SET 47
4.1 SAMPLE CHARACTERISTICS 47
4.2 DATA CHARACTERISTICS 52
4.2.1 cTTO data 52

4.2.2 DCE data 59
4.3 MODELLING RESULTS AND SELECTION 60
4.3.1 Logical consistency 60
4.3.2 Goodness of fit 60
4.3.3 Predictive accuracy 62
4.3.4 Theoretical considerations 63
4.4 VALUE SET 68
4.4.1 Summary of model selection 68
4.4.2 Preferred model and value set 68
4.4.3 Comparison with 3L value set and value sets of other countries 72
5 DISCUSSION AND CONCLUSION 75
5.1 STRENGTHS AND LIMITATIONS IN THE DEVELOPMENT OF THE NEW VALUE SET 75
5.2 IMPLICATIONS OF THE NEW VALUE SET 76
5.3 APPLICABILITY OF THE BELGIAN EQ-5D-5L VALUE SET 77
5.4 FUTURE RESEARCH AND REFLECTIONS 78
■ **APPENDICES 79**
APPENDIX 1. LIKELIHOOD FUNCTIONS IN REGRESSION ANALYSIS 79
APPENDIX 2. REPRESENTATIVENESS OF THE POPULATION AT REGIONAL LEVEL 82
APPENDIX 3. FINAL VALUE SET 85
■ **REFERENCES 100**

LIST OF FIGURES

Figure 1 – The EQ-5D-5L questionnaire (UK English sample version), combined with the visual analogue scale 14

Figure 2 – The composite Time Trade-Off (cTTO): traditional TTO for states better than dead and the lead-time TTO for states worse than dead 18

Figure 3 – The DCE technique 18

Figure 4 – Selected municipalities (PSU) 27

Figure 5 – Female population distribution by province and age of the legal Belgian population on 1 January 2017 (left panel); number of planned interviews for female participants by province and age category (right panel) 29

Figure 6 – Male population distribution by province and age of the legal Belgian population on 1 January 2017 (left panel); number of planned interviews for male participants by province and age category (right panel) 30

Figure 7 – Age and sex groups proportions of the Belgian population versus sample 31

Figure 8 – Pace of the study, number of interviews and impact of training and quality control 33

Figure 9 – Flow chart on the change over time in the number of interviewers 35

Figure 10 – Distribution of the number of interviews per interviewer 36

Figure 11 – Distribution of age, sex, region, health, health-related quality of life, education, and employment status in the EQ-5D-5L survey sample relative to the population 49

Figure 12 – Distribution of observed cTTO values (unconscious state not included) 52

Figure 13 – Distribution of observed cTTO values for the 86 health states (unconscious state not included) 53

Figure 14 – Mean and median cTTO value by severity level (unconscious state not included) 54

Figure 15 – Differences in cTTO valuation by interviewer 55

Figure 16 – Mean and median cTTO values by level and dimension (unconscious state not included) 58

Figure 17 – Preference for alternative A or B in DCE task by relative severities of A and B 59

Figure 18 – Agreement between DCE predicted values and cTTO predicted values for the 86 health states used in the valuation 65

Figure 19 – Scatterplot of predicted (hybrid) and observed (cTTO) utility values for the 86 health states used in the valuation 69
Figure 20 – Kernel density plots of value sets 74
Figure 21 – Distribution of age and sex in the EQ-5D-5L survey sample relative to the population, by region 83

LIST OF TABLES

Table 1 – Examples of instruments and outcomes of HRQoL measures 10
Table 2 – Sampling at the province level 23
Table 3 – Summary statistics for the 86 health states of the cTTO valuation and the unconscious state 56
Table 4 – Evaluation of goodness of fit 61
Table 5 – Evaluation of predictive accuracy 62
Table 6 – Evaluation DCE fit 65
Table 7 – Coefficients of cTTO, DCE and hybrid models of MULT8 and MULT9 with intercept, random effects and correction for heteroscedasticity 67
Table 8 – Coefficients and bootstrapped standard errors of the preferred model and results of the sensitivity analysis 70
Table 9 – Comparison of key characteristics of the value sets produced by different high-performing models 72
Table 10 – Comparison of key characteristics of the Belgian EQ-5D-5L value set, Belgian EQ-5D-3L value and EQ-5D-5L value sets of neighbouring countries 73
Table 11 – EQ-5D-5L value set for Belgium 85

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
5D	Five dimensions of health
5L / 3L	Five / Three severity levels
AIC	Akaike's information criterion
BIC	Bayesian information criterion
cTTO	Composite Time trade-off
DCE	Discrete choice Experiment
DE	German speaking
EQ	Euroqol
EQ-VT	EuroQol Valuation Technology
EUnetHTA	European network for Health Technology Assessment
EU-SILC	European Survey on Income and Living Conditions
FR	French-speaking
HIS	Health interview survey
HRQoL	Health-related quality of life
HUI_3	Health Utilities Index Mark 3
ICER	Incremental cost-effectiveness ratio
MAE	Mean Absolute Error
NL	Dutch speaking
PROM	Patient Reported Outcome Measure
PSU	Primary sampling unit
QALY	Quality-adjusted life-year
RMSE	Root mean squared error
SD	Standard Deviation

SF-6D	Short-Form Six dimension of health
SF-36	Short Form (36) Health Survey
TTO	Time trade-off
UK	United Kingdom
VAS	Visual analogue scale
WTD	Worse than dead

SCIENTIFIC REPORT

1 INTRODUCTION

1.1 Health-related quality of life

With technological progress and the increasing life expectancy of the population, the impact of health interventions on the quality of life of the population is becoming an important concern in public health policies and patient management decisions. The aim is no longer to just extend people's lives but also to improve or try to maintain their health-related quality of life.[1, 2]

What is health-related quality of life?

While quality of life is a very broad concept that can be influenced by all aspects of an individual existence, health-related quality of life (HRQoL) focuses on the way health affects the quality of life, i.e. on the impact of health status on quality of life.[3] According to a more specific definition, HRQoL focuses on the value of health, i.e. the value (also called the utility) assigned to different health states.[4] In 2013, the European network for Health Technology Assessment (EUnetHTA) produced guidelines on HRQoL and utility measures and has defined HRQoL as "*a patient's general subjective perception of the effect of illness and intervention on physical, psychological and social aspects of daily life*".[3, 5, 6]

How to measure health-related quality of life?

Various methods are available to measure HRQoL. Table 1 gives an overview of different types of measures with some examples of instruments.[7]

A first distinction between methods concerns the scope of the measure. A measure can either be disease- or population specific or be generic (i.e. applicable to any condition or population group).[3, 7] Disease- or population-specific instruments are designed to assess particular diseases or patient populations (e.g. Asthma Quality of Life Questionnaire or Child Health Questionnaire) and are claimed to be more responsive in the detection of treatment induced changes, more relevant to patients and clinicians and more sensitive to small differences or changes in HRQoL than generic instruments.[8] However, it is found that the appropriateness of a generic

measure depends on the condition being examined and can be as sensitive as disease-specific measures.[9] There are some major advantages of generic instruments. They allow for broad comparisons of outcomes and interventions across various diseases and populations. As generic instruments are more comprehensive than disease- or population specific instruments, they are more likely to pick up unexpected effects on HRQoL in dimensions not addressed by a disease-specific instrument.[10] Moreover, they improve the comparability between studies.[3, 7, 8]

A second distinction can be made based on the how the values of HRQoL are expressed. Some instruments generate a profile of scores, i.e. a separate score per item or dimension of HRQoL, while others summarize the multidimensional health outcomes into one overall summary score. In preference-based measures, which are a specific case of the latter group, such summary score is denoted as the utility (or preference weighting). Utilities represent individuals' preferences for different health states and are standardized on a scale where 1 is the value of 'full health', 0 is the value of 'dead', and states worse than dead have negative values.[3, 7, 9]

Different elicitation techniques are used to derive utilities. Three of the most commonly used techniques are the time trade-off, the standard gamble and the rating scale (see Box 1 for more details on these techniques).[11-13] Given that the time trade-off and the standard gamble can be cognitively challenging for respondents, other techniques have also been used, such as pairwise discrete choice experiments (DCEs) where respondents have to choose which of two health states they prefer (see Box 1).[14-20]

Table 1 – Examples of instruments and outcomes of HRQoL measures

	Descriptive instrument for HRQoL		Direct assessment of HRQoL without using a descriptive instrument
Result	**Disease- or population specific**	**Generic**	
Profile	Kidney Disease Quality of Life Short Form, Liver Disease Symptom Index	Sickness Impact Profile, Nottingham Health Profile, SF-36	
Summary score	End-stage Renal Disease Symptom Checklist*	Functional Limitations Profile (total summary score derived from dimension scores)	SF-36 question about change in HRQoL over 1 year (not scaled)**
Utility index (as a special case of a summary score)		EQ-5D, Health Utilities Index, SF-6D (indirect valuation)	Valuation of patients' health state with Time Trade-Off, Standard Gamble or Visual Analogue Scale calibrated on 0 to 1 scale (direct valuation).

** combines profile measures with summary score*
*** Is a single item question, answered on an ordinal scale*
Source: Cleemput and Neyt (2015)[7]

In case of preference based measures, a third distinction relates to whose preferences are used to attribute utilities to health states. Preferences and the resulting utilities can be assessed *directly* in the study population (direct evaluation) using for example one of the elicitation techniques described in Box 1. An alternative approach is to use a multidimensional generic HRQoL instrument in combination with a pre-existing set of utility values or *value set.* A value set contains a fixed utility value for each of the possible health states that can be described with the HRQoL instrument. It is produced in a valuation study and is usually based on the preferences of the general public surveyed by means of an elicitation technique (as described in Box 1). Once a value set is determined, it can be used to attribute values to the health states of any study population. The use of multidimensional generic HRQoL instruments in combination with a pre-existing value set is referred to as an *indirect* evaluation. Different instruments exists, such as the Health Utilities Index 2 and 3 (HUI2 and HUI3), the Short Form 6D (SF-6D) and the EuroQol EQ-5D instruments.[3, 13, 21]

Box 1 – Examples of techniques for measuring health state utilities

To measure the utility of a health state, different techniques can be used. Each technique values a specific health state relative to 'full health' with a standardized utility of 1 and 'dead' with a standardized utility of 0. Consider for example a health state *S* with moderate chronic pain and no other health issues.

Standard gamble: Two alternatives lives are compared. Living *Y* years in a health state *S* (alternative 1) compared to a gamble (alternative 2), i.e. opting for a (hypothetical) health intervention with an uncertain result. The health intervention has two possible outcomes: there is a probability *p* to be cured and live in full health for the same number of years *Y*, there is a probability *1-p* to die immediately. In alternative 2, the probability *p* to be in full health is varied until the respondent is indifferent between both alternatives ($p_{indifferent}$). The utility assigned to the evaluated health state *S* is equal to $p_{indifferent}$. In our example, alternative 1 would be to live 10 years with moderate chronic pain. Alternative 2 would be to have an intervention (e.g. surgery or medication) that might result in living 10 years in full health or immediate death. If the respondent is indifferent between both alternatives when the intervention has a probability of success of 80%, the utility of health state *S* would equal 0.80.

Time trade-off (TTO): Two alternatives lives are compared. Living Y years in a health state S (alternative 1) compared to living X years in full health (alternative 2). In alternative 2, the number of years in full health, X is by definition lower or equal to Y and it is varied until the respondent is indifferent between both alternatives ($X_{indifferent}$). The utility assigned to health state *S* is equal to $X_{indifferent}$ /*Y*. In our example, alternative 1 would be to live 10 years with moderate chronic pain. Alternative 2 would be to live in full health for 10 years or less. If the respondent is indifferent between both alternatives with a life span in full health of 8 years, the utility of health state *S* would equal 0.80.

Rating scale (e.g. the Visual Analog Scale or VAS in Figure 1): Respondents must locate the health state *S* on a scale calibrated from 0 (immediate death) to 100 (full health). The utility of the health state S is calculated by dividing the score obtained by 100. In our example, if the respondent would score living with moderate chronic pain at 80 on the rating scale, the utility of health state *S* would equal 0.80.

Pairwise discrete choice experiments: Respondents are presented with a choice between two health states – e.g. health state *A* (alternative 1) with moderate anxiety or depression and severe pain but no other health issues and another health state *B* with severe problems in walking about and no other health issues (alternative 2) – and must indicate which of both alternative states they prefer. Such pairwise choices, however, create only a relative ranking of health states and their attributes, producing a latent scale ranging from the "best" to "worst" health state.[19] Additional information is necessary to transform this to utility values on the full health–dead scale.[19] One possibility is to rescale the DCE-values, using a rescaling factor based on information from other techniques such as TTO, the standard gamble or the rating scale.[19, 22, 23] Another possibility is to add an attribute for duration to the health state and introduce immediate death as a third alternative option.[20, 24] Adding duration as an attribute to health states, however, makes the DCE task more challenging for respondents.[24]

Box 2 – Definition of full economic evaluations

A full economic evaluation is defined as a comparative analysis of at least two health interventions in terms of their costs and health consequences (outcomes). The result of an economic evaluation is expressed as an incremental cost-effectiveness ratio (ICER), reflecting the additional (incremental) cost per additional (incremental) unit of outcome achieved. The lower the ICER, the more cost-effective an intervention is. There are three types of full economic evaluations. They differ in the way health outcomes are assessed:

- **Cost-effectiveness analyses**, where outcomes are expressed in terms of final clinical endpoints. The endpoint can be generic, i.e. life years gained (as recommended by the Belgian guidelines for economic evaluations), or a disease-specific, e.g. ulcers prevented. The latter only allows, however, for comparisons of different interventions for the same indication. Endpoints can also be intermediary clinical outcomes, such as the decrease in blood pressure or improvements in bone mineral density, but this has several disadvantages (lack of comparability, uncertain impact on final and/or patient-relevant outcomes, etc.); Therefore the use of intermediary clinical outcomes is also not recommended by the Belgian guidelines for economic evaluations. **Cost-minimisation analyses** are a specific case of cost-effectiveness analyses where the effects of two treatments are identical.
- **Cost-utility analyses** can also be considered a special case of cost-effectiveness analysis, where outcomes are expressed in terms of utility-based quality-adjusted life years gained (QALYs);
- **Cost-benefit analyses**, where outcomes are expressed in monetary terms, representing the people's willingness to pay.

According to Belgian guidelines for economic evaluations, cost-utility analyses should be used if *"the treatment has an impact on health-related quality of life that is significant to the patient or if there are multiple patient-relevant clinical outcome parameters expressed in different units that cannot be translated into one common unit in a valid way"*.[25]

1.2 Application of health state utility data in clinical and economic evaluations

Public resources are scarce and evidence-informed policy is based on a trade-off between on the one hand the positive impact of health interventions on the quality and duration of life and on the other hand the cost of the interventions to achieve it.[2, 3] Health economic evaluations, or cost-effectiveness analyses, aim at objectifying the trade-offs that are made when deciding to reimburse a new health intervention, by making explicit the additional cost for the additional benefits (more information on economic evaluations can be found in Box 2).[25] When the healthcare budget is fixed, spending more for one treatment means spending less elsewhere in the healthcare sector. Cost-effectiveness evaluations can inform policy makers about where additional resources will generate the highest health gains and hence how scarce resources in health care can be allocated efficiently.[3] However, this requires a common denominator, allowing comparisons between different types of interventions for different diseases. Health state utility data, combined with data on duration of life, can be used for this purpose. The quality-adjusted life-year – or QALY – is a measure that captures both the quality and quantity of life in one metric, and is frequently used as a measure of clinical effectiveness of interventions in health economic evaluations (cost-per-QALY analyses).[3]

To evaluate improvements in HRQoL and life expectancy, i.e. the effectiveness of a health intervention, in a cost-effectiveness analysis, data are required on the HRQoL and life expectancy without the intervention (i.e. with standard of care) and the HRQoL and life expectancy with the intervention. This allows to calculate the number of QALYs gained with the new intervention compared to the standard of care.

The Belgian guidelines for economic evaluations in health care recommend the use of indirect evaluation of HRQoL, i.e. using a pre-existing set of utility values from a generic HRQoL instrument (see above), to calculate QALYs.[25] According to the guidelines, the description of a health state should be done by patients experiencing that particular health state, while the health state utilities should come from the general public.[25] Using utilities derived from the general public rather than patients has the advantage of treating the

same health state in the same way, thereby increasing the comparability across evaluations.[3, 11, 26] Moreover, a valuation by the general public, without a vested interest in the treatment of a particular health state, avoids the impact of strategically overemphasizing the benefits of a new treatment by patients to influence resource allocation decisions.[11, 27] Another potential source of bias when eliciting preferences of patients is patients' coping. Through coping, severe chronic health states could appear less severe than acute mild health states, while the relative importance of treating severe chronic conditions may be higher.[3, 11, 27]

To determine QALYs using indirect utility measurements in a concrete situation, the first step is to ask people, usually patients, to describe their health state using a generic health state description instrument, like the EQ-5D, SF-6D or HUI_3. The second step is to assign utility values to these health states. For this assignment, the pre-existing value set of the instrument is used. QALYs are then calculated by weighting the time spent in a health state by the utility associated with that health state. For the QALYs to have an intuitive meaning (one QALY equals one year in full health and a dead person has zero QALYs), the utility values must be expressed on a scale where 0 is the value for 'dead' and 1 is the value for 'full health' (see section 1.1).[11, 13] Negative values are possible for health states considered worse than dead.[28]

Box 3 – Description of EuroQol[29]

EuroQol is a not for profit organization that comprises an international network of multidisciplinary researchers working on the development of standardized generic instruments to describe and value health-related quality of life.[29] Their current instruments, i.e. the EQ-5D-3L, the EQ-5D-5L and the EQ-5D-Y (for children and adolescents) are the result of nearly 30 years of development and adaptation. Numerous research studies are yet in process to adapt and further develop the EuroQol portfolio. The EuroQol Research Foundation was created to support and instigate scientific research and to govern the intellectual property rights that include, but are not limited to EQ-5D instruments.

1.3 The EQ-5D instrument

Belgian guidelines recommend the use of EQ-5D to measure HRQoL

In order to stimulate the use of generic HRQoL instruments and to promote consistency between studies, the Belgian guidelines for economic evaluations in health care explicitly encourage the use of the EQ-5D instrument to describe the health states. The preference for EQ-5D was motivated by (1) the fact that this instrument has been validated in Dutch and French for Belgium, (2) an EQ-5D-3L value set derived for a part of the Belgian population (Flanders) is available, (3) it is widely used in clinical studies, enabling consistency and comparability.[25, 30] An international review of official economic guidelines indicated that the EQ-5D is the most recommended instrument for the determination of health utilities.[30]

The EQ-5D, developed by the EuroQol Group (see Box 2), is a generic HRQoL instrument commonly used for indirect utility measurement.[30, 31] It is a descriptive system for health states, encompassing 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), combined with a visual analogue scale ranging from "worst imaginable health state" to "best imaginable health state" (see Figure 1).[32, 33] Depending on the version of the instrument, each health dimension has either 3 levels of severity (EQ-5D-3L; 3L: no problems, some/moderate problems, extreme problems/unable to) or 5 levels of severity (EQ-5D-5L; 5L: no problems, slight problems, moderate problems, severe problems, extreme problems /unable to).

Each EQ-5D health state is labelled by a code, e.g. 21531, where each digit represents the severity level of a dimension. By convention, the order of dimensions is mobility, self-care, usual activities, pain/discomfort and anxiety/depression.[33]

The EQ-5D-3L has been criticized for its lack of sensitivity to small changes in health[34, 35] and its limited ability to discriminate between small differences or changes in health.[36] Also ceiling effects, meaning that mild issues with health do not have an impact on the health state description in the EQ-5D-3L, limit the instrument's responsiveness in patients with mild conditions.[32, 35-37]

Figure 1 – The EQ-5D-5L questionnaire (UK English sample version), combined with the visual analogue scale

Under each heading, please tick the ONE box that best describes your health TODAY.

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

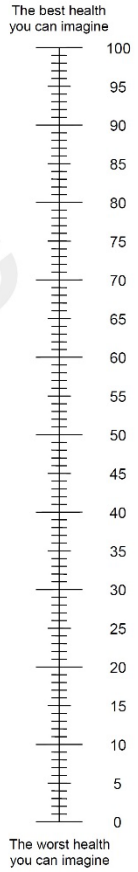
- I am not anxious or depressed ❑
- I am slightly anxious or depressed ❑
- I am moderately anxious or depressed ❑
- I am severely anxious or depressed ❑
- I am extremely anxious or depressed ❑

2

© 2009 EuroQol Research Foundation. EQ-5D™ is a trade mark of the EuroQol Research Foundation. UK (English) v1.2

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Please mark an X on the scale to indicate how your health is TODAY.
- Now, write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



3

© 2009 EuroQol Research Foundation. EQ-5D™ is a trade mark of the EuroQol Research Foundation. UK (English) v1.2

Sample

Note: Reproduced by permission of EuroQol Research Foundation

The EQ-5D-5L has been developed in 2009 to improve the discriminatory power of the instrument and reduce its ceiling effects.[32, 36, 38] The number of response categories per dimension changed from 3 to 5, increasing the number of possible health state descriptions from 243 (3^5) to 3 125 (5^5) while the questionnaire is still easy to administer in clinical trials and intuitive to patients.

No EQ-5D-5L value set exists for Belgium

An EQ-5D value set is a set of utility values for all health states that can be described with the EQ-5D, i.e. the EQ-5D-3L or EQ-5D-5L instrument. As stated in sections 1.1 and 1.2, utilities are used as weights reflecting the health-related quality of life in a particular health state on a full health–dead scale.

An EQ-5D-5L value set has already been determined for a large range of countries (such as France, England, Ireland, the Netherlands, Spain or Germany)[39-44] following a standard protocol defined by the EuroQol group, called the EuroQol Valuation Technology (EQ-VT) protocol (see section 2.1).[45, 46] A valuation study for the EQ-5D-3L has been performed for Flanders in 2003[47] but no value set specific to the EQ-5D-5L instrument is currently available for Belgium.

As an intermediary solution, to not impede research being carried out with the more recent EQ-5D-5L descriptive system, a transformation or mapping procedure was proposed by van Hout et al. (2012).[48] to derive an EQ-5D-5L "crosswalk" value set from the available EQ-5D-3L value sets. However, in order to benefit from the full potential of the EQ-5D-5L, a value set specific to the 5L questionnaire is necessary.

1.4 Study objective

The objective of this study is to derive a value set consisting of utilities for all health states that can be described with the EQ-5D-5L instrument, i.e. an EQ-5D-5L value set for Belgium, using the latest developments in direct valuation methodology and following the EuroQol Group's guidelines for EQ-5D-5L valuation studies.[45, 46] To this end, interviews were conducted in a representative sample of the population and a rigorous estimation strategy pursued to generate a Belgian EQ-5D-5L value set. The current study was requested by Sciensano (see Box 4) and co-financed by the EuroQol Research Foundation (see Box 2) that covered half of the expected study costs.

The new EQ-5D-5L value set updates the currently used EQ-5D-3L value set from 2003 with 243 potential health states to a more extensive value set with 3 125 potential health states. It is expected to be used in the following situations:

- To describe the general population's health and to compare the HRQoL values of patients with the HRQoL values of a reference group from the general population (i.e. a group with similar characteristics (age, sex) as the patient).[49-51] As the EQ-5D-5L questionnaire is included in the Belgian Health Interview Survey (HIS) since 2013 (see Box 4), reference values from the general population (i.e. population norm) can easily be generated once the set of Belgian utilities for all EQ-5D-5L health states becomes available.[50] Such comparisons could for example be used for assessing which patient populations are relatively worse off compared to others, or, in other words, have the highest unmet needs.
- As a patient-reported outcome measure (PROM) for a (more systematic) assessment of the impact of health interventions on the generic HRQoL of patients and to improve quality of care.[52, 53] By collecting data on the descriptive system of the EQ-5D-5L in (patient) populations who did and populations who did not receive the intervention, it is possible to make comparisons between the outcomes in the two groups. The generic nature of the EQ-5D makes it possible

to compare results across health conditions, treatments and care providers.

- To assess the clinical effectiveness of health interventions in terms of improvements in HRQoL, possibly combined with impact on life expectancy. Once a Belgian value set of utilities for all EQ-5D-5L health states is available, it can be used to calculate QALYs.[3, 13, 25]
- To perform cost-effectiveness evaluations of health interventions. A cost-utility analysis – which can be considered as a specific type of cost-effectiveness analysis (see Box 2) – combines the assessment of clinical effectiveness in terms of QALYs with data on the cost of the interventions.[3, 13, 25]

When different HRQoL instruments, with different health states and utility values, are used to calculate QALYs, the comparability of the analyses is jeopardized.[25] With a lack of comparability between evaluations, it is difficult to develop a consistent policy.[3, 25] Standardisation of methods, using the EQ-5D-5L instrument and its EQ-5D-5L value set for Belgium would increase the comparability between cost-utility analyses and by consequence the usefulness of such evaluations for decision making.[25]

Box 4 – The Belgian Health Interview Survey[54]

The Belgian Health Interview Survey (HIS) is a cross-sectional population survey measuring the health and health indicators of the Belgian general population. The HIS is organised every 4-5 years by Sciensano (formely the *Wetenschappelijk Instituut Volksgezondheid* (WIV) – *Institut Scientifique de Santé Publique* (ISP) and collects data from about 10 000 persons in Belgium.[55] A total of five surveys have already been conducted in 1997, 2004, 2008, 2013 and 2018. The EQ-5D-5L questionnaire is included in the survey since 2013.[56-59]

2 METHODS

The methodology for EuroQol valuation studies is standardized internationally, with extensive research on the questionnaire, valuation techniques, potential biases and quality control procedures. Despite these efforts and continuous improvements to the instrument and valuation technology, the study remains challenging, as the measurement of utility values for hypothetical health states (i.e. health states not experienced by the respondent) has many pitfalls. Methodologically, the validity and reliability of the measurement should be guaranteed. This means, for instance, that it should be assessed whether the questions are interpreted as intended. Also the risk of the interviewer bias should be tackled. Efforts are made to reduce the interviewer bias as described in the following sections (e.g. extensive training of interviewers and pilot testing). The following sections focus on the methodological aspects of this study.

2.1 Study protocol

The study follows the protocol developed by the EuroQol group, called the EuroQol Valuation Technology (EQ-VT V2.1) protocol (see Oppe et al. 2014[46] and Stolk et al. 2019[45]). This section describes the main requirements defined in this protocol and how they were applied.

2.1.1 Valuation techniques

Both composite Time Trade-Off (cTTO) and Discrete Choice Experiment (DCE) techniques were used as required by the EQ-VT protocol.[28, 45, 46] All tasks were administered using the EQ-VT software in the presence of an interviewer (see sections 2.1.2 and 2.1.3).

Composite TTO technique

The cTTO combines a traditional TTO technique for health states better than dead, and a lead-time TTO technique for health states worse than dead (see Figure 2).[60]

At the beginning of the exercise, people must choose between two lives: life A or life B. In life B, people live 10 years in the described EQ-5D-5L state and then die. In life A, people live 10 years in full health and then die. Then, according to the respondent's choice, the number of years in life A is varied while B remains fixed at 10 years. The exercise is stopped when the respondent is indifferent between life A and life B. In this case, the utility associated with the health state in life B equals the time in life A when indifferent divided by 10 (i.e. the number of years in life B), see also Box 1 for more information. For example, the evaluation on the left side in Figure 2 shows that the respondent is indifferent between 5 years in full health (indicated by x) and 10 years in the health state associated with life B, indicating a utility value of $0.5 \left(=\frac{5}{10}\right)$ for the evaluated health state.[60-62]

If the respondent prefers life A over life B even when the number of years in life A equals zero, this means that the respondent considers life B (i.e. with the described EQ-5D-5L state) as worse than dead (WTD). In this case, valuation is switched to a Lead Time-TTO technique, where a 10 years lead time period in full health is attached to both life A and life B. Again, people have to choose between life A or life B but life B lasts 20 years consisting of two periods (10 years in full health, followed by 10 years in the described EQ-5D-5L state) and life A varies between 0 and 10 years in full health according to the respondent's choice (see Figure 2). In this case, the utility associated with the health state in life B equals time in life A when indifferent (x in Figure 2, right side) subtracted by 10 years (i.e. the lead time), and then divided by 10 (i.e. the number of year in life B minus the lead time). For example, the evaluation on the right side in Figure 2 shows that the respondent is indifferent between 5 years in full health (indicated by x) and 10 years in full health followed by 10 years in the health state associated with life B. The corresponding utility value is $-0.5 \left(=\frac{(5-10)}{(20-10)}=\frac{-5}{10}\right)$ for the evaluated health state.[60-62]

Respondent can trade in time with a precision of 6 months, resulting in cTTO values that range from -1 (trading in all the lead time) to 1 (trading no time at all) in increments of 0.05, or a total of 41 different utility values.[28, 45, 60]

Only a selection of health states were evaluated by cTTO valuation and used as input in the data analysis to generate an EQ-5D-5L value set (see sections 3 and 4). As defined in the EQ-VT protocol, a set of 86 health states were evaluated. These health states were carefully selected in a simulation exercise based on their predictive power in generating a full value set.[63] They were divided in ten blocks with 10 health states per block.[44, 64] The blocked design allows to achieve a mix of states with different severities to be presented to the respondents.[63, 64] Each block included one of the five mildest state (i.e. 21111, 12111, 11211, 11121, and 11112) and the worst state (55555). Hence, each of the mildest state is included in two of the ten blocks and all respondents evaluate the worst state. The remaining 80 health states were distributed over the 10 blocks in such a way that the full severity range was more or less covered within a block while all blocks also had the same mean severity. Each block was randomly assigned to respondents and was presented in a random order. Such random selection was done automatically by the EQ-VT software.

The unconscious state

In the previous valuation study for the EQ-5D-3L performed for Flanders more than 10 years ago[47], the unconscious state was included as possible health state. However, this state was not included in the new EQ-VT 2.1 protocol.[53] In Belgium, it was deemed important for clinical research to also obtain a value for this state, and therefore the state "unconscious" has been added as an 11th state in two of the ten valuation blocks (hence approximately 20% of the respondents have valued this state). It was always the last state to be valued to avoid interference with the standard cTTO procedure in the EQ-VT protocol.

Figure 2 – The composite Time Trade-Off (cTTO): traditional TTO for states better than dead and the lead-time TTO for states worse than dead

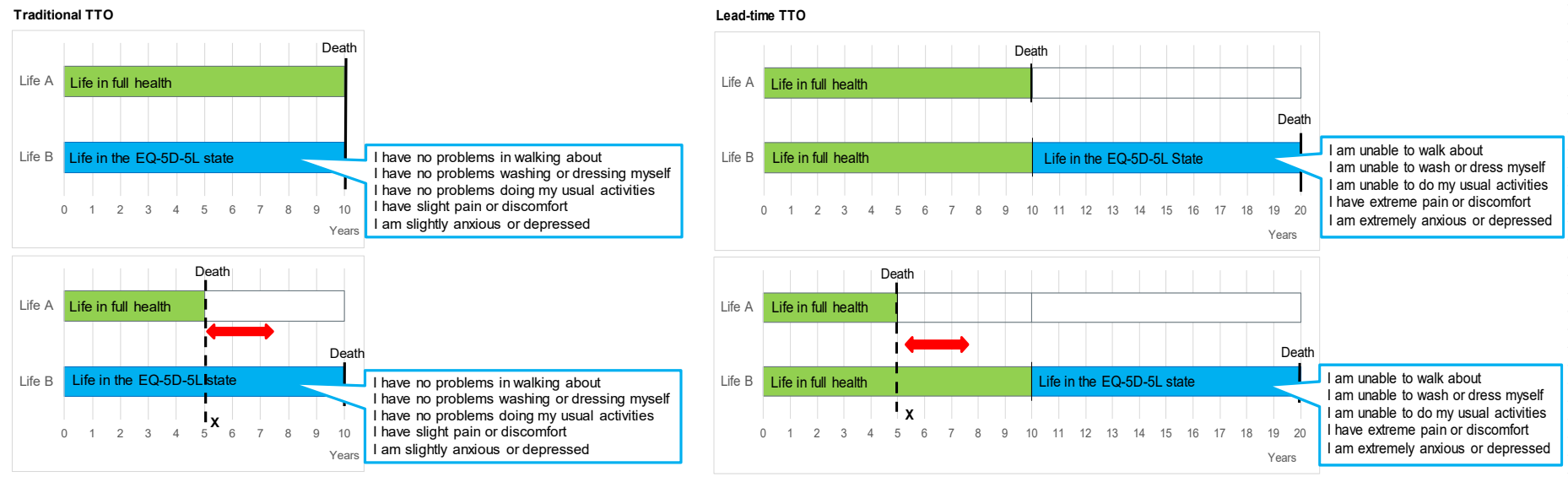
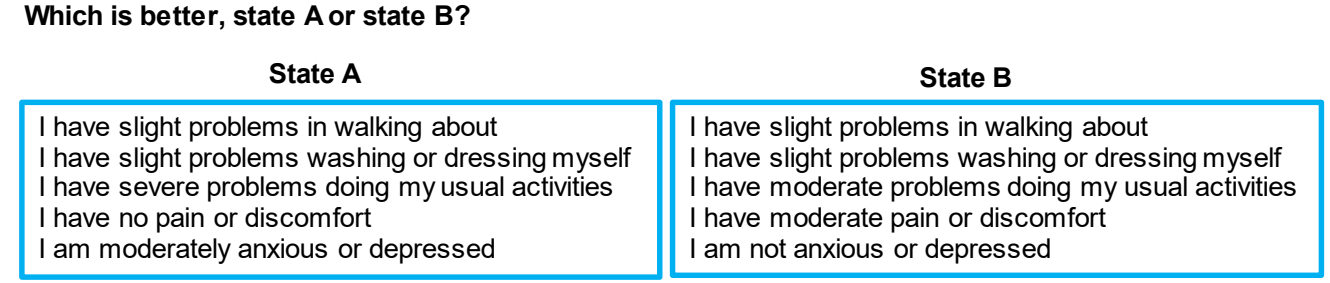


Figure 3 – The DCE technique



DCE technique

The second part of the valuation involved a discrete choice experiment (DCE), comprising seven paired choice tasks. In each choice task, respondents are asked to indicate their preference between two EQ-5D-5L states (example given in Figure 3).

As defined in the EQ-VT protocol, 196 choice pairs of EQ-5D-5L health states were divided over 28 blocks of seven pairs.[44, 64] As for the cTTO, the choice pairs were carefully selected to optimize the predictive power.[63] The health states in a choice pair were comparable in terms of their severity index and no choice pairs had to be evaluated where one alternative logically dominates the other.

Each block was randomly assigned to respondents and presented in a random order (the order of each pairs but also its position on the screen (state A or B)). Such random selection was done automatically by the EQ-VT software.

2.1.2 EQ-VT Software

To promote consistency in the way studies are performed, the EuroQol Group has developed a digital technology, i.e. the EQ-VT software that has been tested extensively. By means of a laptop, interviewers and respondents had access to the EQ-VT platform (adapted to the V2.1 protocol). The interviewer guides the respondent through the interview. The software registers all the answers as well as the time spent in the different sections of the interview (see section 2.1.3), a feature that is used in the quality control of the interviews (see section 2.1.7).

Interviewer scripts within this EQ-VT platform were translated into Dutch and French by a professional translation agency following the EuroQol Group's guidelines. While a German script was available for the study performed in Germany, such a script was not adapted to the 2.1 version of the protocol and the corresponding software and could therefore not be used for this study. Hence, for people living in the German-speaking community, the French version of the EQ-VT software had to be used and a specific bilingual German and French speaking interviewer was recruited. A paper version of the script and summary of all health states translated in German by a professional translation agency were provided to this interviewer (n = 22 interviews).

In March 2020 – before the end of the current study – the contract with the IT Company that had developed the EQ-VT platform expired and was discontinued. Therefore, a new platform was developed. To ensure similar experiences for the respondents and the interviewers between the two platforms, the new platform was essentially the same. In the current study, 48 interviews (all in French) were performed with this new platform (see also section 2.4).

During the transfer to this new platform, it appeared that in the Dutch version, one severity level of one dimension was not correctly labelled, i.e. level 2 in the mobility dimension "Een beetje problemen met *lopen*" was used instead of "Een beetje problemen met *rondwandelen*". Because all Dutch interviews were conducted with the old software, they made use of this wrong label. Such problem was kept in mind during data analysis (see also section 4.2.1), but the impact is expected limited because from September 2018 onwards, all interviewers also had a printed questionnaire with the correct label to help the respondent during the interview to correctly locate the health state described on the screen in the EQ-5D-5L questionnaire and to facilitate assessing the severity of the state (see also the section below on the number of interviewers and additional trainings).

2.1.3 Structure of the interview

In accordance with the EQ-VT protocol, interviews were performed face-to-face using the EQ-VT software (computer-assisted face-to-face interviews). Each interview followed the same process:

1. General welcome
2. Introduction and respondents' characteristics
    - Self-reported health of the respondent on the EQ-5D-5L questionnaire

- Self-reported health of the respondent on the EuroQol visual analogue scale (EQ-VAS)
- Background questions on socio-demographic information

3. Composite Time Trade-Off (cTTO) experiment (see also section 2.1.1)

   - Instructions and examples of cTTO tasks, including two exercises on being in a wheelchair and three exercises on EQ-5D health states (*training phase*).
     The aim of the training phase is to make the respondent familiar with the cTTO valuation and the iterative sequence of the task and to illustrate both the traditional TTO and the lead-time TTO.[45] In the wheelchair example, the respondent is asked to value "being in a wheelchair" using the cTTO elicitation procedure (see section 2.1.1). If the respondent values "being in a wheelchair" as better than dead (traditional TTO), he/she is asked in a second exercise to imagine a health state much worse than "being in a wheelchair" to move the evaluation to the lead-time TTO. If, on the other hand, the respondent values "being in a wheelchair" as worse than dead (lead-time TTO), he/she is asked in a second exercise to imagine a health state much better than "being in a wheelchair" to move the evaluation to the traditional TTO. After the wheelchair example, the respondent is asked to value three practice health states of varying severity.
   - cTTO valuation of 10 EQ-5D-5L states and possibly the unconscious state (*valuation phase*)
   - cTTO debriefing (*feedback phase*)
     Respondents answer three feedback questions with regard to the clarity of the question and complexity of the valuation exercise.[61]
   - Feedback module (*feedback phase*)
     In the feedback module, each respondent was presented a ranking of the ten health states ordered by their responses in the cTTO valuation, with the highest valued state on top, the lowest at the bottom and states valued as 'equal' placed side by side. The respondents were then asked to confirm their choices or to identify which health states they considered to be incorrectly ranked. No option for re-valuing the selected states by the respondent was allowed.[41, 45, 65-67]

4. Discrete Choice Experiment (DCE) (see also section 2.1.1)

   - Instructions on DCE tasks (*training phase*)
   - DCE valuation of 7 pairs of EQ-5D-5L states (*valuation phase*)
   - DCE debriefing (*feedback phase*)
     Respondents answer three feedback questions with regard to the clarity of the question and complexity of the valuation exercise.[61]

2.1.4 Target sample

According to the EQ-VT protocol, the target sample must consist in 1 000 individuals who completed a valid interview (i.e. interviews that passed the quality control, as described in section 2.1.7). Such sample size was determined to ensure sufficient precision in the estimation for the TTO-tasks.[46, 63]

2.1.5 Number of interviewers

The interviewers play an essential role in guiding the respondent through the interview.[28] Research has shown that a one-to-one setting allows an interviewer to provide detailed instructions and feedback where appropriate, improving the accurateness and quality of cTTO valuation data.[45, 61] The demands placed on the interviewers are high and they are trained before working in the field (see section 2.1.6).

In order to achieve high data quality, the EQ-VT protocol stipulates that no less than 8 and no more than 14 interviewers should be recruited to ensure a maximum and a minimum number of interviews per interviewer, i.e. between 70 and 130 interviews per interviewer. A total of 11 interviewers were therefore recruited at the beginning of the study (see also section 2.4.1). In addition, the quality of the conducted interviews is monitored (see section 2.1.7).

2.1.6 Training and pilot tests

Data collection was preceded by a training of all interviewers, including a pilot study to reduce interviewer's bias. Recruitment of interviewers and management of the work in the field was subcontracted to Profacts, a market research agency.

In line with the EQ-VT protocol, the EuroQol group trained the principal investigators of the study (i.e. 2 members of the KCE and 2 members of Profacts during a 2-days training in Rotterdam) and the principal investigators subsequently trained the interviewers during a one-day training.

The training of interviewers included at least 5 pilot interviews. Results of the pilot tests were discussed with the EuroQol group and a feedback was given to the interviewers before the start of interviews in the field.

2.1.7 Quality control

The quality control process took place after each block of around 10 interviews per interviewer, in collaboration with the EuroQol team and the subcontractor that was responsible for the field work (+/- every 2 weeks). This quality control is described in the studies of Ramos et al. (2017) and Stolk et al. (2019) and the main elements are summarized in this section.[45, 68]

In the quality control procedure, compliance with the protocol was reviewed and an interview was flagged as being of potentially poor quality if one of the following criteria were met:

- The worse than dead (WTD) task was not explained in the wheelchair examples during the training phase (see the interview guide in section 2.1.3);
- The time spent on explaining the cTTO task in the wheelchair examples was less than 3 minutes;
- A respondent spent less than 5 minutes to complete the 10 cTTO tasks; or
- The value for state 55555 was not the lowest and it was at least 0.5 higher than that of the state with the lowest value.

If 4 or more out of the 10 first interviews of an interviewer were flagged, these interviews had to be removed and a retraining of the interviewer was required. After a further 10 interviews, performance and compliance were re-evaluated. If again 40% or more interviews were flagged, the interviewer had to be excluded and all his/her interviews had to be removed from the sample.

During the quality control, the following elements were also discussed: few WTD responses, clustering of responses at certain values (spikes), inconsistencies involving to worst EQ-5D-5L health state (55555), or low values for mild states.

The subcontractor had to provide each interviewer with a summary of their "performance" after each quality control and hold a face-to-face contact in case of highlighted problems (with eventually a re-training, and/or an exclusion).

2.2 Sampling process

2.2.1 Inclusion and exclusion criteria

Data were collected in a sample of the Belgian population. All people aged between 18 and 100 years old residing in Belgium and listed in the National Register were considered. The random selection was done by the National Register using the sampling procedure described in section 2.2.2. The following type of residents were excluded from the sampling:

- Residents without a physical address or not listed in the National Register (e.g. homeless, irregular migrants)
- People in prison

2.2.2 Sampling procedure

As the study objective was to establish an EQ-5D-5L value set for Belgium (see also 1.4), we wanted our sample to be representative of the Belgian population, although the EQ-VT protocol does not require representative sampling. At the same time, we wanted to keep the study feasible in terms of time frame and resources. A convenience sample of participants that e.g. respond to a public campaign are less suited for this purpose as the selection probability of each respondent is usually unknown, and this makes inference about the population more difficult and bias prone.[69] In contrast, a random sample of the Belgian population based on a known probability of inclusion allows for much better inference about the population.

We chose a multistage, stratified, cluster sampling with unequal probability design to construct our sample. Although it is a more complex sampling method than simple random sampling, it comes with several advantages. As already mentioned, one advantage is in logistics. For single random sampling where one just randomly selects from the Belgian population, interviews risk being geographically scattered. A clustered and stratified design allows for a larger sample to be interviewed in the same time frame and with similar resources than a simple random sampling would. Stratification also brings the advantage of having some control on ensuring sufficient representation of the population on certain variables like age and sex. At the same time, stratification can increase efficiency of sample estimates by reducing variability (and thus standard errors) compared to random sampling without stratification.[69] In contrast, clustering and weighting can increase variability compared to simple random sampling. However, the advantages of feasibility and stratification were considered to outweigh the disadvantages of the clustering and weighing aspects of the design. The sampling procedure was similar in nature to the procedure used by e.g. Sciensano for the Belgian Health Interview Survey.[70]

To assure a sufficient coverage of the Belgian territory, municipalities were sampled as clusters within provinces taking into account unequal probability of inclusion based on population size within the provinces and municipalities. This to find a balance between sufficient representativeness and feasibility (see the sections below for details). Additionally, we further stratified sampling on age categories and sex to obtain a sufficient balance in respondent characteristics. We chose age and sex as strata because of the availability at the National Register for sampling and because for the later application of the EQ-5D-5L value set, these variables were deemed important to have weighed proportionally on the construction the EQ-5D-5L value set. Age was categorised in eight categories to avoid too small strata ([18,30), [30,40), [40,50), [50,60), [60,70), [70,80), [80,90), [90,100)).

The EQ-VT protocol necessitates about 1 000 successful interviews to be able to construct the EQ-5D-5L value set in a reliable way (see section 2.1.4). We assumed a conservative estimated response rate of 10% in deciding how many potential participants to sample. For each interview, 10 potential participants were sampled at random in the same stratum. A total of 10 000 potential participants were then sampled.[a]

2.2.2.1 Province

The sampling uses province population to determine the number of interviews per province. Brussels region was considered separately. Whenever we use ‘province’ in the text below, Brussels region is considered as a separate entity.

[a] Note that not all 10 000 potential participants were contacted. For each required interview in a specific stratum, only one or a few potential participants needed to be contacted in order to find someone willing to participate in the survey. Exceptionally, all 10 potential candidates were contacted without successfully recruiting a respondent.

Based on the legal Belgian population on 1 January 2017 (as retrieved from Statistics Belgium), Table 2 details the number of interviews per province (column 'Rounded number of interviews').[71]

Table 2 – Sampling at the province level

Level	Population N	Population %	Number of interviews	Rounded number of interviews	Selection probability of an individual in %
Brussels region[a]	917 907	10.2%	101.68	100	0.0109%
Antwerpen	1 467 210	16.3%	162.53	160	0.0109%
Limburg	706 089	7.8%	78.22	80	0.0113%
Oost-Vlaanderen	1 205 461	13.4%	133.53	130	0.0108%
Vlaams Brabant	899 110	10.0%	99.6	100	0.0111%
West-Vlaanderen	973 765	10.8%	107.87	110	0.0113%
Brabant Wallon	315 176	3.5%	34.91	40	0.0127%
Hainaut	1 058 246	11.7%	117.23	120	0.0113%
Liège	875 791	9.7%	97.01	100	0.0114%
Luxembourg	219 354	2.4%	24.3	20	0.0091%
Namur	389 340	4.3%	43.13	40	0.0103%

[a] *Although not a province, Brussels region is considered separately.*

2.2.2.2 Municipality

Within each province, 10 interviews were clustered within selected municipalities. Following the algorithm used by the Belgian HIS in 2013 for selecting primary sampling units (PSUs), 100 municipalities (100 groups of 10 interviews) were chosen at random.[70] Although the algorithm chose municipalities at random within a province, it took into account the relative population size to assure that both smaller and larger municipalities were part of the selection. As a consequence, larger municipalities could be chosen multiple times.

The steps of the algorithm are described below. In the second column, a fictional province with 10 municipalities is used to illustrate the steps. We use a fictional province as a real province would make the example too wieldy.

	Step	Running example
1.	For each stratum (province + Brussels) all municipalities were listed by population size (from large to small).	
2.	A list was made with the cumulative population size equalling the population size of the municipality + the population size of the municipalities listed before.	

Municipality	Population	Cumulative population
A	23 320	23 320
B	13 805	37 125
C	13 046	50 171
D	12 128	62 299
E	9 236	71 535
F	8 924	80 459
G	8 587	89 046
H	6 772	95 818
I	6 378	102 196
J	6 227	108 423

3.	An interval (I) was calculated being equal to the total population size of the stratum (T) divided by the number of groups (N) of 10 interviews to be selected: $I = T/N$. Each group is attributed to a municipality	Continuing our example, suppose 20 interviews need to be conducted in the fictional province. This corresponds to two groups of 10 interviews or two municipalities that need to be selected. The interval then is $I = \frac{108\,423}{2} = 54\,212$.
4.	An integer random number (R) is selected where $0 < R < I$.	In our example of the fictional province, suppose the random number is 15 457.
5.	A list of N values (x_i) is generated (one for each group of interviews per province), starting with R and adding each time a multiplier of I: $x_i = R + n \times I$ with n an integer so that $0 \leq n < N$. We refer to these values as $\boldsymbol{interval_{interviewgroup}}$.	
6.	For each interview group J per province, a municipality is selected by comparing the $interval_{interviewgroup\ J}$	The selection of **the municipality for the first group of 10 interviews** (Municipality A):

Interview group	Calculation	$\boldsymbol{interval_{interviewgroup}}$
0	(0 × 54 212) + 15 457	15 457
1	(1 × 54 212) + 15 457	69 669

calculated in the step 5 with the cumulative population ($cumpop_I$) calculated in step 2 for each municipality I. The municipality is chosen for which: $min(cumpop_I - interval_{interviewgroup\ J}) > 0$.

Municipality	Cumulative population	$Interval_{interview\ group\ 0}$	Difference cumulative population and $interval_{interview\ group\ 0}$	
A	23 320	15 457	7 863	smallest difference of interval < cumulative population
B	37 125	15 457	21 668	
C	50 171	15 457	34 714	
D	62 299	15 457	46 842	
E	71 535	15 457	56 078	
F	80 459	15 457	65 002	
G	89 046	15 457	73 589	
H	95 818	15 457	80 361	
I	102 196	15 457	86 739	
J	108 423	15 457	92 966	

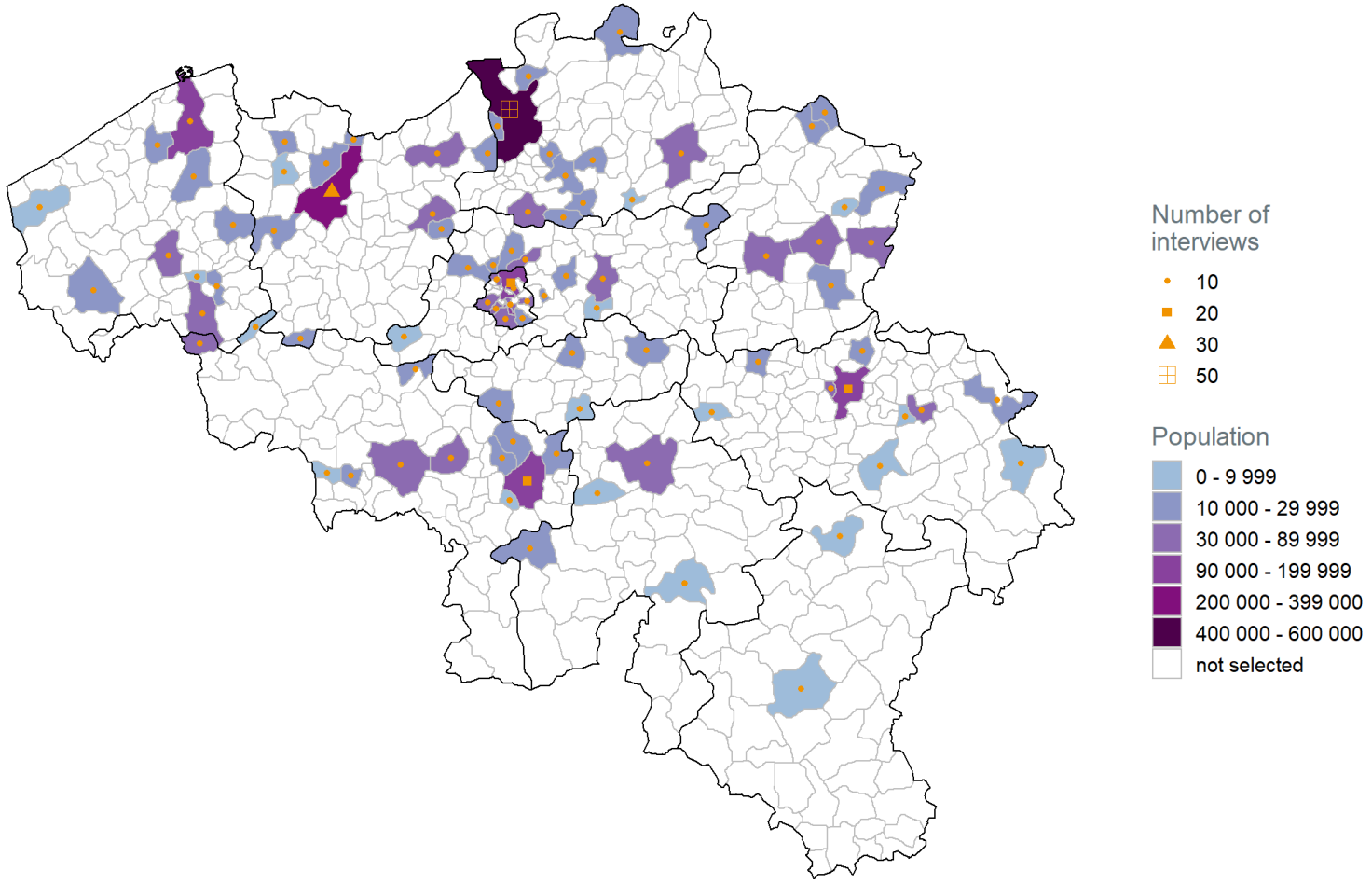
The selection of the municipality for the second group of 10 interviews (Municipality E):

Municipality	Cumulative population	$Interval_{interview\ group\ 1}$	Difference cumulative population and $interval_{interview\ group\ 1}$	
A	23 320	69 669		
B	37 125	69 669		
C	50 171	69 669		
D	62 299	69 669		
E	71 535	69 669	1 867	smallest difference of interval < cumulative population

	F	80 459	69 669	10 791	
	G	89 046	69 669	19 378	
	H	95 818	69 669	26 150	
	I	102 196	69 669	32 528	
	J	108 423	69 669	38 755	

The result of applying the algorithm to all provinces + Brussels is shown in the Figure 4 below.

Figure 4 – Selected municipalities (PSU)



2.2.2.3 Age and sex

Based on the population distribution in each province, interviews in the municipalities were attributed to subgroups based on 10 year age categories and sex. To obtain a representative sample for the Belgian population, the interviews were allocated using the following algorithm:

1. If there are more interviews per age and sex category than there are selected municipalities in the province, allocate from each age category and sex one interview to each of the selected municipalities. E.g. in the province “Antwerpen”, 12 municipalities were selected. For the province “Antwerpen”, there are 13 interviews in the female [30,40) category to allocate to these 12 municipalities. In this step, each municipality is allocated one interview in the female [30,40) category. If there are less interviews in the age and sex category than municipalities, go to step 3.
2. Repeat step 1 until there are less interviews in the age and sex category than there are selected municipalities in the province. In the example in the previous step, there is only one interview left, so the algorithm advances to step 3.
3. Randomly assign the remaining interviews in the age and sex category to municipalities in the province. In the example of the previous steps, the remaining interview in the female [30,40) category is randomly assigned to one of the 12 municipalities in the province “Antwerpen”.

Figure 5 and Figure 6 show the number of allocated interviews per age category and sex for each province, alongside the population distribution in the province. Rounding of the proportions caused a slight deviation of the total number of interviews by province. Based on the largest rounding errors, the number of interviews for six age – sex - province combinations were manually adjusted by adding or reducing the number of interviews by 1.

Figure 5 – Female population distribution by province and age of the legal Belgian population on 1 January 2017 (left panel); number of planned interviews for female participants by province and age category (right panel)

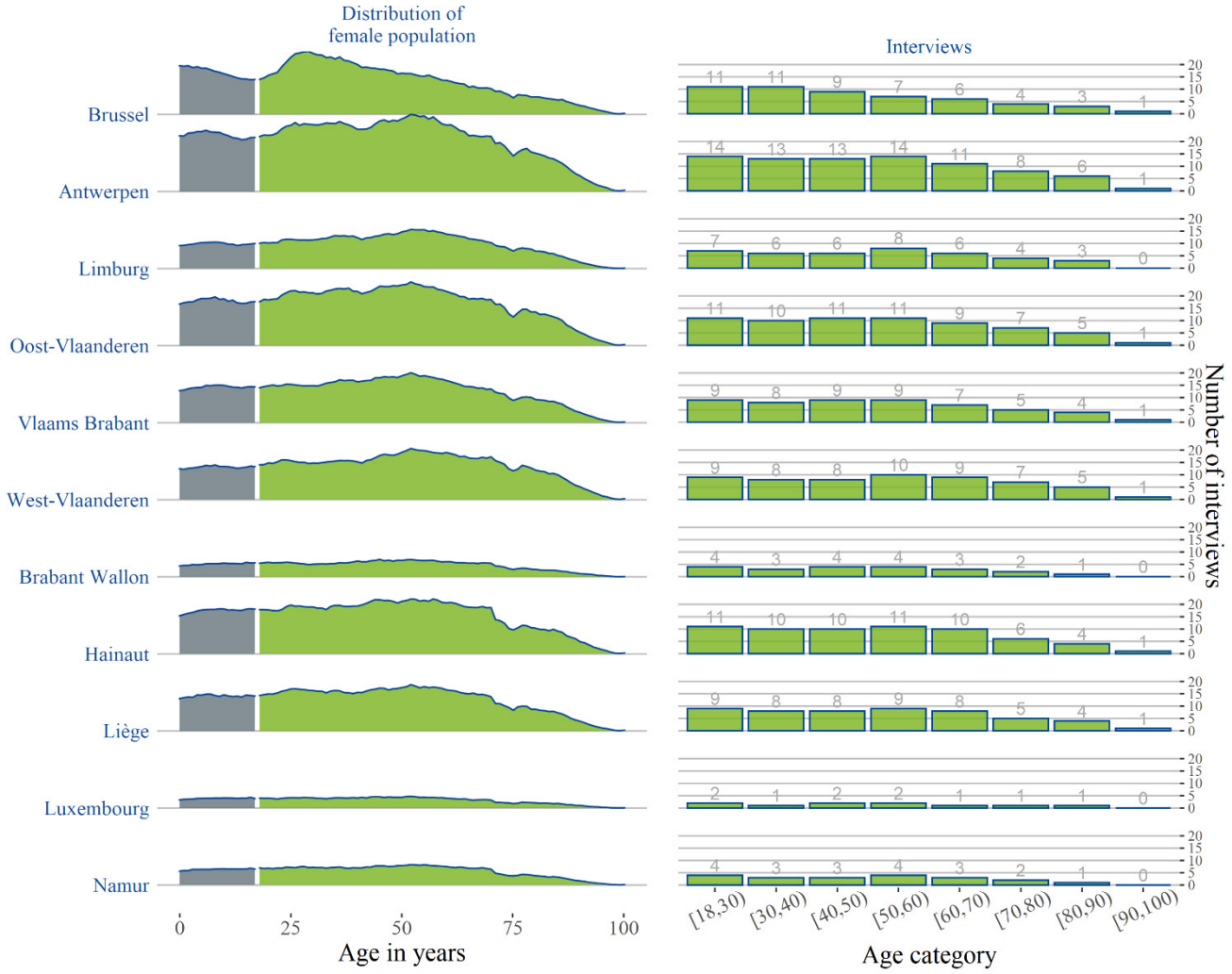
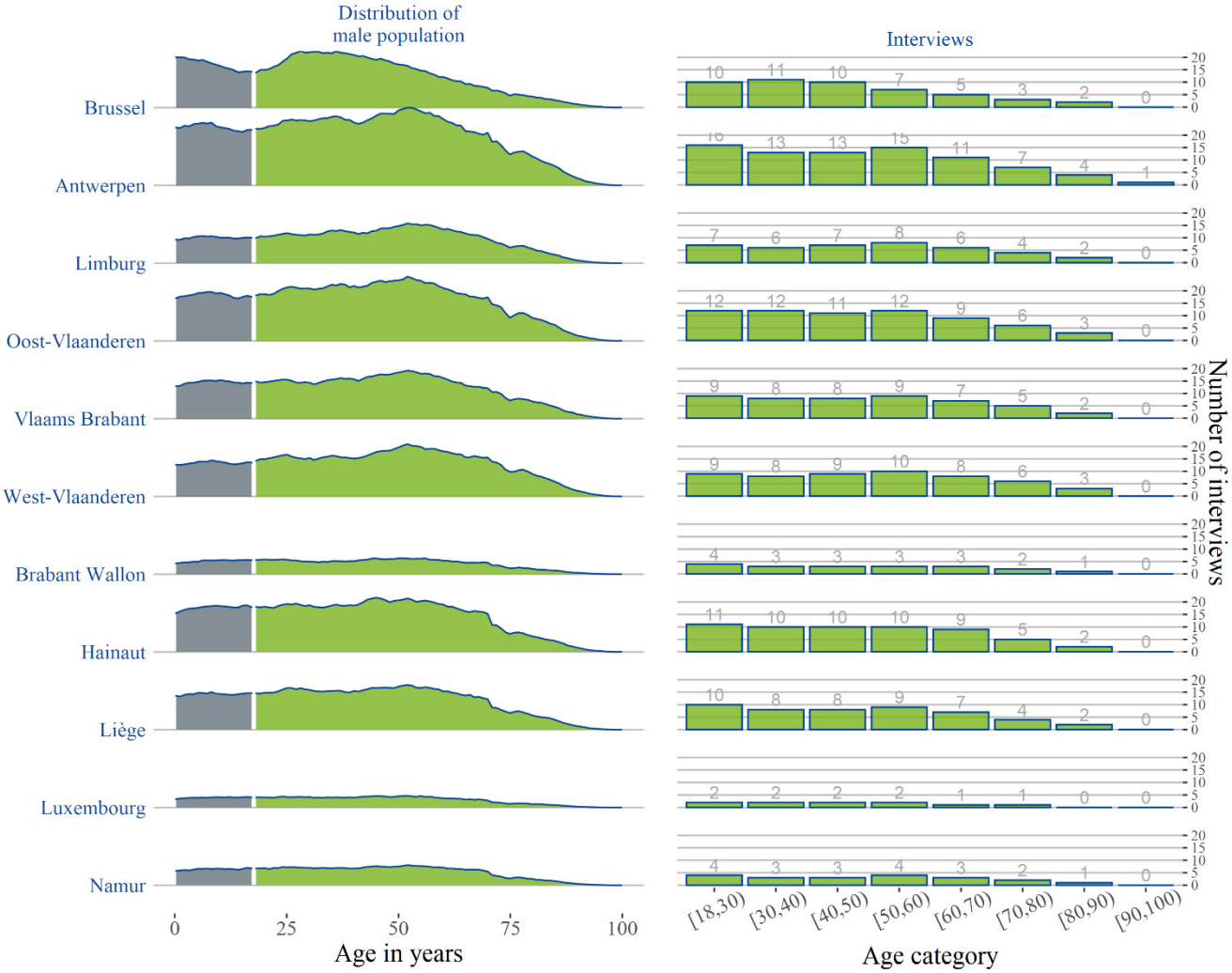
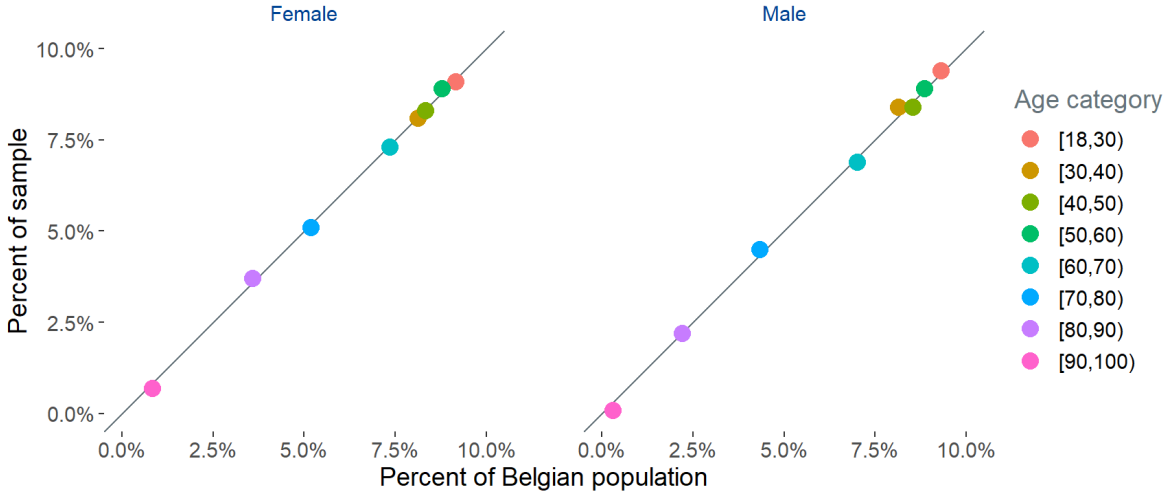


Figure 6 – Male population distribution by province and age of the legal Belgian population on 1 January 2017 (left panel); number of planned interviews for male participants by province and age category (right panel)



To verify that the overall age category and sex distribution of the sampled individuals corresponded to the distribution of age category and sex of the Belgian population, Figure 7 compares the agreement of the proportions in sample and population.

Figure 7 – Age and sex groups proportions of the Belgian population versus sample



2.3 Study approval and privacy protection aspects

The study research protocol was approved by the Ethical Committee of the University hospital of Gent (B670201835610).

The approval of the National Register Sectorial Committee was also obtained for the use of the National Population Register during the recruitment of participants. The use of data from the National Register imposed the use of a third party for the work on the field (legal requirement of the National Register Sectorial Committee). Recruitment of interviewers and management of the data collection work in the field were therefore outsourced to an external team via a public tendering process. The selected subcontractor was Profacts, a market research company.

To ensure the respect of the law of 8 December 1992 on the privacy protection concerning the processing of personal data, as amended by the Act of 11 December 1998, a contract on data protection was signed by Profacts, including for example the obligation to not use data for other purposes than this study and to destroy these data at the end of the study.

2.4 Data collection

2.4.1 Initial recruitment of interviewers

As stated in section 2.3, recruitment of interviewers and management of the work on the field was outsourced to Profacts, a market research company. According to the EQ-VT protocol, Profacts had to recruit between 8 and 14 interviewers (see section 2.1.5). A total of 11 interviewers were therefore recruited at the beginning of the study: 6 Dutch-speaking interviewers, 4 French-speaking interviewers, and 1 bilingual French- and Dutch-speaking interviewer. The main selection criterion for interviewers was not their educational background, but their experience with handling complex interviews. The Dutch-speaking interviewers were assigned to the Flemish region, the French-speaking interviewers were assigned to the Walloon region and the bilingual French- and Dutch-speaking interviewer was assigned to the region of Brussels-Capital.

Because of the quality control and difficulties encountered during the work in the field, additional interviewers were recruited during the data collection (see section 2.4.3). Moreover, at the beginning of the study, Profacts was not able to recruit a bilingual German- and French speaking interviewer (for people of the German-speaking community in Belgium). Such interviewer was also recruited at a later stage (see section 2.4.3).

2.4.2 Recruitment of respondents

The National Register transmitted to Profacts the contact details (name and address) of people randomly selected according to the sampling procedure described in section 2.2.2. Interviewers then received from Profacts the contact details of respondents living in their assigned area.

To increase the response rate and to facilitate the work of interviewers, a letter explaining the project was sent to all people selected in the sample prior to a contact by phone or in person. A link towards a website explaining the study and providing the possibility to already give their consent to participate was included in the letter.

Next, a first contact with respondents was made either by phone or in person (door-to-door), during which the willingness to participate in the study was discussed. In case of positive response, the interview was either conducted directly or scheduled at a later moment. Each respondent had to sign an informed consent document, preserved by Profacts. It was stipulated that Profacts had to make at least 3 contact attempts to recruit respondents in case it was difficult to get in touch, with at least one attempt after 17:00 or during the weekend.

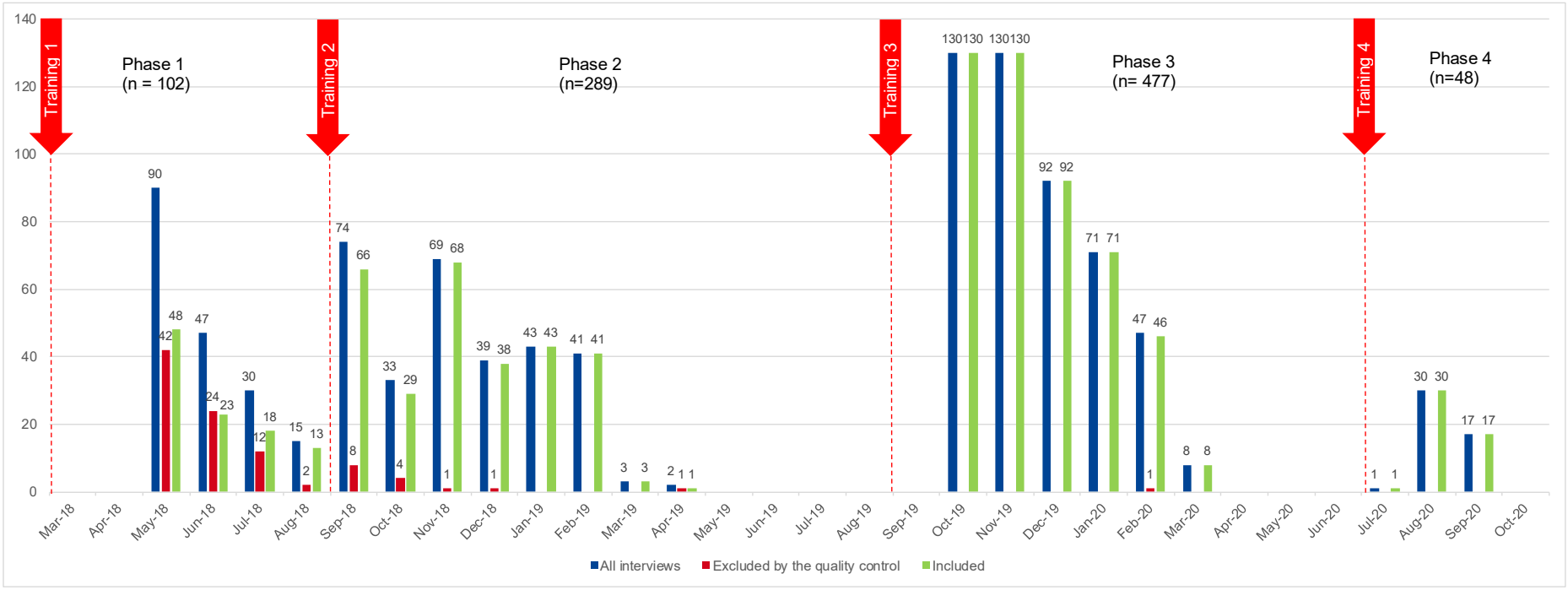
At the start of the study, no financial or other incentive was given to respondents. As it became clear that it was difficult to recruit participants, it was decided to provide a compensation of €15 to the respondents.

2.4.3 Data collection time frame and recruitment of new interviewers

Data collection was initially foreseen between 1 March 2018 and 31 August 2018 (i.e. a 6 month period). Nevertheless, in practice, the data collection process ran from 1 May 2018 to 30 September 2020. While the initial training of interviewers was performed on 5-6 March 2018, data collection in the field could only start from May 2018 due to an unexpected long delay in obtaining data from the National Register. Then, difficulties in recruiting participants, combined with a higher than expected loss of interviews in the quality control process, slowed down the progress of the study. Finally, while the study was nearing completion in February 2020, the COVID-19 crisis further extended the deadlines.

The pace of the study, the number of interviews, as well as the impact of training and quality control is described in Figure 8.

Figure 8 – Pace of the study, number of interviews and impact of training and quality control



N= number of interviews included (i.e. that passed the quality control)

As shown in Figure 8, the data collection for the Belgian EQ-5D-5L valuation study can be divided in 4 phases, each phase beginning with a new training of interviewers.

- Phase 1: The first training of interviewers, as foreseen in the EQ-VT protocol was given on 5 March 2018 for Dutch-speaking interviewers and on 6 March 2018 for French-speaking interviewers. Nearly half of the interviews (44%) during this period did not pass the quality control (Figure 8).
- Phase 2: Because the quality of interviews was judged insufficient, despite the feedback given after each block of about 10 interviews, an additional one-day training was organized on 30 August 2018 for Dutch-speaking interviewers and on 31 August 2018 for French-speaking interviewers. At that moment, 4 interviewers did not pass the quality control and were excluded (as well as all their interviews) and 4 new interviewers were enrolled and trained. It was also decided that each interviewer would receive a printed version of the EQ-5D-5L questionnaire in their language to help respondents to correctly locate

the health state described on the screen in the EQ-5D-5L questionnaire and so to facilitate the assessment of a state's severity.

In October 2018, Profacts also recruited a bilingual French- and German-speaking interviewer to work in the two selected municipalities belonging to the German-speaking community. This interviewer was trained on 24 October 2018. However, because of health problems, the interviewer was never able to start the field work (no interview performed). This interviewer is therefore not included in the flow chart described in Figure 9.

- Phase 3: In May 2019, because the pace of the study was judged insufficient and because Profacts was no longer able to keep enough interviewers active in the field to successfully carry out the rest of the study, it was decided to temporarily halt the study and to recruit and train new interviewers (including a bilingual French- and German-speaking interviewer). Such a choice was done in consultation and in agreement with the EuroQol group. Because the limit of 14 interviewers required in the EQ-VT protocol was exceeded, a more intensive training and follow-up was decided to reduce as much as possible interviewers' bias. This intensive training was given in September 2019 and included:
  - A one-day training, divided in 2 groups (French-speaking and Dutch-speaking interviewers). During this training, interviewers also performed one pilot.
  - After the first one-day training, three additional pilot tests per interviewer were done on family and friends (at the interviewer's convenience, e.g. at home). The results of the 4 pilot tests were assessed for compliance with the quality control requirements and individual feedback was provided.
  - A second day of training, divided in 2 groups (the same groups as the first training day), consisted in (1) a debriefing on the 4 pilot tests already conducted, (2) an additional 6 pilot test on a sample of interviewees, which were monitored by researchers from Profacts and the KCE and (3) a final debriefing. The dropout rate (i.e. the percentage of interviews flagged in the quality control process, see section 2.1.7) on these pilots had to be ≤ 30% before an interviewer was authorized to work in the field. This means that in case of a higher dropout rate, additional pilots were done.
- Phase 4: Finally, in March 2020, because of the COVID-19 pandemic and the imposed lock-down, it was decided, in consultation and in agreement with the EuroQol group, to stop the data collection before having reached the target of 1 000 and to only conduct a few number of interviews in the region of Brussels and in French-speaking Brabant (province Brabant-Wallon) in order to increase the representativeness of the sample (not enough interviews were performed in these two areas). These last interviews could only be conducted once the lockdown measures were released (i.e. from July 2020). Three interviewers working in these area were retrained on 7 July 2020 (during a 2 hours training) to ensure they were ready to work again in the field after the break imposed by the lockdown. In order to carry out these last interviews in respect with the legal protective measures, 2 choices were proposed to the respondents: performing the interview at distance via Microsoft Teams (n = 5 interviews) or face-to-face (n= 43 interviews) by wearing a mask and respecting distancing and other protective rules. The resurgence in the number of COVID-19 cases and the second lockdown led to the termination of the study. At that time, 916 interviews were performed.

A flow chart describing the exclusions and additional recruitments of interviewers, as well as their training is shown in Figure 9. Throughout the study, a total of 22 interviewers have conducted fieldwork, leading to an average of 42 interviews per interviewer. Figure 10 presents the distribution of the number of interviews per interviewer.

Figure 9 – Flow chart on the change over time in the number of interviewers

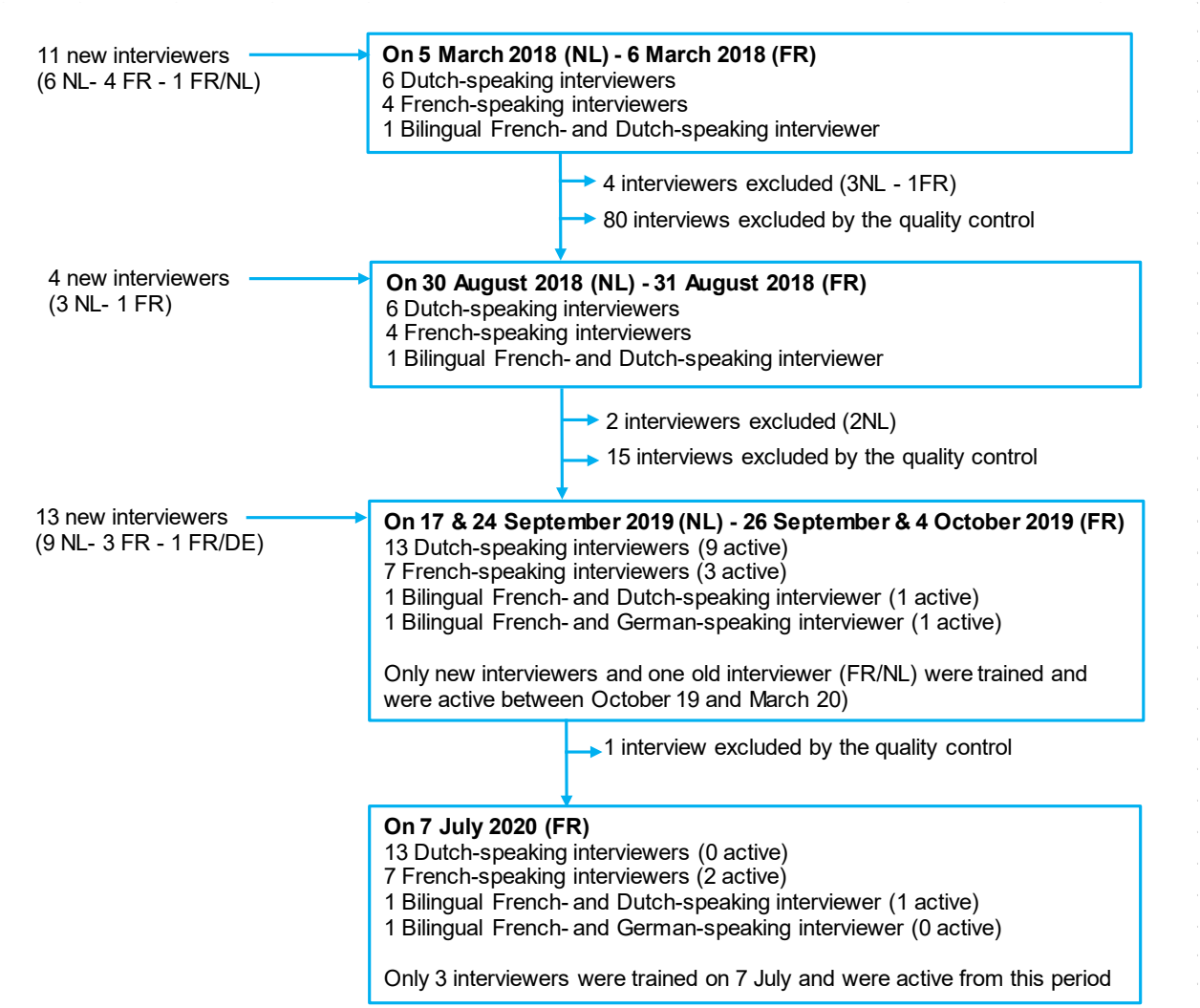
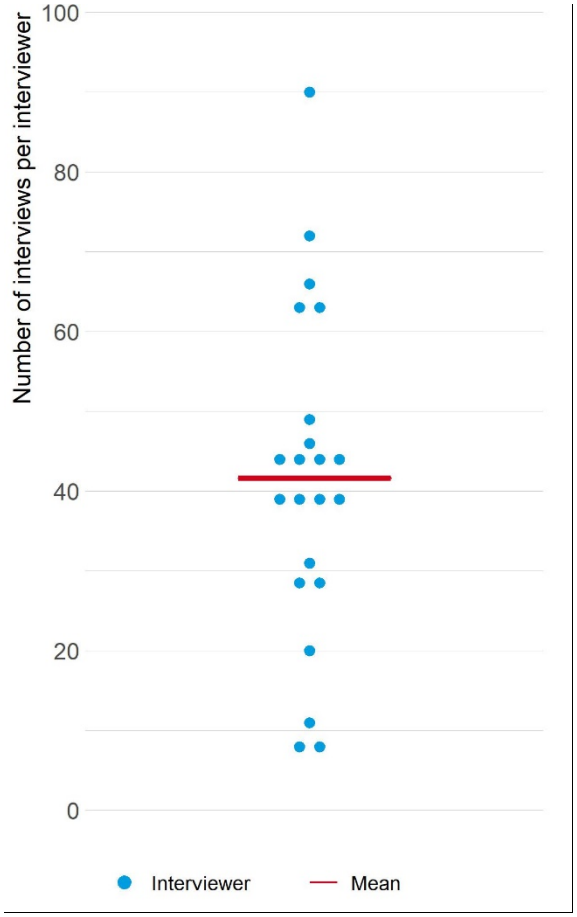


Figure 10 – Distribution of the number of interviews per interviewer



3 DATA ANALYSIS

The objective of the data analysis is to create a value set consisting of utility values for all 3 125 health states defined by the EQ-5D-5L system, which is representative for the general Belgian adult population. To achieve this objective, we need to decide on a regression model that generates utility values for all 3 125 health states based on the available valuation information from the collected interviews (see section 2.1.1).

The data analysis proceeded in four stages. First, a modelling dataset was created by excluding unusable and implausible data that were not removed in the EuroQol quality control process (see section 2.1.7 and 2.4). The representativeness of the resulting sample was assessed. Second, an extensive range of potential regression models was estimated. Third, each of the estimated model was evaluated and a set of selection criteria was applied to decide on the final regression model. Fourth, using the final regression model, the EQ-5D-5L value set for Belgium was generated.

3.1 Data preparation and representativeness

3.1.1 Creation modelling dataset

Starting from the 916 interviews that passed the EuroQol quality control process, the sample was further reduced to obtain the modelling dataset consisting of 892 respondents with 8 920 cTTO observations for the 86 health states, 189 cTTO observations for the unconscious state and 6 244 DCE observations for 392 health states assigned to 196 choice pairs.

Full sample

Each interview can be related to an anonymized id from the National Register so that compliance with the sampling plan (see section 2.2.2) can be assessed and post-stratification weights can be computed (see section 3.1.2). In the dataset, it occurred three times that two interviews referred to the same person id. Based on the age and sex of the respondent it was possible to identify the three interviews with a correct person id. However, it was not possible to correctly link the other three interviews to the id of the interviewed person. Without a valid person id, it is not possible to assign a post-stratification weight and these three interviews become unusable. The remaining sample consist of 913 respondents and is referred to as the full sample.

Modelling sample

In order to obtain the modelling sample, additional data cleaning was performed. Based on the cTTO valuation, we excluded respondents whose responses could justifiably be considered implausible. The exclusion criteria were drawn from previous valuation studies and exclusions were minimised to avoid potential bias in the results from researcher judgements.[42, 43, 60, 72-75] In total, 21 interviews were excluded, based on the following rules:

1. Respondents whose reported utility values are positively associated with the severity of the evaluated health states (N = 2 respondents). Such positive association would indicate a preference of the individual for deteriorating health. The severity of a health state is defined as the level sum score (i.e. health state 11111, indicating full health has a severity of 5=1+1+1+1+1, while the worst health state 55555 has a severity of 25=5+5+5+5+5). The positive association is evaluated for each respondent using a linear regression with the utility value as dependent variable and the severity as explanatory variable.
2. Respondents who gave all 10 evaluated health states the same utility value (N = 15 respondents).
3. Respondents who gave the same utility value to the mildest evaluated health state (with severity 6) and the worst evaluated health state (with severity 25), while the utility values for other health states are varied (n = 4 respondents).

Other exclusion rules applied in certain valuation studies were not considered. First, in deviation from the EQ-VT protocol and valuation studies in e.g. France[40], Germany[39] and Ireland[41], health states flagged in the feedback module (see section 2.1.3) were not excluded. It concerns 361 observations – about 4% of the data remaining after applying exclusion rules 1 to 3 (see above) – that were flagged by 241 respondents – about 27% of the sample. These fractions are in line with previous research.[65] The main

reason for not excluding data based on the feedback module was to not further reduce the sample size, which was already smaller than targeted given the discontinuation of the data collection due to the COVID-19 crisis (see section 2.4.3). This decision was taken in consultation and in agreement with the EuroQol group. The robustness of our results with regard to this choice will be tested (see section 3.4).

Second, we did not systematically exclude inconsistent valuation of logically dominated states in addition to the exclusion rules already laid out above.[72] A health state logically dominates another health state if it is better in at least one dimension and no worse in all other dimensions (e.g. health state 22222 logically dominates health state 32223) and hence a rational person would prefer the former over the latter. Given the complexity of the valuation exercise, we did not further restrict based on the dominance/rationality principle.

In line with most other valuation studies, no further restrictions were imposed based on the DCE data.[39, 40 , 41-43, 76] Potential exclusion criteria could relate to the occurrence of suspicious answer patterns in the DCE data, i.e. always choice A, always choice B or alternating choice A-B (or B-A). Such patterns occurred in 21 interviews and given this low number could reflect genuine preferences.

3.1.2 Representativeness

Post-stratification weights

Both for the modelling sample of 892 respondents and for the full sample of 913 respondents, post-stratification weights were calculated. These weights were used to adjust the estimation to correct for differences between the planned and realised interviews and subsequently to obtain preference values representative for the Belgian population. The post-stratification weights were calculated as follows:

1. We checked for 'empty strata', combinations of province, age category and sex where interviews were planned but where, in the end, no interviews were conducted. We found 7 out of 151 original strata to be empty. For the calculation of post-stratification weights, the empty strata need to be merged with other non-empty strata to create post-survey strata. We chose to define the post-survey strata as close as possible to the original sampling and merge the empty stratum with the stratum with the same province and sex, but the age category just below, so the interviews in that stratum represented those empty strata as well.
2. For each post-survey stratum (province, age category, and sex), the number of people that each interview represents was calculated as the population size in a given post-survey stratum divided by the number of interviews realized in that stratum.
3. The post-stratification weight of an interview was then calculated as the number of people that each interview represents divided by the Belgian over 18 population size on 1 January 2017 (as in the original sampling).[71]

Unless specified otherwise, **all further analyses** – descriptive analysis and regression analysis – **made use of these sampling weights.**

Validation of representativeness

The representativeness of the sample was validated with respect to the age and sex composition as well as the health status, quality of life, activity status and educational attainment in the current survey relative to population statistics from administrative sources and representative reference surveys in these domains.

First, for the population makeup, administrative data on the population on 1 January 2017 were obtained from Statistics Belgium.[71]

Second, information on activity status, educational attainment and self-assessed health was obtained from the European Survey on Income and Living Conditions (EU-SILC).[77] The EU-SILC is an annual survey organized by Statistics Belgium with about 12 000 respondents in Belgium that serves as reference source for comparative statistics on socioeconomic status and living conditions. Given that most interviews in our survey were collected in 2019, wave 2019 of the EU-SILC (with interviews from April 2019 to November 2019) was used in the validation exercise.[78] Note that information

in wave 2018 was quite similar to information from wave 2019, and would not materially change the conclusions.

Third, information on self-assessed health and health-related quality of life was obtained from the health interview surveys (HIS). The HIS is a cross-sectional survey organized by Sciensano every 4 to 5 years with a sample size of about 10 000 respondents.[55] (see also Box 4) It is the main source to assess population health. In the validation exercise, we use information from wave 2018 (from January 2018 to December 2018), the most recent wave, as well as wave 2013, which is the only wave with information on the visual analogue scale (VAS).[57, 59]

3.2 Model construction

The EuroQol protocol does not specify a standard regression model to be used to generate utility values for an EQ-5D-5L value set. The regression models selected for the EQ-5D-5L value set vary between countries, reflecting the differences in underlying data characteristics, evaluation process and model selection criteria. Nonetheless, a set of four core regression specifications – one additive model and three multiplicative models (see below) – exists and serves as backbone for the approach followed in the current study.[79-82] First, the modelling strategy for the estimation of cTTO models is discussed. While cTTO was the main technique to determine utility values within the EQ-5D framework, it has been accompanied by a set of DCE questions in more recent valuations. DCE has a strong theoretical foundation in random utility theory and has been increasingly used to quantify stated preferences for health.[44, 83, 84] Therefore, also models estimated using the DCE data are presented as well as the hybrid approach, combining cTTO and DCE data in one model.

3.2.1 Modelling cTTO valuation

The pursued estimation strategy consists in the estimation and evaluation of a large number of potential regression models. Each model can generally be described as follows:

$$DU_{ij} = 1 - U_{ij} = I + \beta_{cTTO} X_j + \varepsilon_{ij} \quad (1)$$

where U_{ij} and DU_{ij} are, respectively, the utility and disutility related to health state j reported by respondent i; I is the intercept in the regression; $\beta_{cTTO} X_j$ represents one of the four core regression specifications that we consider (see below) with coefficients (β_{cTTO}) and variables (X_j) representing the valuation of the different dimensions and levels; ε_{ij} is the error term.

Dependent variable

Disutilities rather than utilities were estimated as dependent variable.[85] Disutility is defined as the deviation in utility from the full health state with utility value 1. The reason for this choice is that utility values can be both positive and negative – they range from -1 to 1 (see section 2.1.1) – which can complicate the estimation of coefficients. Disutility, on the other hand, is always positive and ranges from 0 to 2. Note that we treat the dependent variable as a continuous variable even though this is in reality not the case as respondents can only give 41 distinct values (see section 2.1.1).[81]

Four core regression specifications

Generally, there are four main regression specifications to model EQ-5D-5L value sets.[79-82] All specifications are made up of the same 20 binary variables, but of a different set of estimated coefficients. For each of the five dimensions – i.e. mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), anxiety/depression (AD) –, four binary variables were defined (x_{dl}) indicating whether or not the health state is characterized by a problem on dimension d at level l – i.e. levels 2, 3, 4 or 5, with level 1 being the reference level.

A first, widely used, specification, is the additive 20-coefficients model, referred to as ***ADD20***.[23, 39-44, 62, 73-76, 86-92] Each binary variable has its own coefficient (β_{dl}) that indicates the disutility related to having problems on dimension d at level l. The ADD20 is the most flexible specification with the highest number of estimated coefficients and can be mathematically described as follows[42, 79]:

$$\beta_{cTTO}X_j = \sum_d \sum_l \beta_{dl}x_{dl} \quad (2)$$
$$= \beta_{MO2}x_{MO2} + \beta_{MO3}x_{MO3} + \beta_{MO4}x_{MO4} + \beta_{MO5}x_{MO5}$$
$$+ \beta_{SC2}x_{SC2} + \beta_{SC3}x_{SC3} + \beta_{SC4}x_{SC4} + \beta_{SC5}x_{SC5}$$
$$+ \beta_{UA2}x_{UA2} + \beta_{UA3}x_{UA3} + \beta_{UA4}x_{UA4} + \beta_{UA5}x_{UA5}$$
$$+ \beta_{PD2}x_{PD2} + \beta_{PD3}x_{PD3} + \beta_{PD4}x_{PD4} + \beta_{PD5}x_{PD5}$$
$$+ \beta_{AD2}x_{AD2} + \beta_{AD3}x_{AD3} + \beta_{AD4}x_{AD4} + \beta_{AD5}x_{AD5}$$

The remaining three specifications are multiplicative models, which are more restrictive, i.e. with fewer coefficients. Given their nonlinear nature, they are less frequently used.[79, 80, 93] Nonetheless, the results in valuation studies that consider a multiplicative model suggests that multiplicative models might outperform the ADD20 model in terms of predictive accuracy.[79, 80, 93]

The simplest multiplicative specification, referred to as ***MULT8***, consists of 8 coefficients. In this specification, there is one coefficient for each dimension (β_d), indicating that each dimension has a different disutility weight with β_d representing the disutility of having a problem on dimension d at level 5. It is assumed that the relative distance between the levels is identical for all dimensions. There are coefficients L_l for levels 2, 3 and 4, with levels 1 and 5 standardized at a value of, respectively, 0 and 1. The coefficients L_l have a value between 0 and 1 and indicate the disutility for level l in proportion to level 5. For example, the disutility of having moderate problems (level 3) with mobility is $\beta_{MO} \times L_3$; if L_3 would be 0.45, this implies that the utility decrement of a moderate problem in walking about amounts to 45% of the utility decrement of being unable to walk about. The mathematical formulation of the MULT8 specification is as follows:

$$\beta_{cTTO}X_j = \sum_l \left(\sum_d \beta_d x_{dl}\right) L_l \quad (3)$$
$$= (\beta_{MO}x_{MO2} + \beta_{SC}x_{SC2} + \beta_{UA}x_{UA2} + \beta_{PD}x_{PD2} + \beta_{AD}x_{AD2})L_2$$
$$+ (\beta_{MO}x_{MO3} + \beta_{SC}x_{SC3} + \beta_{UA}x_{UA3} + \beta_{PD}x_{PD3} + \beta_{AD}x_{AD3})L_3$$
$$+ (\beta_{MO}x_{MO4} + \beta_{SC}x_{SC4} + \beta_{UA}x_{UA4} + \beta_{PD}x_{PD4} + \beta_{AD}x_{AD4})L_4$$
$$+ (\beta_{MO}x_{MO5} + \beta_{SC}x_{SC5} + \beta_{UA}x_{UA5} + \beta_{PD}x_{PD5} + \beta_{AD}x_{AD5})$$

The ***MULT9*** specification extends the MULT8 specification with one additional coefficient, L_5. This coefficient allows to make a distinction at level 5 between the dimensions mobility, self-care and usual activities on the one hand and pain/discomfort and anxiety/depression on the other hand. The reason for this distinction is the difference in wording of level 5 in the different dimensions, described as “unable to” in the former three dimensions and “extreme” in the latter two dimensions. Hence, in MULT9, it is assumed that the relative distance between levels 1 to 4 is identical for all dimensions, but might differ between levels 4 and 5 for the first three and the last two dimensions. The mathematical formulation of the MULT9 specification is as follows:

$$\beta_{cTTO}X_j = (\beta_{MO}x_{MO2} + \beta_{SC}x_{SC2} + \beta_{UA}x_{UA2} + \beta_{PD}x_{PD2} + \beta_{AD}x_{AD2})L_2 \quad (4)$$
$$+ (\beta_{MO}x_{MO3} + \beta_{SC}x_{SC3} + \beta_{UA}x_{UA3} + \beta_{PD}x_{PD3} + \beta_{AD}x_{AD3})L_3$$
$$+ (\beta_{MO}x_{MO4} + \beta_{SC}x_{SC4} + \beta_{UA}x_{UA4} + \beta_{PD}x_{PD4} + \beta_{AD}x_{AD4})L_4$$
$$+ (\beta_{MO}x_{MO5} + \beta_{SC}x_{SC5} + \beta_{UA}x_{UA5}) + (\beta_{PD}x_{PD5} + \beta_{AD}x_{AD5})L_5$$

The fourth and last specification is the MULT11, further extending the MULT9 specification. In ***MULT11***, it is assumed that the difference between the first three and the last two dimensions not only affects the relative distance between levels 4 and 5 (as is the case for MULT9), but the relative

distance between all levels. Hence for the two subgroups of dimensions a separate set of level coefficients is defined, i.e. LE_l for the "extreme"-group and LU_l for the "unable to"-group. As for the MULT8 specification, level 1 and level 5 are standardized for each set at 0 and 1, respectively. It can be mathematically specified as follows

$$\begin{aligned}\beta_{cTTO}X_j = {} & (\beta_{MO}x_{MO2} + \beta_{SC}x_{SC2} + \beta_{UA}x_{UA2})LU_2 + (\beta_{PD}x_{PD2} + \beta_{AD}x_{AD2})LE_2 \\ & + (\beta_{MO}x_{MO3} + \beta_{SC}x_{SC3} + \beta_{UA}x_{UA3})LU_3 + (\beta_{PD}x_{PD3} + \beta_{AD}x_{AD3})LE_3 \\ & + (\beta_{MO}x_{MO4} + \beta_{SC}x_{SC4} + \beta_{UA}x_{UA4})LU_4 + \beta_{PD}x_{PD4} + \beta_{AD}x_{AD4})LE_4 \\ & + (\beta_{MO}x_{MO5} + \beta_{SC}x_{SC5} + \beta_{UA}x_{UA5}) + (\beta_{PD}x_{PD5} + \beta_{AD}x_{AD5})\end{aligned} \quad (5)$$

As the mathematical formulations in Equations (2) to (5) make clear, the multiplicative specifications are constrained variants of the more flexible additive model, with fewer degrees of freedom.

A large number of potential alternative options

Within these four core specifications, a range of potential alternatives are tested. We allow for the following alternative options:

1. **With or without intercept**, i.e. I in Equation (1). The intercept can be interpreted as a disutility that applies to all health states that deviate from full health, irrespective of the severity of the health state.
2. **With or without random effects.** Random effects introduce a respondent-specific component in the error term, i.e. ε_{ij} in Equation (1) can be decomposed as $\nu_i + \eta_{ij}$ with ν_i the random component of respondent i. In case no random effects are assumed ν_i is set at zero, otherwise ν_i is characterized by a normal distribution with zero mean and variance σ_r^2. The random effect can be interpreted as individual variation around the intercept.[23, 43, 72, 74, 76, 79, 80, 85, 86, 89, 93]
3. **With or without accounting for heteroscedasticity.** A feature found in several valuation studies is that respondents disagree more strongly on the cTTO valuation of health states that diverge more severely from full health, or put differently, the variability in reported utility values increases with worsening health states.[39-42, 62, 75, 86, 87, 89, 90]This is indicative of heteroscedasticity in the data and can be taken into account by modelling the variance of error term.[81, 94] Following Ramos-Goñi et al. (2016), we estimate the natural logarithm of the standard deviation as a function of the same 20 binary variables that are used to estimate disutility in Equation (1).[94] The same core specification – ADD20, MULT8, MULT9, or MULT11 – is used as for the estimation of disutility.
4. **With or without censoring.** The minimum reported cTTO utility value is capped at -1 by design. It may be possible that some respondents would like to value certain health states even lower.[82] If that is the case, the (dis)utility values are censored, i.e. we observe a capped value instead of the real value. In the estimation, censoring can be taken into account.[39-44, 62, 72, 73, 75, 76, 86, 88-92]

The different alternative options can be combined, leading to a large number of potential regression models to be estimated and evaluated.

3.2.2 *Modelling DCE valuation*

In the DCE task, respondents compared a pair of health states A and B and chose the better one, i.e. the health state with the higher utility value. These choices give information on the relative preference of one health state over another. Contrary to the cTTO valuation, the DCE valuation did not provide direct utility values that are anchored to a scale where a value of 1 represents full health and a value of 0 represents dead, but rather relative values between levels and dimensions.[42] The DCE coefficients can, however, be rescaled to make them more directly comparable to those obtained from the cTTO analysis (see section 3.2.3). The DCE model is not used to generate the final utility values, but rather serves as input in the model selection process (see below) and assess the agreement between the results from cTTO valuation and the DCE valuation.

Let us consider the following utilities for health states A and B in pair p evaluated by respondent i:

$$U_{ipA} <?> U_{ipB} \quad (6)$$

As for cTTO valuation, we model disutilities rather than utilities. The disutilities can be specified in a similar way as in Equation (1) using the same four core specification.

$$DU_{ipA} = 1 - U_{ipA} = I + \beta_{DCE} X_{pA} + \varepsilon_{ipA} <?> DU_{ipB} = 1 - U_{ipB} = I + \beta_{DCE} X_{pB} + \varepsilon_{ipB} \quad (7)$$

The comparison of disutilities can be translated in a binary choice variable (C) indicating whether disutility is highest for health state A ($C = 1$) or for health state B ($C = 0$)

$$\text{If } DU_{ipA} - DU_{ipB} = \beta_{DCE}(X_{pA} - X_{pB}) + (\varepsilon_{ipA} - \varepsilon_{ipB}) > 0 \text{ then } C_{ip} = 1 \quad (8)$$

$$\text{Else if } DU_{ipA} - DU_{ipB} = \beta_{DCE}(X_{pA} - X_{pB}) + (\varepsilon_{ipA} - \varepsilon_{ipB}) \leq 0 \text{ then } C_{ip} = 0$$

The DCE model uses this binary choice variable C_{ip} as dependent variable, while difference between the binary dimension-level variables for health states A and B are used as independent variables. Under the assumption that the errors follow an extreme value distribution, the coefficients can be estimated by a logit model. The interpretation of the coefficients is similar to the cTTO models, except that the scale is different (see above).

Note that by taking the difference between health states A and B, the intercept has disappeared, because we assumed that the intercept is the same for all health states (see Equation (1)). If the intercept is significant, this would signify that there is a systematic disutility difference between health states selected as choice A and health states selected as choice B, i.e. that the intercept in health state A is significantly different from the intercept in health state B.[81]

For the DCE valuation, the four core specifications are estimated without correction for heteroscedasticity and random effects (it can be shown that the random effect cancels out as was the case for the intercept). By design, the DCE data are not censored.

3.2.3 Hybrid model

The aim of both the cTTO and DCE valuation is to provide information on the utility values of health states. One could argue that the same value-function underlies the responses in both exercises and thus expect that the coefficients obtained from the cTTO models are in close agreement with the coefficients from the DCE models.

In order to maximize the available information in the estimation and hence improve accuracy, a hybrid model can be used to estimate simultaneously a single set of coefficients on the cTTO and DCE data. For a detailed description of the hybrid model, we refer the interested reader to Ramos-Goñi et al. (2017).[94] The hybrid model has been frequently selected as final model to generate EQ-5D-5L utility values.[39-42, 86, 95]

The main underlying assumption in the hybrid model is that the coefficients in the DCE model (β_{DCE}) can be rescaled to match the coefficients from the cTTO model (β_{cTTO}). Also in the current study, proportionality between both coefficients is assumed, or put differently, the same scaling parameter applies to all coefficients:

$$\beta_{DCE} = \frac{\beta_{cTTO}}{\theta} \Rightarrow \beta_{HYB} = \beta_{cTTO} = \beta_{DCE}\theta \quad (9)$$

Although the coefficients from the hybrid model (β_{HYB}) are estimated on both cTTO and DCE data, they have exactly the same interpretation as the coefficients from the cTTO model in terms of utility decrements on a scale where value 1 represents full health and value 0 represents dead. Moreover the rescaling parameter θ can also be used to rescale coefficients from the DCE model and make them more comparable to the coefficients from the cTTO model (see section 3.2.2).

The cTTO part of the hybrid model can be specified in various ways – with/without random effect, intercept, correction for heteroscedasticity, censoring etc. – as explained in section 3.2.1, while the DCE part of the hybrid model is specified as in section 3.2.2 without constant term. Of course the same core regression specification – ADD20, MULT8, MULT9, or MULT11 – applied to both the cTTO and DCE part of the hybrid model. One optimization function was used to obtain consistent coefficient estimates (see section 3.2.4)

3.2.4 Fitting of the models

The coefficients for the cTTO, DCE and hybrid models were estimated using maximum likelihood estimation. Certain specifications could also have been estimated using other procedures, e.g. the ADD20 model for cTTO data without random effects can be estimated using a least square procedure. However, a maximum likelihood estimation procedure was preferred since it can be applied to all models discussed in sections 3.2.1 to 3.2.3 and hence all models could be fitted using the same SAS procedure, i.e. Proc NLMIXED in SAS 9.4.[96] The likelihood functions that were maximized are specified Appendix 1.

3.3 Model selection

A large range of regression models were evaluated and compared with the aim to select one final model that is best able to predict utility values for all health states defined by the EQ-5D-5L descriptive system based on the observed cTTO utility values and DCE choices for a small set of these health states. Model selection was based on logical consistency of the coefficient estimates, goodness of fit, predictive accuracy and theoretical considerations.

In a first stage, logical consistency, goodness of fit and predictive accuracy were assessed for the models using cTTO data to get a ranking of best potential specifications. In a second stage, theoretical considerations as well as the results from the DCE regression models were used to consider the use of a censored model and/or a hybrid model.

3.3.1 Logical consistency

The estimated coefficients of the final model must be logically consistent. Coefficients are considered logically consistent if the disutility in a health dimension does not increase with the severity level. This implies the following:

- **ADD20-model**: $I \geq 0$ and for all dimensions (d) $\beta_{d5} \geq \beta_{d4} \geq \beta_{d3} \geq \beta_{d2} \geq 0$
- **MULT8-model**: $I \geq 0$, for all dimensions (d) $\beta_d > 0$, and for the levels $1 \geq L_4 \geq L_3 \geq L_2 \geq 0$
- **MULT9-model**: $I \geq 0$, for all dimensions (d) $\beta_d > 0$, and for the levels $\min(1, L_5) \geq L_4 \geq L_3 \geq L_2 \geq 0$
- **MULT11-model:** $I \geq 0$, for all dimensions (d) $\beta_d > 0$, and for the levels $1 \geq LU_4 \geq LU_3 \geq LU_2 \geq 0$; $1 \geq LE_4 \geq LE_3 \geq LE_2 \geq 0$

3.3.2 Goodness of fit

The second criterion assesses the ability of the estimated models to fit the observed data. A variety of measures exist to evaluate goodness of fit. We opt for the mean absolute error (MAE), a frequently used measure for predictive accuracy, and the Bayesian information criterion (BIC), a measure that combines goodness of fit and model complexity.[43, 97] The goodness of fit of the estimated regression models was evaluated by first ranking the models by their performance on each measure and then summing up both ranks.

The **mean absolute error** is computed as the average of the sum of the absolute values of the difference between the predicted and observed utility value for a health state (see Equation (10)). Note that the MAE was not assessed at the observation level (i.e. a health state evaluated by a respondent), but at the population mean of the health state, as population-level health state utilities are the main outcome of our analysis. Therefore, in the computation of the MAE, the weighted average observed utility of a health state is compared to the predicted value derived from the regression model. A lower MAE is favoured as it indicates a better match between

observed and predicted population values. The MAE is preferred to the root mean squared error (RMSE), since it is less sensitive to an occasional large error.

$$MAE_M = \frac{1}{N}\sum_j \left| \frac{\sum_i w_i U_{ij}}{\sum_i w_i I_{ij}} - \widehat{U}_{Mj} \right| \quad (10)$$

Where MAE_M is the MAE value of regression model M; N equals the number of observed health states that are taken into account in the computation of the MAE, i.e. 86 health states in case of the cTTO data; w_i is the population weight of respondent i; I_{ij} is a dummy indicating that respondent i evaluated health state j; U_{ij} is the utility and disutility related to health state j reported by respondent; and $\widehat{U}_{Mj}$ represent the predicted utility value for health state j based on the estimated coefficients in regression model M.

The ability of the model to have a good fit with the data is important, but a narrow focus on goodness of fit may be misleading as the inclusion of additional coefficients always increases a model's fit. An over-fitted model is, however, not able to distinguish the factors driving the utility values from random variation and does not produce trustworthy predictions. Therefore, also a second measure was used, the **Bayesian information criterion**. It presents a trade-off between the goodness fit of the model (evaluated by the likelihood value) and the complexity or parsimony of the model (evaluated by the number of estimated parameters) and reduces the risk of selecting overfitted models. A model with a lower BIC value is preferred and reflects a better fit and/or fewer estimated coefficients. Only models using the same input data can be compared to each other using the BIC value. Hence, for example cTTO models and hybrid models cannot be compared.[41, 44] The BIC is related to another often used measure, the Akaike's information criterion (AIC), but the BIC penalizes more heavily for additional parameters.[97]

$$BIC_M = -2\ln(L_M) + k_M ln(n) \quad (11)$$

Where BIC_M is the BIC of regression model M; $\ln(.)$ is the natural logarithm function; L_M is the likelihood value of regression model M; k is the number of estimated coefficients of regression model M; and n is the number of observations used in the estimation.

3.3.3 Predictive accuracy

The final regression model predicts utility values for 3 125 health states based on observations for 86 cTTO health states and – in case the hybrid model is selected – 392 health states in the DCE choice pairs. Hence the ability of a model to predict unobserved values is a vital element of the model's performance and predictive accuracy is thus considered a more important selection criterion than goodness of fit.

Cross-validation techniques were used to assess the predictive accuracy of all cTTO models. In this case, the modelling data were split into two parts: a training set and a validation set. First, the regression model was fitted on the training data. Next, out-of-sample utility values were predicted for health states that are withheld from the training data but observed in the validation set. Finally, the MAE was calculated for these health states (see section 3.3.2). In addition, the logical consistency of each of the fitted models in the cross-validation was assessed (see section 3.3.1).

Based on Rand-Hendriksen et al. (2017), three different cross-validations were used, using a different split between training and validation data.[79]

1. **Leave-one-state-out:** one health state was left out of the modelling dataset to create a training dataset with 85 health states and to predict the utility of the left-out state. Hence, the majority of the data was available to fit the model. This subdivision was replicated 86 times, sequentially removing each of the observed health states once. The final MAE value for the leave-one-state-out cross-validation is the average of the MAE resulting from the 86 replications.
2. **Leave-one-block-out:** in the cTTO valuation, the health states to be evaluated were grouped ten blocks (see section 2.1.1). A block was randomly assigned to a respondent. Hence health states in a particular block were all valued by the same set of individuals and are not fully

independent. In case only one state is left out, information on how these respondents valued other health states was still observed in the training data potentially increasing the accuracy of out-of-sample predictions. Therefore a second cross-validation technique consisted in sequentially leaving out each of the ten blocks of health states to create the training data, effectively reducing the data available for model fitting by about 10%. For each of the 10 replications, this procedure removed all information of a subset of respondents. The MAE was not calculated for all health states in the left out block, as the worst state (55555) was included in all blocks and the mildest health state (with severity 6, or only one decrement from full health) was included in at least one other block (see section 2.1.1). Hence, the predictions for these states were not assessed. The final MAE value for the leave-one-block-out cross-validation is the average of the MAE resulting from the 10 replications.

3. **Leave-random-block-out:** the downside of the leave-one-block-out cross-validation is that it can only be replicated 10 times. Therefore, a third cross-validation was carried out, in which new blocks were randomly generated. Using random Latin squares, the original 10 blocks were mixed to create 10 new random pseudo-blocks so that each new pseudo-block consisted of one health state from each of the original blocks, including once the worst state and one of the 5 mildest states. This process of creating new blocks was replicated 10 times, leading to 100 pseudo-blocks that can be sequentially left out. As in the leave-one-state-out cross-validation, this technique does not lead to the exclusion of full interviews. The final MAE value for the leave-random-block-out cross-validation is the average of the MAE resulting from the 100 replications.

The predictive accuracy of the estimated regression models was evaluated by summing up the MAE resulting from each of the three different cross-validation techniques and next ranking the models from low to high on the combined MAE.

3.3.4 Theoretical considerations

Finally, theoretical considerations may steer model selection. Some model features may be desirable from a theoretical point of view, but may reduce the model's performance in terms of predictive accuracy or goodness of fit as measured above. We attempt to substantiate the choices based on theoretical considerations with suggestive evidence. These choices relate to the need to correct for heteroscedasticity, the presence of censored data and the use of the hybrid model.

Heteroscedasticity

As discussed in section 3.2.1, a feature found in several valuation studies is that the variability in reported utility values increases with worsening health states. This is indicative of heteroscedasticity in the data and can be taken into account by modelling the variance of error term. We examined the presence of heteroscedasticity in the collected data by looking at the variance of the error terms in the cTTO models as function of the predicted disutility values. If heteroscedasticity is found, this would be an argument in favour of selecting a model that corrects for this, even if models without such correction perform better in terms of goodness of fit or predictive accuracy.

Censored data

The minimum reported cTTO utility value is capped at -1 by design. It is possible that some respondents would like to value certain health states even lower. If that is the case, the (dis)utility values are censored, i.e. we observe a capped value instead of the real value. This can in particular be the case when respondents have given the lowest valuation to multiple health states.

While there are arguments in favor of censoring, and a number of value studies have accounted for it, it is impossible to discern from the data whether or not respondents want to value health states lower than they currently do when given the opportunity. Alternative explanations exist. One possibility is that respondents are (nearly) indifferent between health states once a critical level of ill health is reached and hence might value multiple health states at the lowest possible utility value of -1.

The possible censored nature of the data can taken into account in the estimation by adjusting the maximum likelihood function (see section 3.2.4 and Appendix 1). This would deliberately lead to a lower valuation of the more severe health states than observed and reduce the performance of the model in terms of goodness of fit or predictive accuracy. Hence, the choice to correct for censoring is typically a theoretical consideration. Nonetheless, a comparison of DCE and cTTO results can give some further indication. The DCE data are uncensored by design. If one argues that the same value-function underlies the responses in both cTTO and DCE, the agreement of the cTTO results with the DCE results can provide suggestive evidence on the need for treating the cTTO data as censored. Versteegh et al. (2016) propose to explore the "DCE fit", i.e. the mean absolute difference between the DCE utility values for all 3 125 health states and the cTTO value set generated once with and once without censoring.[43] To generate the DCE value set, the rescaling factor θ from the hybrid model with the same core specification was used to make both value set more comparable. The model (censored or uncensored) with the lowest DCE fit is preferred.

The hybrid model

The cTTO as well as DCE valuation technique have strengths and weaknesses and the hybrid model (see section 3.2.3) has been proposed as a pragmatic compromise to combine stated preferences from both methods.[39, 44, 86, 89] The hybrid model has been selected in multiple recent valuation studies as preferred model.[23, 39-42, 75, 76, 86, 88, 90-93, 95]

While the desirability of the hybrid model is difficult to assess, we do argue that it can be of added value when the same value-function underlies the responses in both cTTO and DCE. This implies that there needs to be a sufficient level of agreement between the coefficients from the DCE and cTTO model when estimated separately. Given that both data types are used simultaneously, it is expected that coefficients in the hybrid are estimated with greater precision as the available information is maximized, and, secondly, that the biases due to shortcomings of either data type are potentially reduced.[44]

In case strong agreement between the DCE results and cTTO results was found, this would justify the use of the hybrid model.

3.4 Preferred model and value set

Once the final model was selected, it was fitted on the modelling dataset. Confidence intervals of the coefficients were derived using the Rao-Wu bootstrap with 1 000 bootstrap samples.[69] To generate the bootstrap samples, the original strata were merged to increase the minimum number of interviews per strata to approximately 15. The bootstrap strata were defined at the level of a province (combining the neighbouring provinces Namur and Luxembourg), sex and age group ([18-40), [40-60), 60+).

Sensitivity analysis

The robustness of the coefficients of the final regression model were tested by re-estimating the model coefficients using 3 different data samples as input: (1) the full sample (see section 3.1.1); (2) the modelling data restricted to the interviews collected prior to the COVID-19 pandemic, i.e. without the data from 48 respondents obtained in phase 4 (see section 2.4.3); (3) the modelling data excluding observations flagged in the feedback module (see sections 3.1.1 and 2.1.3). These robustness checks allow to assess the impact of implausible data on the preference values and to explore whether the COVID-19 pandemic has had an impact on the valuation of certain health dimensions, e.g. anxiety and depression.

Value set

The coefficients of the preferred model were used to generate the EQ-5D-5L value set for Belgium with utility values for all 3 125 health state. In addition, the utility value for the unconscious health state was not derived from the regression model, but equals the weighted average of the stated preferences. The new Belgian value set was compared to the previous value set based on the EQ-5D-3L questionnaire and the EQ-5D-5L crosswalk values derived from this set using the mapping function proposed by van Hout et al. (2012).[47, 48]

4 RESULTS: THE BELGIAN EQ-5D-5L VALUE SET

4.1 Sample characteristics

The modelling dataset consists of 892 respondents. The representativeness of the sample for the Belgian population was assessed in terms of age, sex, region, health status, health-related quality of life, education and employment status using population statistics from administrative sources and reference surveys in these domains. Using the post-stratification weights, we conclude from Figure 11 that the sample is representative for the Belgian population.

In panel A of Figure 11, the **age and sex composition** of the survey sample (red and blue bars) was compared to the makeup of the Belgian population. Age refers to the age at time of sampling and not at the time of interview. As the sampling procedure was designed on the Belgian population reported by Statistics Belgium on 1 January 2017 (see section 2.2.2), this cross-section of the population served as reference (black line).[71] In general, there is a very good visual match between survey and population distribution. In addition, a chi-squared test confirmed that the survey and population distribution are not statistically different. This was expected as the post-stratification weights were computed to fit the age and sex composition within the strata (see section 3.1). There is only a minor overrepresentation of women aged between 60-69 (7.7% in the survey sample; 7.3% in the Belgian population) and a deficit in the number of respondents aged between 90 and 99 for both sexes (Females: 0.4% in the survey sample; 0.8% in the Belgian population / Males: 0.1% in the survey sample; 0.3% in the Belgian population).

The **regional composition** of the survey sample and the population is also in close agreement (see panel B of Figure 11). There is a slight overrepresentation of respondents of the German community, given that two municipalities from this community were randomly selected in the sampling process (see section 2.2). An age and sex composition by means of a population pyramid at the regional level is presented in 0 and further substantiates the representativeness of the survey sample.

Self-assessed health as reported by the individuals in the survey sample corresponds very closely to information obtained from the EU-SILC (wave 2019) and the HIS (wave 2018), as shown in panel C of Figure 11. For example, about 75% of the individuals in our survey reported to be in good or very good health (row 3), a fraction that is situated in between those observed in the EU-SILC and the HIS.

Comparative statistics on **quality of life** are found in panel D of Figure 11. Row 1 compares the fraction of individuals who reported not to have any problem on any of the 5 dimensions, i.e. whose health state can be defined as 11111 in the EQ-5D-5L framework. Individuals indicating not to experience any problem represent a lower share of our sample relative to the HIS (34.6% in the survey sample; 40.7% in the HIS). The other side of the picture is that a higher fraction indicated to experience health issues, i.e. any problem from level 2 to 5, and this for all dimensions except self-care. It is reassuring, though, that the relative differences between the dimensions in the share of individuals reporting any problem shows a similar pattern in our data and the HIS. In particular, the fraction of individuals reporting any pain or discomfort is high, almost 60%, whereas about 20% to 30% of the individuals indicated issues in the dimensions of mobility, usual activities or anxiety/depression. About 6% of the individuals reported to have problems with self-care.

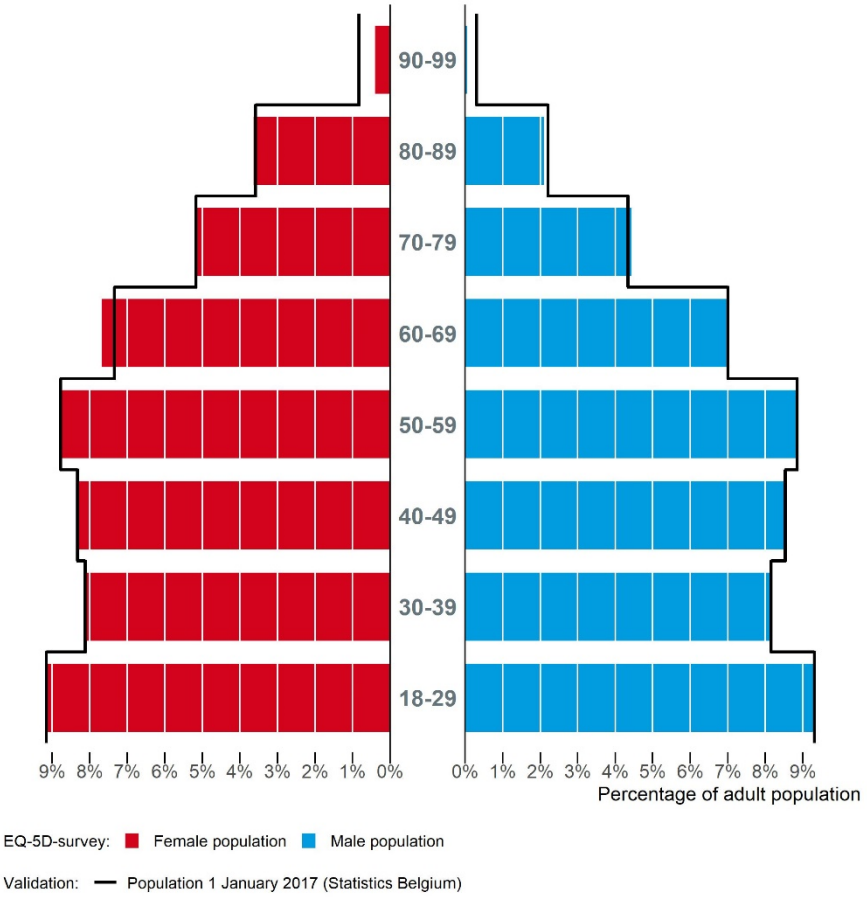
It is documented that the COVID-19 pandemic has contributed to a worsening of mental health in the population[98], which might affect the assessment of the health status of individuals who participated in the survey before or after the first COVID-19 lockdown in Belgium between March 2020 and June 2020. As discussed in section 2.4.3, interviews in phase 1 to 3 were collected prior to the first COVID-19 lockdown in Belgium (844 respondents in the modelling dataset), while phase 4 took place between the first and second lockdown (48 respondents in the modelling dataset). When excluding respondents from phase 4, the statistics on quality of life were not significantly nor meaningfully different (results not shown in Figure 11).

The mean **visual analogue score (VAS)** for the modelling sample was 77.6 (standard deviation of 16.0), with a decreasing score by age. Panel E of Figure 11 presents the comparison of the mean scores by age group in our survey and in wave 2013 of the HIS. The important time lag between the outcomes of both surveys makes a comparison less straightforward, but it is encouraging that the average scores are similar in magnitude. The VAS is only slightly higher in the lower age groups in our data relative to the HIS.

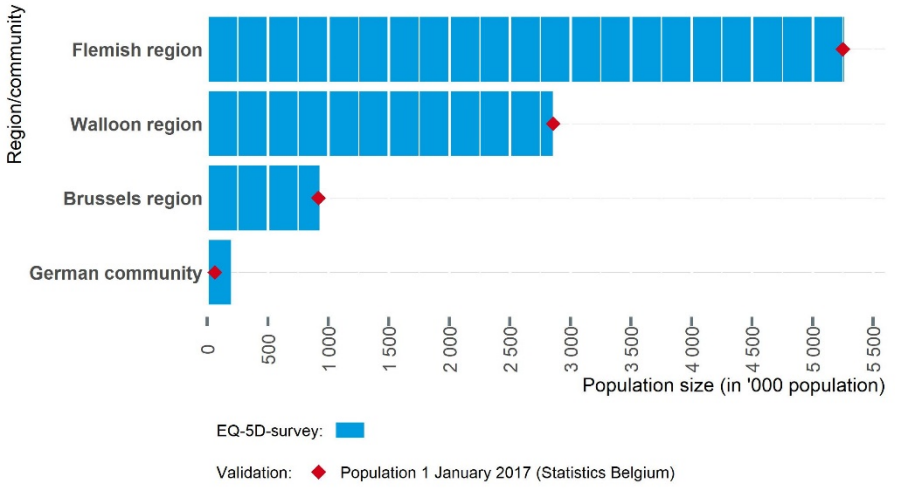
Socioeconomic information on **educational attainment** (subdivided in 3 groups) and **activity status** (subdivided in 5 groups) are summarized in Panel F and G of Figure 11. The socioeconomic profile of respondents in our survey is in line with those in wave 2019 of the EU-SILC. Among the survey participants, there is a slight overrepresentation of individuals with tertiary education (42.0% in the survey sample; 37.7% in the EU-SILC) and employees (47.5% in the survey sample; 45.2% in the EU-SILC) and a lower share of inactive individuals (15.5% in the survey sample; 19.0% in the EU-SILC). These differences are, however, small and likely related with each other, i.e. individuals with tertiary education have lower inactivity and unemployment rates.[99] Moreover, these patterns can (partly) be explained by the nature of the tasks expected from the respondent, which may dissuade the participation of individuals with lower educational attainments or lead to a higher rate of excluded interviews in the quality control process (see section 2.1.7).

Figure 11 – Distribution of age, sex, region, health, health-related quality of life, education, and employment status in the EQ-5D-5L survey sample relative to the population

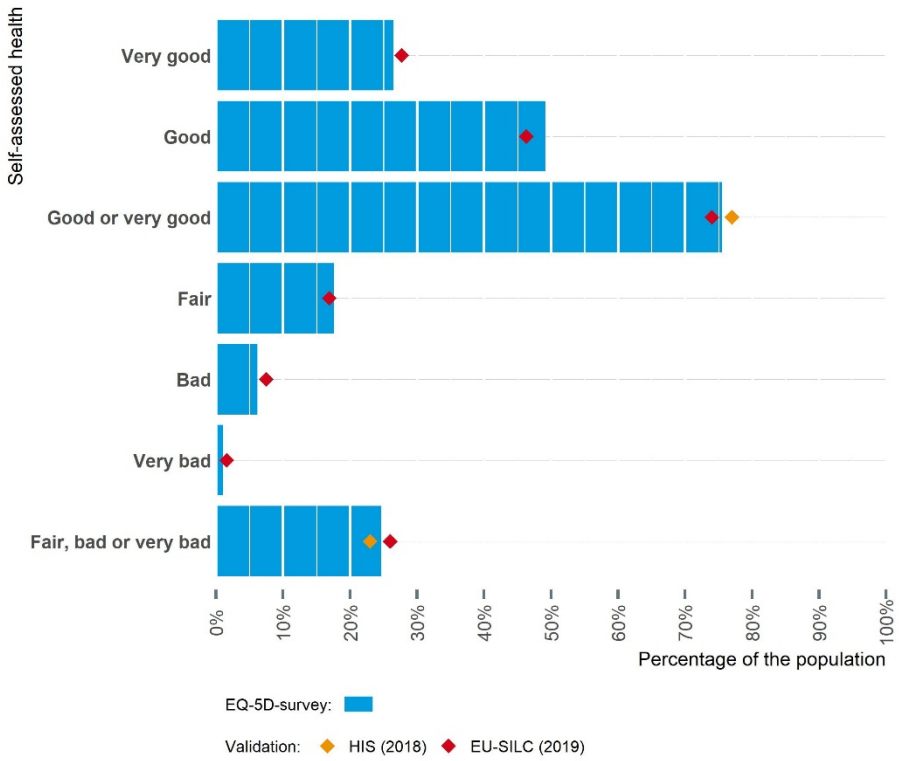
A. Population pyramid with age-sex distribution



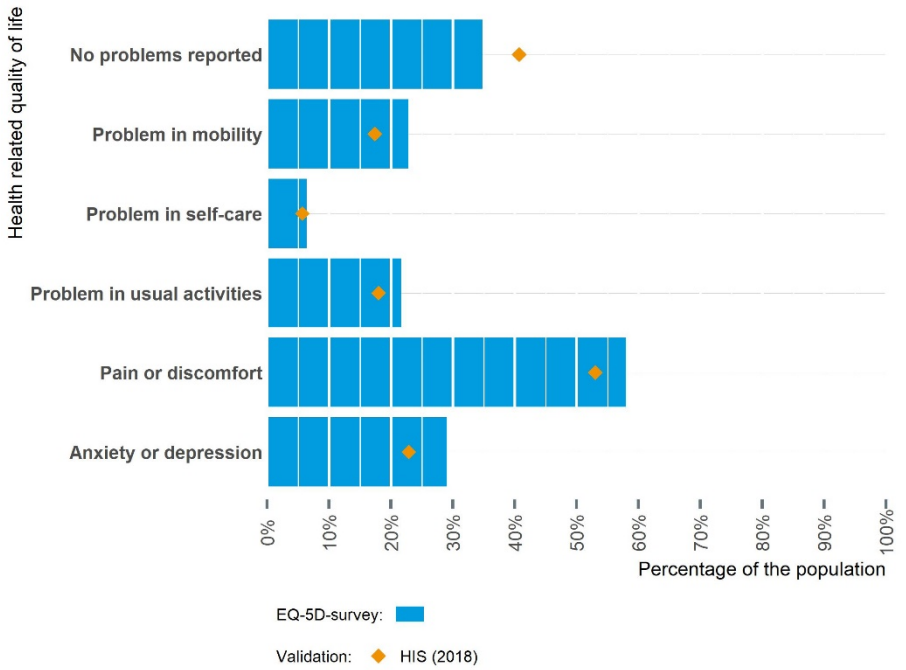
B. Regional distribution



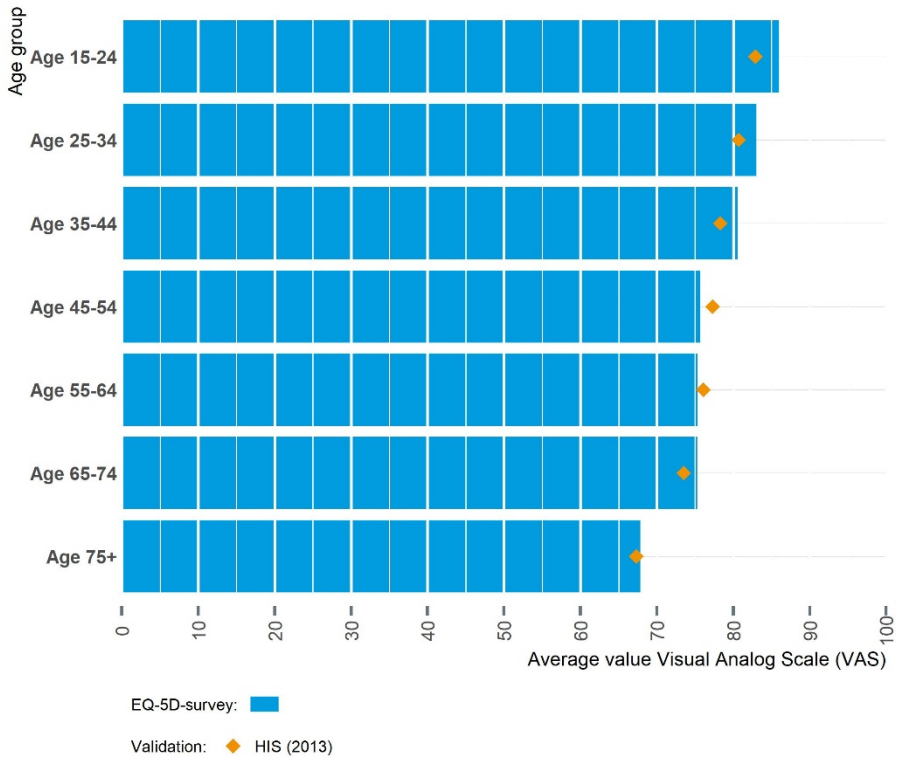
C. Distribution self-assessed health



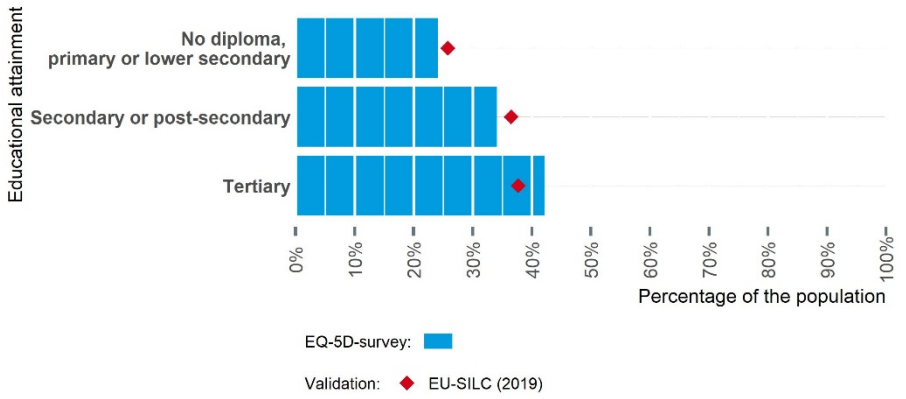
D. Health-related quality of life



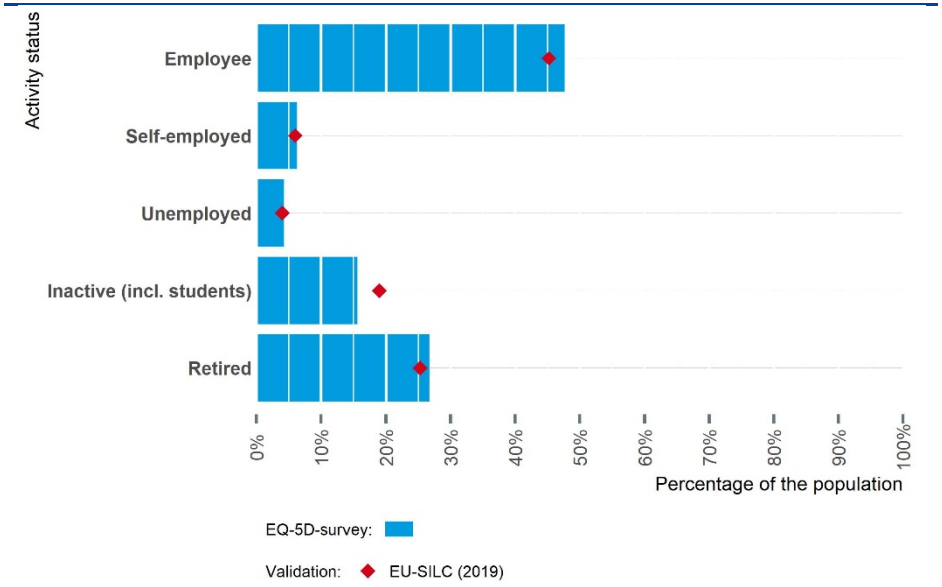
E. Visual Analog Scale (VAS)



F. Education



G. Activity status



4.2 Data characteristics

A number of descriptive analyses were conducted to assess features of both cTTO and DCE data.

4.2.1 cTTO data

Clustering at the lowest and highest values in the distribution of cTTO values

The distribution of observed cTTO values is presented in Figure 12. The full range of potential cTTO values was found in the data. Overall 25.2% of cTTO values are negative and were thus valued worse than dead. About 9.1% of the cTTO tasks resulted in the lowest possible value of -1, indicating that respondents exhausted all available lead time. 161 respondents (18%) valued multiple heath states at -1.

Clustering was also observed for other key values, such as 0.5, 0.95 and 1. About 9.5% and 9.9% of the responses have a cTTO value of 1 and 0.95, respectively, implying no or a very limited willingness to trade-off life-years to avoid mild health problems. It is also clear from Figure 12 that respondents show digit preference with higher percentages of responses for round numbers e.g. the percentage of cTTO values at 0.60 clearly exceeds the percentages at 0.55 or 0.65. Responses at -0.5, 0 and 0.5 represent, respectively, 2.8%, 3.6% and 7.4% of all observations.

Variation in the valuation of health states is important and increasing with severity

The results in Figure 12 mask the important variation observed in the valuation of health states. It is not only the worst state that was valued at -1 nor did only mild health states receive a cTTO value of 1. The individual variation in cTTO values by health states is presented in Figure 13. A higher percentage of responses for a specific combination is visualized by the size and transparency of the bubble. Positive values are presented in green, negative values in red and a value of 0 in grey. For the 5 mildest health states, the cTTO values are heavily concentrated at 1 or 0.95, with a high level of agreement between the respondents. The divergence in preferences gradually increases and from severity 11 onwards, health states systematically receive negative valuations, while below this threshold, negative values occur sporadically. At severity level 17 or higher, a lower share of high cTTO values is observed (between 0.5 and 1).

Figure 12 – Distribution of observed cTTO values (unconscious state not included)

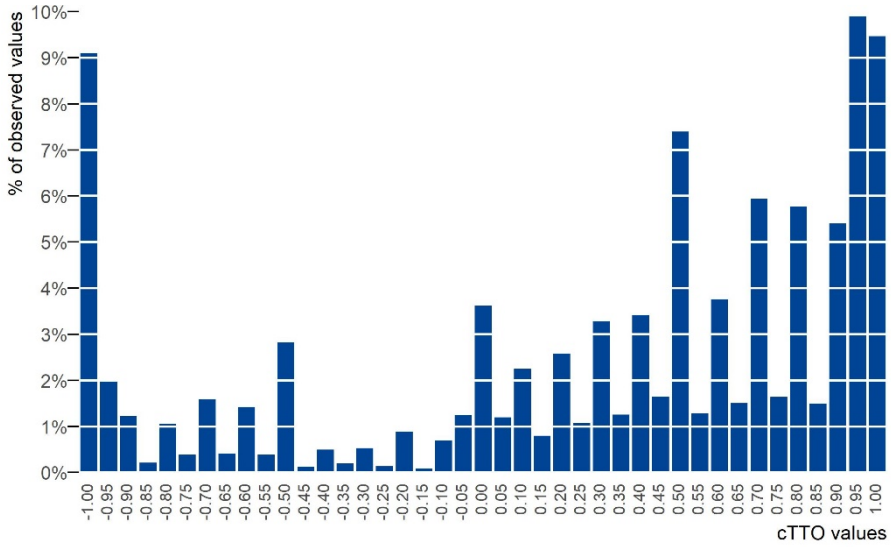
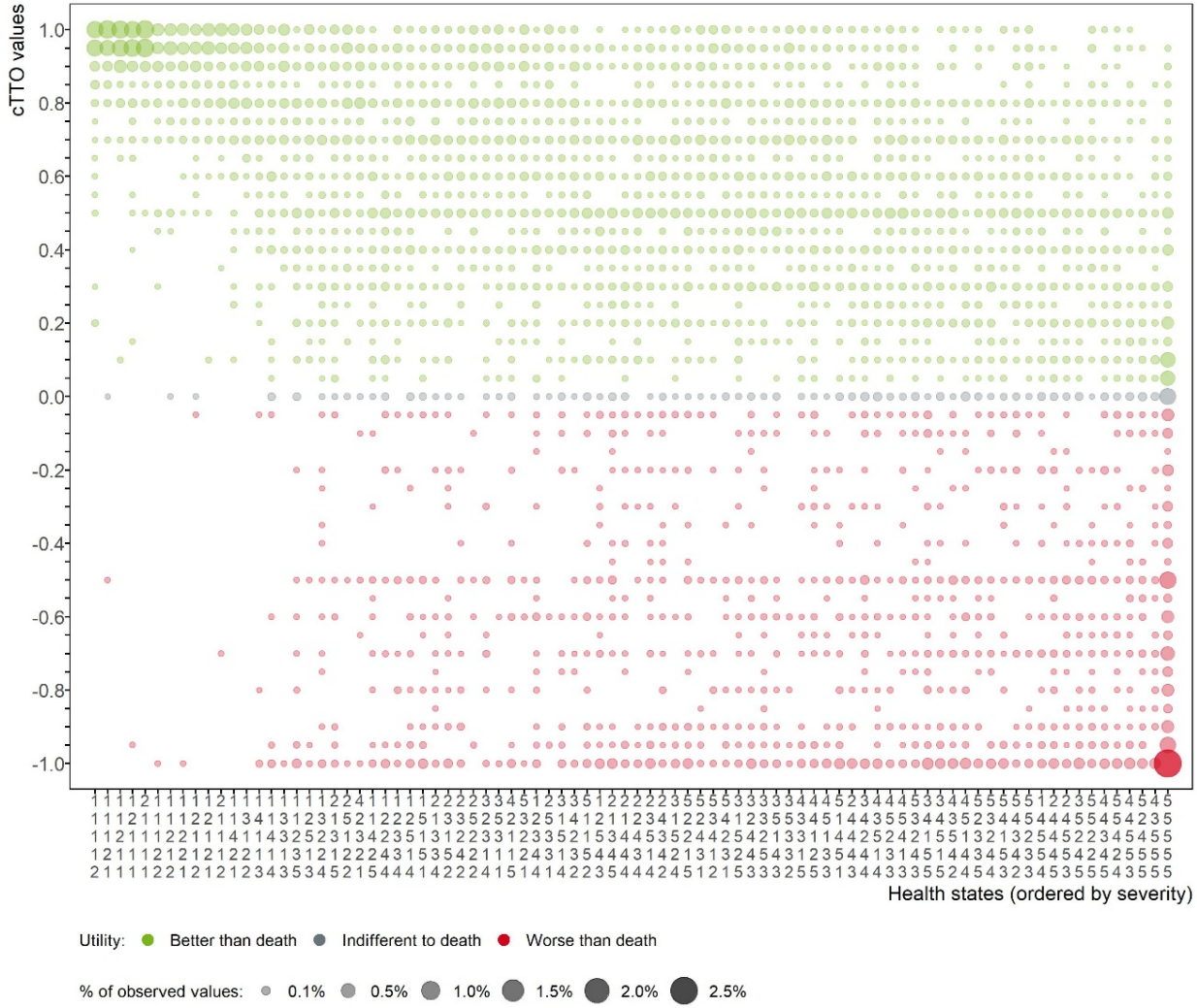


Figure 13 – Distribution of observed cTTO values for the 86 health states (unconscious state not included)



The relation between severity, variability, mean and median cTTO values is depicted more clearly in Figure 14. This can be seen as a simple test of face validity of our data, where the expectation is that cTTO values decrease as health states worsen. The results in Figure 14 confirm that both mean and median cTTO values are lower as severity levels increase. The variability as measured by the standard deviation increases strongly up to level 13 and remains more stable thereafter. This is indicative of heteroscedasticity in the data. Except for the worst health state, the median value is above the mean, indicating that distribution of values is skewed to the left. A number of (very) low cTTO scores pulls down the mean value below the median. This is exactly what is shown in Figure 12 and Figure 13. When respondents prefer to value a health state worse than dead, they generally chose a strong negative value, i.e. lower or equal to -0.5, rather than between 0 and -0.5.

Variation in negative cTTO values partly influenced by interviewer

While individual preferences impact upon the cTTO valuation, the observed variability in values and the importance attached to each dimension (see below), there also appears to be an influence of the interviewer on the valuation. In spite of a strict interview protocol, quality control process and training of the interviewers, differences in the cTTO valuation between interviewers were observed.

Figure 15 presents the difference in cTTO valuation at the interviewer level. Each interviewer is represented by a unique symbol while the color of the symbol is related to the number of conducted interviews. The results indicate that there is particularly important variation in the fraction of responses clustered at 1 and at -1 and in the fraction of responses valued worse than dead (cTTO value < 0). Figure 12 showed that at the upper end of the distribution, there was a clustering of values at both 0.95 and 1. When taking both values together, the results in Figure 15 indicate much less variability between interviewers in the fraction of responses having one of both values. The fraction of responses valued at -1 ranges from 0% to 32% across interviewers, while the fraction of negatively valued health states ranges from 4% to 50%. The upper and lower ends of the ranges were not determined by interviewers with lower numbers of conducted interviews. On the other hand, the fraction of excluded interviews (see section 3.1.1 for the exclusion criteria), the fraction of interviews flagged in the feedback modules (see section 2.1.3 for information on the feedback module) and the percentage of responses clustered at 0.5, 0 or -0.5 appears to be quite consistent across interviewers.

Figure 14 – Mean and median cTTO value by severity level (unconscious state not included)

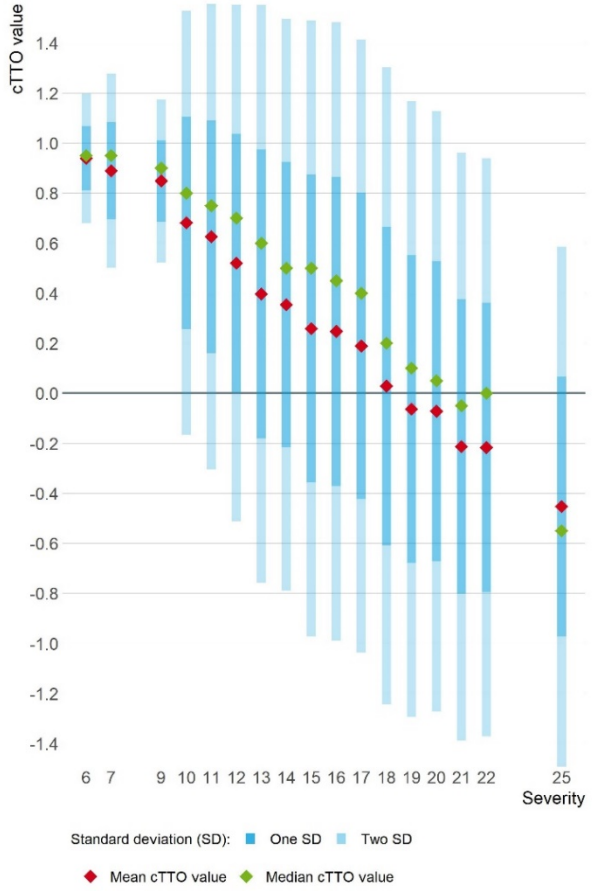
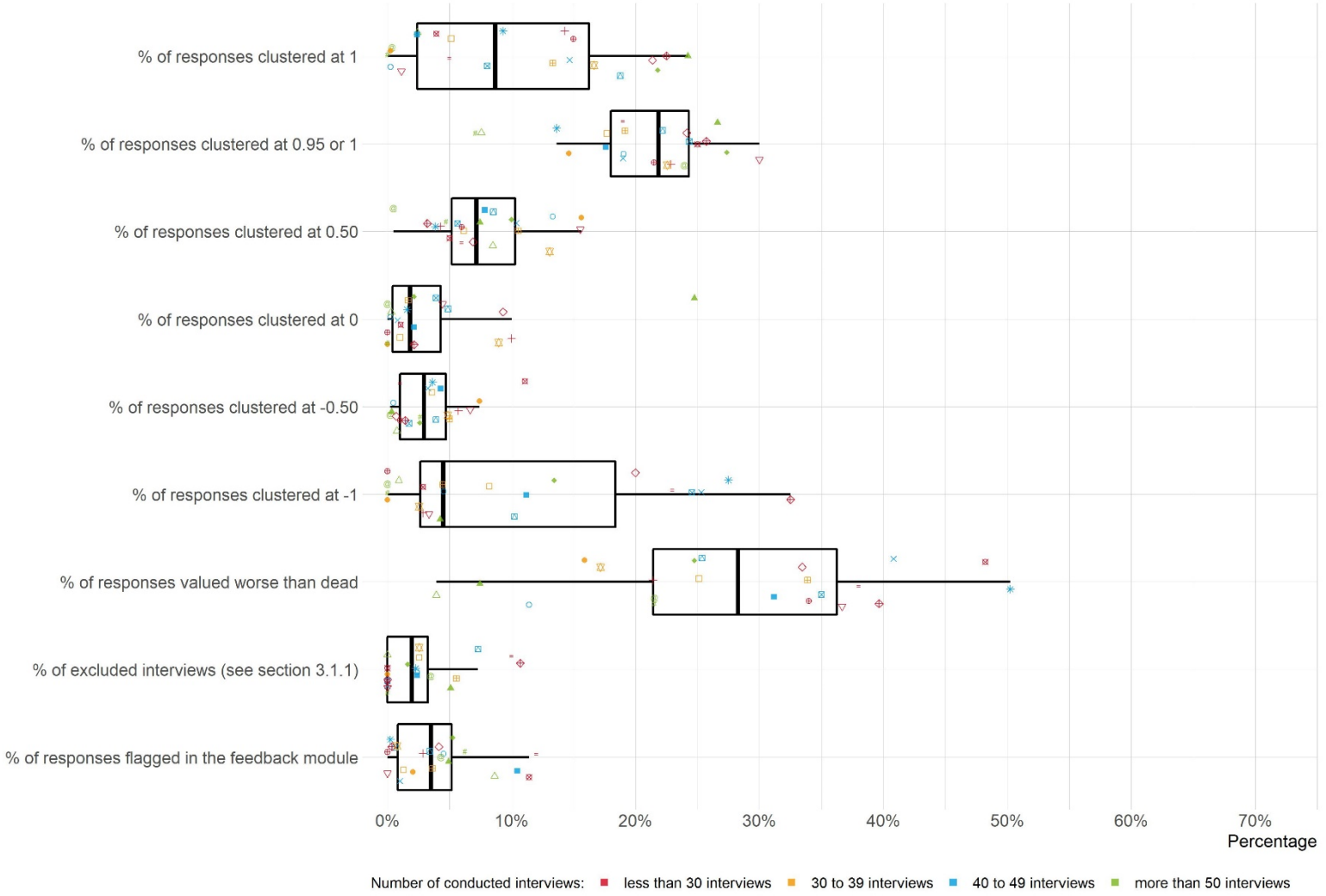


Figure 15 – Differences in cTTO valuation by interviewer



The unconscious state and the worst state are least preferred

In Table 3, summary statistics (mean, median, variance, p10, p25, p75, p90) are provided for all 86 health states of the cTTO valuation as well as the unconscious state. The unconscious state was valued lowest (mean at -0.454), while the health state with slight problems in walking about and no problems in the other dimensions – state 21111 – was valued most closely to full health (mean at 0.958). The 5 mildest states have the highest cTTO values, ranging from 0.923 to 0.958. The mean cTTO value for the worst state (55555) and the unconscious state are nearly identical, but the variability in valuation is lower for the unconscious state. In addition to these two states, 14 other states were negatively valued on average.

Table 3 – Summary statistics for the 86 health states of the cTTO valuation and the unconscious state

Profile	N	Mean (standard error)		P10	P25	Median	P75	P90	Variance	Profile	N	Mean (standard error)		P10	P25	Median	P75	P90	Variance
11112	171	0.926	(0.011)	0.85	0.95	0.95	1.00	1.00	0.019	**31525**	92	0.269	(0.057)	-0.70	0.10	0.40	0.60	0.90	0.306
11121	169	0.945	(0.011)	0.90	0.95	0.95	1.00	1.00	0.020	**32314**	82	0.444	(0.063)	-0.65	0.20	0.65	0.90	0.95	0.328
11122	89	0.874	(0.026)	0.70	0.90	0.95	1.00	1.00	0.061	**32443**	96	0.183	(0.066)	-1.00	-0.15	0.45	0.65	0.85	0.425
11211	190	0.941	(0.008)	0.90	0.95	0.95	1.00	1.00	0.011	**33253**	82	0.177	(0.072)	-0.90	-0.50	0.45	0.70	0.90	0.424
11212	88	0.896	(0.017)	0.75	0.90	0.95	1.00	1.00	0.027	**34155**	96	-0.247	(0.061)	-1.00	-0.90	-0.10	0.30	0.50	0.358
11221	93	0.892	(0.023)	0.75	0.85	0.95	1.00	1.00	0.046	**34232**	89	0.560	(0.050)	0.00	0.50	0.70	0.85	0.95	0.221
11235	93	0.265	(0.065)	-1.00	0.00	0.50	0.75	0.85	0.385	**34244**	88	-0.061	(0.063)	-1.00	-0.55	0.00	0.40	0.80	0.371
11414	79	0.428	(0.063)	-0.60	0.30	0.60	0.80	0.95	0.317	**34515**	93	0.005	(0.069)	-1.00	-0.70	0.20	0.55	0.70	0.431
11421	92	0.828	(0.020)	0.60	0.75	0.85	0.95	1.00	0.037	**35143**	79	0.236	(0.066)	-0.90	0.00	0.45	0.65	0.80	0.347
11425	94	0.376	(0.055)	-0.55	0.20	0.50	0.80	0.95	0.285	**35245**	93	-0.151	(0.065)	-1.00	-0.90	0.00	0.40	0.65	0.391
12111	182	0.923	(0.013)	0.80	0.95	0.95	1.00	1.00	0.030	**35311**	89	0.670	(0.041)	0.25	0.55	0.80	0.90	1.00	0.146
12112	88	0.881	(0.019)	0.75	0.85	0.95	0.95	1.00	0.032	**35332**	94	0.605	(0.043)	0.30	0.50	0.70	0.90	0.95	0.175
12121	96	0.903	(0.015)	0.75	0.90	0.95	1.00	1.00	0.022	**42115**	94	0.281	(0.059)	-0.60	0.00	0.45	0.70	0.80	0.322
12244	92	0.219	(0.056)	-0.95	0.00	0.35	0.50	0.80	0.291	**42321**	89	0.670	(0.039)	0.30	0.60	0.80	0.90	0.95	0.134
12334	82	0.479	(0.063)	-0.50	0.40	0.70	0.90	0.95	0.324	**43315**	89	0.257	(0.065)	-0.95	0.05	0.45	0.70	0.90	0.369
12344	90	0.178	(0.062)	-0.80	-0.15	0.35	0.60	0.85	0.336	**43514**	88	0.150	(0.067)	-1.00	-0.10	0.30	0.60	0.80	0.411
12513	90	0.651	(0.043)	0.20	0.50	0.70	0.95	1.00	0.159	**43542**	96	0.056	(0.067)	-1.00	-0.50	0.35	0.55	0.80	0.434
12514	93	0.278	(0.065)	-0.95	0.00	0.50	0.75	0.90	0.389	**43555**	89	-0.217	(0.062)	-1.00	-0.80	0.00	0.30	0.55	0.334
12543	96	0.184	(0.064)	-1.00	-0.25	0.40	0.60	0.90	0.395	**44125**	90	0.165	(0.061)	-0.95	-0.10	0.30	0.55	0.80	0.319
13122	94	0.869	(0.013)	0.70	0.80	0.90	0.95	1.00	0.016	**44345**	90	-0.086	(0.063)	-1.00	-0.60	0.00	0.40	0.55	0.341
13224	89	0.376	(0.061)	-0.70	0.20	0.55	0.80	0.90	0.324	**44553**	88	-0.231	(0.064)	-1.00	-0.90	-0.10	0.20	0.50	0.373
13313	92	0.793	(0.027)	0.50	0.70	0.90	0.95	1.00	0.069	**45133**	96	0.339	(0.064)	-1.00	0.30	0.60	0.75	0.90	0.405
14113	89	0.681	(0.045)	0.30	0.55	0.80	0.95	1.00	0.180	**45144**	93	-0.183	(0.066)	-1.00	-0.80	0.00	0.30	0.70	0.394
14554	90	-0.034	(0.061)	-1.00	-0.60	0.10	0.40	0.65	0.323	**45233**	92	0.383	(0.058)	-0.80	0.25	0.50	0.75	0.90	0.316
15151	89	0.131	(0.073)	-1.00	-0.50	0.40	0.70	0.85	0.464	**45413**	94	0.317	(0.057)	-0.70	0.15	0.50	0.70	0.90	0.306

Profile	N	Mean (standard error)		P10	P25	Median	P75	P90	Variance
21111	180	0.958	(0.004)	0.90	0.95	0.95	1.00	1.00	0.004
21112	90	0.888	(0.022)	0.70	0.80	0.95	1.00	1.00	0.040
21315	89	0.446	(0.062)	-0.60	0.25	0.65	0.90	1.00	0.335
21334	82	0.441	(0.064)	-0.70	0.35	0.70	0.80	0.95	0.338
21345	88	-0.002	(0.063)	-1.00	-0.50	0.15	0.50	0.70	0.364
21444	79	0.093	(0.063)	-0.95	-0.10	0.25	0.50	0.70	0.317
22434	94	0.405	(0.052)	-0.30	0.30	0.50	0.70	0.90	0.253
23152	88	0.215	(0.068)	-1.00	-0.20	0.50	0.70	0.90	0.427
23242	82	0.345	(0.067)	-0.90	0.25	0.55	0.75	0.95	0.369
23514	96	0.163	(0.071)	-1.00	-0.50	0.50	0.70	0.90	0.489
24342	82	0.221	(0.066)	-0.80	-0.10	0.45	0.65	0.80	0.356
24443	89	0.164	(0.067)	-1.00	-0.05	0.35	0.60	0.75	0.394
24445	89	-0.215	(0.062)	-1.00	-0.80	0.00	0.25	0.50	0.334
24553	94	0.032	(0.061)	-0.95	-0.50	0.20	0.50	0.70	0.341
25122	92	0.710	(0.034)	0.35	0.60	0.80	0.95	1.00	0.110
25222	79	0.668	(0.044)	0.35	0.55	0.80	0.90	1.00	0.152
25331	79	0.689	(0.041)	0.40	0.50	0.80	0.95	1.00	0.136
31514	79	0.264	(0.065)	-0.90	0.00	0.40	0.65	0.80	0.342
31524	89	0.335	(0.065)	-0.95	0.20	0.50	0.70	0.95	0.367
51152	94	0.113	(0.061)	-0.90	-0.40	0.30	0.50	0.70	0.347
51451	93	-0.030	(0.072)	-1.00	-0.80	0.15	0.50	0.80	0.467
52215	96	0.055	(0.066)	-1.00	-0.50	0.30	0.50	0.75	0.424
52335	89	0.121	(0.064)	-1.00	-0.50	0.25	0.60	0.80	0.358
52431	89	0.481	(0.059)	-0.70	0.40	0.70	0.90	1.00	0.309
52455	92	-0.196	(0.058)	-1.00	-0.70	0.00	0.20	0.50	0.321
53221	90	0.690	(0.039)	0.30	0.60	0.80	0.95	1.00	0.128
53243	79	0.170	(0.066)	-0.95	-0.05	0.35	0.60	0.80	0.352
53244	79	-0.042	(0.066)	-1.00	-0.60	0.15	0.35	0.60	0.352
53412	82	0.498	(0.060)	-0.50	0.40	0.65	0.90	0.95	0.295
54153	89	0.031	(0.071)	-1.00	-0.65	0.15	0.60	0.70	0.446
54231	93	0.441	(0.057)	-0.55	0.30	0.60	0.80	0.95	0.298
54342	90	0.183	(0.056)	-0.70	-0.10	0.30	0.50	0.75	0.270
55225	82	0.056	(0.076)	-0.95	-0.65	0.20	0.70	0.95	0.478
55233	92	0.347	(0.062)	-0.70	0.20	0.50	0.80	0.95	0.362
55424	88	-0.060	(0.064)	-1.00	-0.65	0.10	0.40	0.60	0.383
55555	892	-0.453	(0.017)	-1.00	-1.00	-0.55	0.00	0.20	0.270
Unconscious	189	-0.454	(0.035)	-1.00	-1.00	-0.50	0.00	0.05	0.233

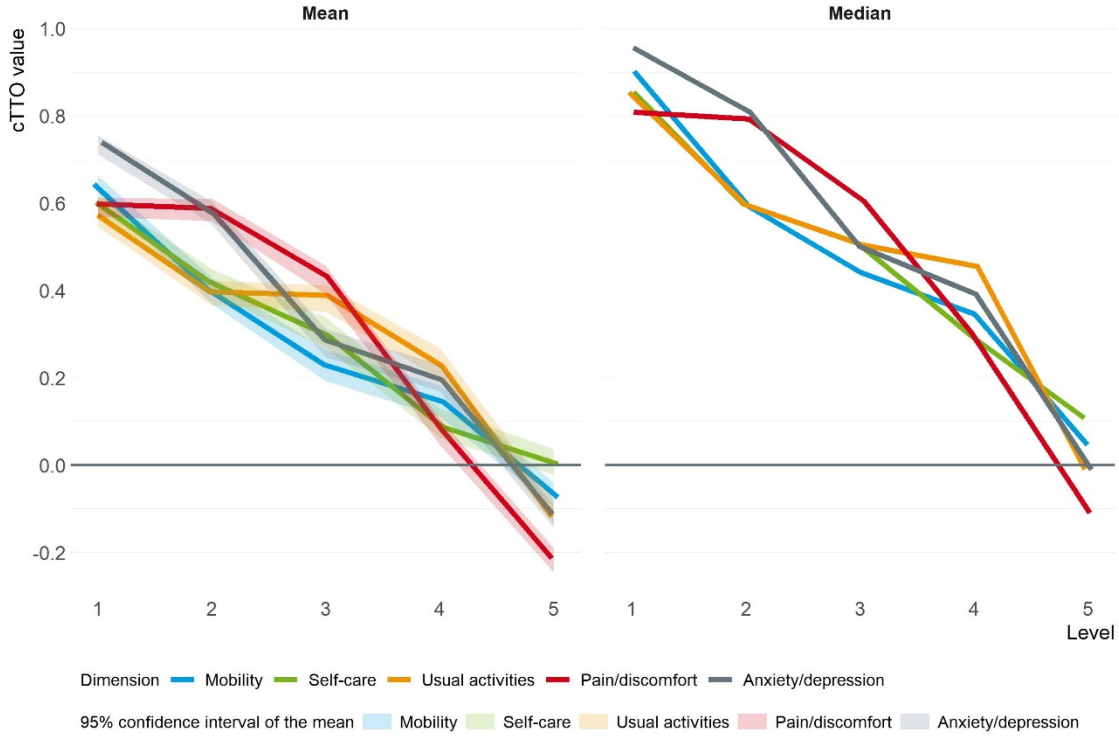
Extreme pain or discomfort leads to the highest utility loss

For each dimension, we subdivided the observed health states by level of severity in that dimension and computed the mean and median cTTO value. In Figure 16, the mean and median values were plotted against the level in the dimension. For each dimension, there is an overall negative relation between the level of the health issue and the utility value although certain transitions between levels do not lead to a utility loss. The mean and median value for no pain/discomfort and slight pain/discomfort are (nearly) the same, as well as the mean value for slight and moderate problems doing usual activities. This might lead to problems with logical consistency when estimating the ADD20 regression models (see section 3.3.1).

For most dimensions, the drop in median utility is most pronounced when going from severity level 4 to 5. There is no clear pattern for the mean value. While pain or discomfort appear to be relatively acceptable at level 2 or 3 compared to issues of similar severity in other dimensions, we find that health states with severe or extreme pain or discomfort are associated with the lowest utility values.

Note that there are no indications in Figure 16 that the mistranslation of severity level 2 for mobility in the Dutch questionnaire (see section 2.1.2) has led to any valuation problems.

Figure 16 – Mean and median cTTO values by level and dimension (unconscious state not included)

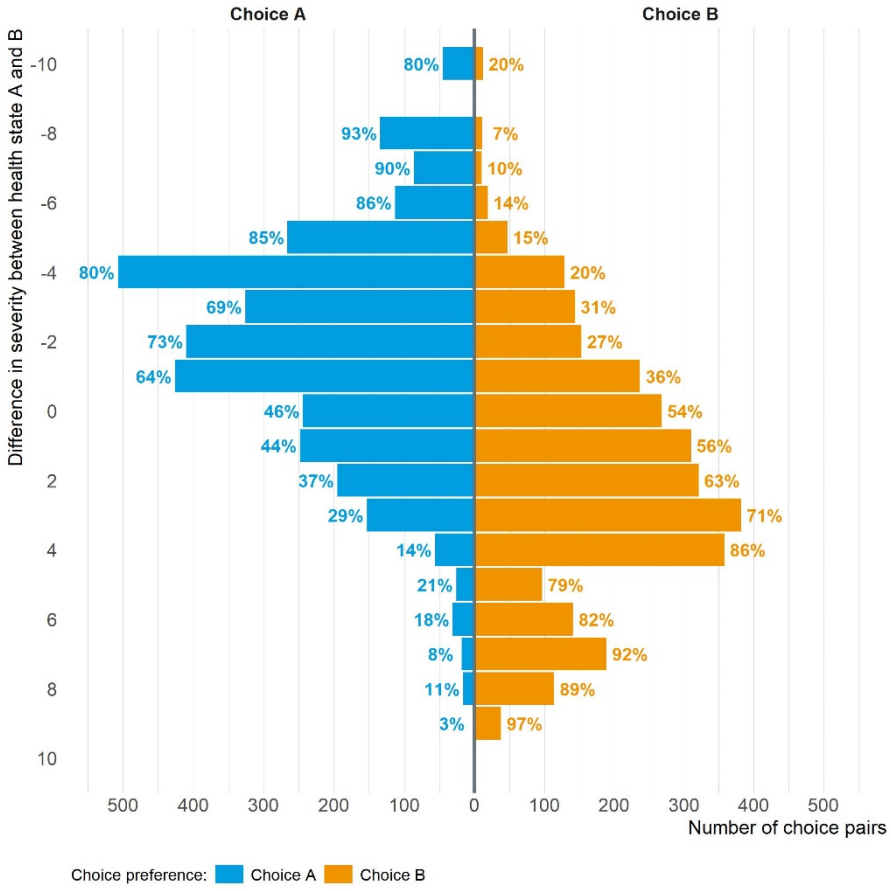


4.2.2 DCE data

In the DCE task, no choice pairs had to be evaluated where one alternative logically dominates the other alternative (see section 2.1.1). However, there can be differences in the severity of the choice alternatives. In Figure 17, the number of responses with preference for choice A or B was assessed in relation to the difference in severity between both health states. The vertical axis indicate the severity difference between health state A and B with positive values in case the severity of A exceeds B, a value of zero in case of equal severities and negative values when the severity of A is lower than B. In addition to the number of responses with preference for choice A or B on the horizontal axis, also the fraction of choice A or B per severity difference is indicated next to each bar.

The results confirm the hypothesis that the DCE choice was driven by the relative severity difference between both states. The health state with the lowest severity is favoured with the preference being more pronounced as the severity difference grows. In absence of a severity difference, a more or less equal split between choices A and B is observed. This supports the face validity of the DCE data.

Figure 17 – Preference for alternative A or B in DCE task by relative severities of A and B



4.3 Modelling results and selection

A wide range of potential models were estimated and evaluated using the model selection criteria with the aim to narrow down the number of alternatives to one preferred model. Ideally, the preferred model performs strongly on all criteria.

4.3.1 Logical consistency

The appropriateness of the models was assessed by evaluating the logical consistency of the estimated coefficients. In the preferred model, the disutility in a dimension is not allowed to decrease when health conditions worsen.

Exclusion of the ADD20 models for inconsistencies in the coefficients

For all cTTO models as well as the hybrid models, the ADD20 specification had an inconsistent ordering of levels 2 and 3 in the usual activities dimension, with disutility of level 2 exceeding disutility of level 3. This was already identified as a potential risk in section 4.2.1. Also, the DCE model with the ADD20 specification shows inconsistencies for level 2 and 3 in the dimensions mobility and usual activities. The high level of flexibility in the estimation of the coefficients in the ADD20 models comes at a cost in terms of consistency. Therefore, the ADD20 models as such are not further considered. Alternatively, an adjusted version of the ADD20 specification enforcing consistency in the usual activities dimension was estimated; in what follows, it is referred to as adjusted-ADD20.

Illogically ordered coefficients were found in almost none of the multiplicative specifications for the cTTO, DCE and hybrid models estimated on the modelling dataset. The exception is the MULT11 model estimated on the DCE data and the cTTO and hybrid version of the MULT11 model without intercept, without random effects, and with correction for heteroscedasticity and censoring. In both cases, the point estimates for the coefficients of levels 2 and 3 in the dimensions mobility, self-care and usual activities were inconsistent.

4.3.2 Goodness of fit

All cTTO models were evaluated in terms of goodness of fit. Table 4 presents the 15 models with the best performance, ordered by the sum of the ranks in BIC value and MAE value. The following conclusions can be drawn.

First, multiplicative models perform better in terms of BIC, while additive models perform better in terms of MAE.

cTTO models with intercept, random effects and correction for heteroscedasticity rank best in terms of goodness of fit

Second, models with intercept (12 out of 15 models), random effects (12 out of 15 models) and with correction for heteroscedasticity (10 out of 15 models) rank highest and are thus preferred. The inclusion of random effects leads in particular to improvements in the likelihood value and hence BIC value, implying that the model is a more likely representation of the underlying data. In terms of MAE, including random effects can lead to both better and worse performance, as can be seen when comparing for example models with rank 1 and 8, 2 and 9 or 6 and 10. A correction for heteroscedasticity leads to both a better BIC and MAE value, compare e.g. models with rank 1 and 5, 2 and 11, 4 and 14, or 6 and 15. Whether or not to include an intercept has only a minor effect on the BIC value, but utility values predicted by a model with intercept are generally in closer agreement with the observed utility values leading to lower MAE value. Only a limited number of models listed in Table 4 do not have an intercept, simply because they are outperformed in terms of MAE by similar models with intercept.

Third, models that accounted for censoring do not perform very well with respect to goodness of fit. As discussed in section 3.3.4, this was expected as the predicted values from a censored model deliberately deviate from the observed values, under the assumption that some respondents would prefer to value certain health states lower than allowed in the cTTO valuation. Hence the choice to account or not for censoring should be based on other considerations (see below).



Table 4 – Evaluation of goodness of fit

	Specification	Options	Sum of ranks	BIC value	BIC rank	MAE value	MAE rank
1	Adjusted-ADD20	Intercept, random effects, correction for heteroscedasticity	11	10 541	7	0.052	4
2	MULT11	Intercept, random effects, correction for heteroscedasticity	15	10 513	5	0.055	10
3	Adjusted-ADD20	No intercept, random effects, correction for heteroscedasticity	15	10 542	8	0.054	7
4	MULT9	Intercept, random effects, correction for heteroscedasticity	21	10 497	2	0.056	19
5	Adjusted-ADD20	Intercept, random effects	24	10 657	15	0.054	9
6	MULT8	Intercept, random effects, correction for heteroscedasticity	24	10 492	1	0.056	23
7	Adjusted-ADD20	No intercept, random effects	29	10 658	16	0.055	13
8	Adjusted-ADD20	Intercept, correction for heteroscedasticity	30	12 354	28	0.051	2
9	MULT11	Intercept, correction for heteroscedasticity	32	12 321	21	0.055	11
10	MULT8	Intercept, correction for heteroscedasticity	34	12 308	18	0.055	16
11	MULT11	Intercept, random effects	38	10 638	13	0.057	25
12	Adjusted-ADD20	No intercept, random effects, correction for heteroscedasticity and censoring	39	12 390	31	0.054	8
13	Adjusted-ADD20	Intercept, random effects, correction for heteroscedasticity and censoring	43	12 390	29	0.054	14
14	MULT9	Intercept, random effects	45	10 631	10	0.057	35
15	MULT8	Intercept, random effects	46	10 626	9	0.058	37

Note: MAE = Mean Absolute Error, BIC = Bayesian Information Criterion

4.3.3 Predictive accuracy

A cross-validation analysis was applied to all cTTO models. Three different techniques were used: leave-one-state-out, leave-one-block-out, and leave-random-block-out (see section 3.3.3). Table 5 presents a summary of the results for the 15 models with the lowest sum of MAE of the three different techniques. Models that were also among the best performers in terms of goodness of fit are indicated in bold (see Table 4). The overlap is substantial with 8 out of 15 models.

The MULT8 specification with intercept, random effects and correction for heteroscedasticity is the best performing model in terms of predictive accuracy, with a high rank in each of the cross-validations. It ranked 6th in terms of goodness of fit. Most models in Table 5 have an intercept. The impact of including random effects is less clear, with 8 out of 15 models having random effects, but 4 of them are ranked in the top 5. Only 5 models in Table 5 correct for heteroscedasticity and none for censoring.

The logical consistency of all cross-validation estimations was assessed and revealed a number of inconsistencies. Various alternative versions of the adjusted-ADD20 specification had illogically ordered coefficients, in particular with regard to levels 2 and 3 as well as 4 and 5 in the self-care dimension and level 1 and 2 in the pain/discomfort dimension. It was decided to not further enforce consistency in these dimensions. Inconsistencies were also identified in a limited number of multiplicative models. Models were excluded in case more than 5% of the cross-validation estimations showed inconsistent orderings. They are indicated in red in Table 5. Note that this exclusion did not alter our final judgement on the preferred model.

Table 5 – Evaluation of predictive accuracy

	Specification	Options	Sum of MAE	MAE state out	rank state out	MAE block out	Rank block out	MAE random blocks	Rank random blocks
1	**MULT8**	**Intercept, random effects, correction for heteroscedasticity**	**0.192**	**0.062**	**2**	**0.066**	**1**	**0.064**	**4**
2	MULT8	Intercept, random effects	0.193	0.063	5	0.066	3	0.064	6
3	MULT8	Intercept	0.194	0.064	11	0.066	4	0.063	3
4	**Adjusted-ADD20**	**No intercept, random effects**	**0.194**	**0.063**	**4**	**0.068**	**8**	**0.063**	**1**
5	**MULT9**	**Intercept, random effects**	**0.195**	**0.064**	**8**	**0.067**	**5**	**0.064**	**7**
6	**MULT9**	**Intercept, random effects, correction for heteroscedasticity**	**0.195**	**0.062**	**1**	**0.068**	**11**	**0.065**	**12**
7	MULT9	Intercept	0.196	0.065	15	0.067	6	0.064	5
8	Adjusted-ADD20	Intercept, random effects	0.197	0.065	13	0.068	9	0.064	8
9	Adjusted-ADD20	No intercept	0.197	0.066	19	0.068	12	0.063	2
10	**MULT11**	**Intercept, random effects**	**0.197**	**0.064**	**12**	**0.069**	**14**	**0.065**	**11**

11	**MULT11**	**Intercept, correction for heteroscedasticity**	**0.198**	**0.064**	**10**	**0.066**	**2**	**0.068**	**35**
12	MULT11	Intercept	0.198	0.066	18	0.068	13	0.064	10
13	**MULT8**	**Intercept, correction for heteroscedasticity**	**0.198**	**0.064**	**7**	**0.070**	**17**	**0.065**	**16**
14	Adjusted-ADD20	Intercept	0.199	0.066	27	0.068	10	0.064	9
15	**Adjusted-ADD20**	**Intercept, random effects, correction for heteroscedasticity**	**0.200**	**0.066**	**24**	**0.067**	**7**	**0.066**	**22**

4.3.4 Theoretical considerations

Some model features may be desirable from a theoretical point of view, but may reduce the model's performance in terms of predictive accuracy or goodness of fit as measured above. Additional considerations were used to substantiate the choice to (not) correct for heteroscedasticity, to (not) correct for censoring and to (not) use the hybrid model.

The error term is heteroscedastic

The descriptive analysis in section 4.2.1 demonstrated that the variability in reported utility values increases with worsening health states, an indication of heteroscedasticity in the cTTO data. Moreover, a visual inspection of the variability in the error terms by the predicted values of the estimated models clearly showed that the error term is heteroscedastic. An observation that was corroborated by two formal tests of heteroscedasticity, the White's test and the modified Breusch-Pagan test.[100] Both tests were performed for each of the four specifications – ADD20, MULT8, MULT9 and MULT11 (see section 3.2.1) – using an intercept, but without random effects or censoring and confirm the presence of heteroscedasticity ($p<0.001$).

Hence, irrespective of the performance in terms of goodness of fit or predictive accuracy, there is a clear argument to favour a model that corrects for heteroscedasticity.

Censoring cTTO values reduces correspondence with DCE results

The DCE data are uncensored by design. Under the assumption that both the cTTO reported utilities and DCE choices were driven by the same underlying preferences, an assessment of the agreement between the DCE model and the cTTO models with and without correction for censoring, might provide suggestive evidence on the need for treating the cTTO data as censored. Note that the DCE predicted values were rescaled using the rescaling factor from the hybrid model to facilitate the comparison with the cTTO results.

Table 6 and Figure 18 present the results of the assessment for the 2 best performing cTTO models in term of predictive accuracy that account for heteroscedasticity and have a strong goodness of fit. It concerns the MULT8 and MULT9 model with intercept, random effects and correction for heteroscedasticity, ranked, respectively, 1st and 5th for predictive accuracy and 6th and 4th on goodness of fit. The evaluation of other specifications has been made and leads to the same conclusions, but it is not presented here.

The DCE fit in Table 6, which was calculated as the mean absolute difference between the health states predicted by the DCE model and the cTTO models (see section 3.3.4), indicates that predicted values from the uncensored cTTO model are in better agreement with the predicted values from the DCE model than the predicted values from the censored model. The same conclusion holds when comparing the DCE fit of the censored and uncensored hybrid models. This suggests that treating the cTTO data as censored does not improve the estimation of the underlying preferences.

A similar conclusion can be drawn from the scatterplot in Figure 18 that graphically illustrates the relation between the predicted values of the cTTO and DCE models of the 86 health states observed in the cTTO valuation. Values (depicted by the blue dots) below [above] the diagonal (solid red line) imply that the utility value predicted by the cTTO model is lower [higher] than the value predicted by the DCE model. A perfect fit would imply that all values are located on the diagonal. A good fit on average would mean that the blue dots are balanced around the diagonal and that the linear regression line (dashed black line) between the DCE values and the cTTO values is very close to the diagonal. The comparison of panels A and B as well as C and D indicates that the uncensored model is on average in better agreement with the DCE results. On the one hand, for the worst state the agreement between the DCE and the censored model is better. On the other hand, in the middle part of the distribution, the censored cTTO values are well below the DCE values. One explanation is that throughout the cTTO valuation, a large range of health states have been given the lowest possible utility value of -1 (see Figure 13), leading to lower predicted utility values for all these health states when the data are assumed to be censored.

The consequence is that censoring leads to a significant increase in the number of negatively valued health states (see Table 6), which is in fact already substantial in the uncensored models. For comparison, the percentage of negatively valued states in the value set was about 15.5% in the Netherlands, 15.1% in Germany, 8.2% in Spain, 6.9% in France, 5.1% in England, and 1.8% in Canada.[39, 40, 43, 44, 72] Although the value sets of other countries should not be treated as benchmark because preference may differ profoundly between societies, it is striking that in all countries, the percentage of states valued worse than dead is lower than the predictions in the uncensored models in Table 6.

We conclude that the available evidence favours the use of an uncensored model.

High level of agreement between DCE and cTTO predicted values justifies the use of a hybrid model

Panel A and C in Figure 18 demonstrate a high level of correspondence between the predicted values from the uncensored cTTO models and DCE models. The point estimates of the coefficients of the corresponding cTTO and DCE models as well as the hybrid model are given in Table 7. The coefficients of the DCE model were rescaled using the rescaling factor of the hybrid model. As discussed in section 3.2.2, no intercept is included in the DCE model. As a validity check, the DCE model with intercept was also estimated and the results indicated that the intercept was not significant at $p<0.05$, justifying the assumption.

The preference ranking of the dimensions is consistent between the DCE and cTTO model. Disutility is lowest and not significantly different for the dimensions self-care and usual activities, mobility has a slightly higher level of disutility, the disutility associated with anxiety and depression is substantially higher and the most important source of disutility is the dimension pain and discomfort. The results from the MULT8 and MULT9 model are highly similar.

The level coefficients are slightly different. In the DCE model, there is a higher utility decrement going from level 1 to 2 and from level 4 to 5, and a lower decrement moving from level 2 to 3 and from level 3 to 4, compared to the results from the cTTO model. In both models, the transition from level 3 to 4 is accompanied with the sharpest drop in utility.

The high level of agreement between the cTTO and DCE results, in particular in the preference ranking of the dimensions, warrants the use of a hybrid model in which the coefficients can be estimated with higher precision.

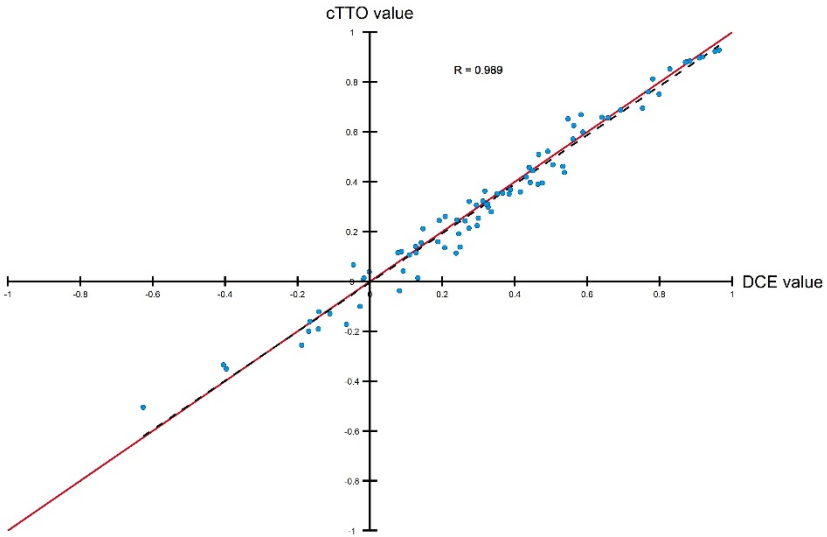
Table 6 – Evaluation DCE fit

	MULT8 – intercept, random effects, correction for heteroscedasticity		MULT9 – intercept, random effects, correction for heteroscedasticity	
	Not censored	Censored	Not censored	Censored
DCE fit	0.040	0.045	0.041	0.045
Negatively valued states (out of 3 125 states)	498 (15.9%)	584 (18.7%)	505 (16.1%)	596 (19.1%)

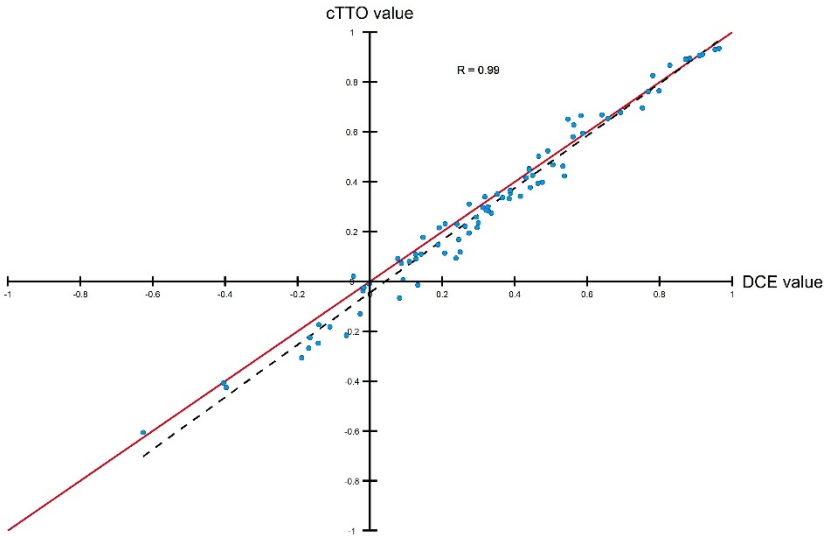
Figure 18 – Agreement between DCE predicted values and cTTO predicted values for the 86 health states used in the valuation

MULT8 specification – intercept, random effects, correction for heteroscedasticity

A. Not censored

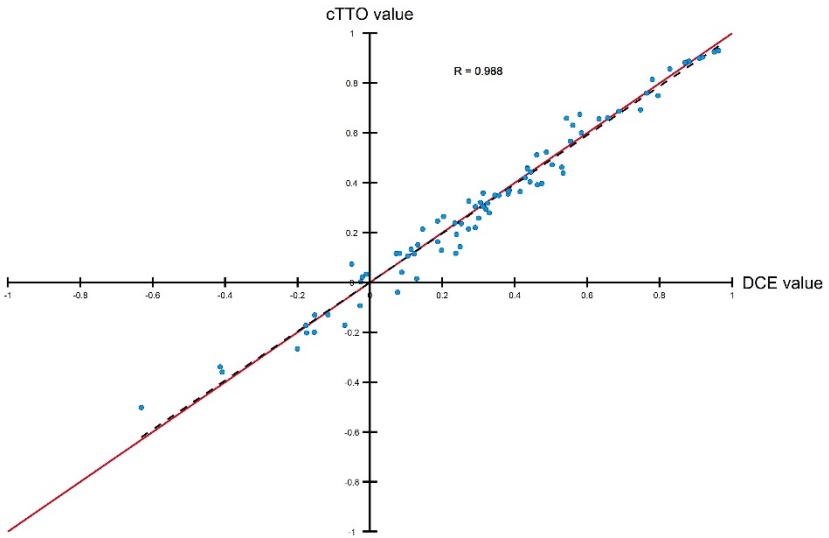


B. Censored

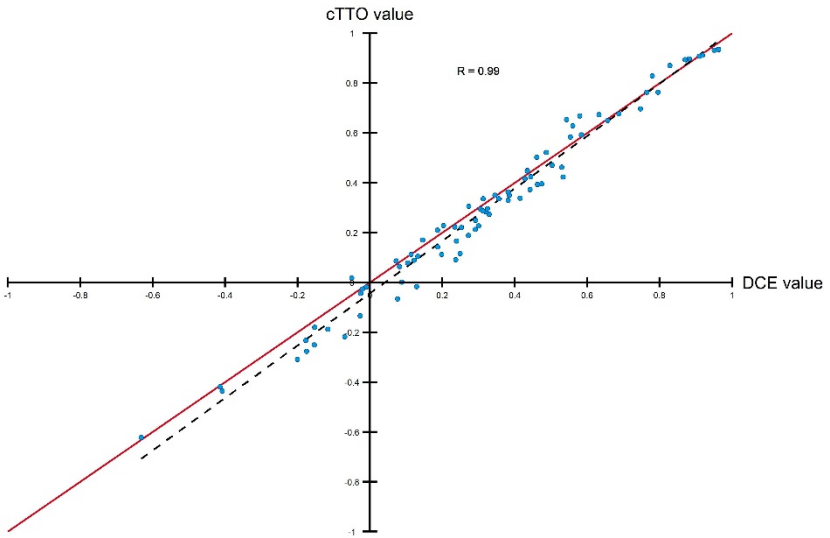


MULT9 specification – intercept, random effects, correction for heteroscedasticity

C. Not censored



D. Censored



Note: The diagonal is presented by the red line and indicates an equal value for DCE and cTTO. The dashed black line is the linear regression line between the cTTO values and the DCE values. The R-value is the correlation between the cTTO values and the DCE values.

Table 7 – Coefficients of cTTO, DCE and hybrid models of MULT8 and MULT9 with intercept, random effects and correction for heteroscedasticity

MULT8 – intercept, random effects, correction for heteroscedasticity				MULT9 – intercept, random effects, correction for heteroscedasticity			
Coefficients	cTTO	DCE	Hybrid	Coefficients	cTTO	DCE	Hybrid
$\boldsymbol{Intercept}$	**0.055**		**0.038**	$\boldsymbol{Intercept}$	0.054		0.037
$\boldsymbol{\beta_{MO}}$	0.214	0.265	0.227	β_{MO}	0.210	0.264	0.222
$\boldsymbol{\beta_{SC}}$	0.162	0.199	0.166	β_{SC}	0.161	0.200	0.169
$\boldsymbol{\beta_{UA}}$	0.160	0.200	0.181	β_{UA}	0.157	0.200	0.179
$\boldsymbol{\beta_{PD}}$	0.476	0.507	0.482	β_{PD}	0.455	0.491	0.456
$\boldsymbol{\beta_{AD}}$	0.438	0.456	0.439	β_{AD}	0.416	0.443	0.415
$\boldsymbol{L_2}$	0.100	0.178	0.139	$\boldsymbol{L_2}$	0.102	0.184	0.144
$\boldsymbol{L_3}$	0.254	0.236	0.258	$\boldsymbol{L_3}$	0.268	0.242	0.266
$\boldsymbol{L_4}$	0.847	0.706	0.788	$\boldsymbol{L_4}$	0.885	0.726	0.824
				L_5	1.056	1.037	1.061

MULT9 not statistically different from MULT8

The distinguishing factor between the MULT8 and MULT9 specification is the L_5 coefficient relaxing the proportionality assumption between the levels in all dimensions by allowing the level 5 coefficient in the dimensions pain/discomfort and anxiety/depression to differ from the level 5 coefficient in the dimensions mobility, self-care and usual activities (see section 3.2.1). The level 5 coefficient for the latter three dimensions is standardized at 1. In other words, the MULT8 specification can be seen as a special case of the MULT9 model with L_5 constrained to 1. As can be seen in Table 7, the L_5 coefficient is, only marginally different from 1. A likelihood ratio test on the cTTO, DCE and hybrid models does not reject the null hypothesis (at $p<0.05$) that the constrained MULT8 model provides as good a fit for the data as the more flexible MULT9 model. In addition, a Wald test indicates that the L_5 coefficient is not statistically different from 1 in the cTTO, DCE and hybrid model (at $p<0.05$).

4.4 Value set

4.4.1 Summary of model selection

A wide range of potential models were estimated and submitted to a thorough selection procedure. Taking all elements into consideration, the parsimonious hybrid MULT8 model with intercept, random effects and correction for heteroscedasticity has been selected as preferred model. We summarize the main reasons once more:

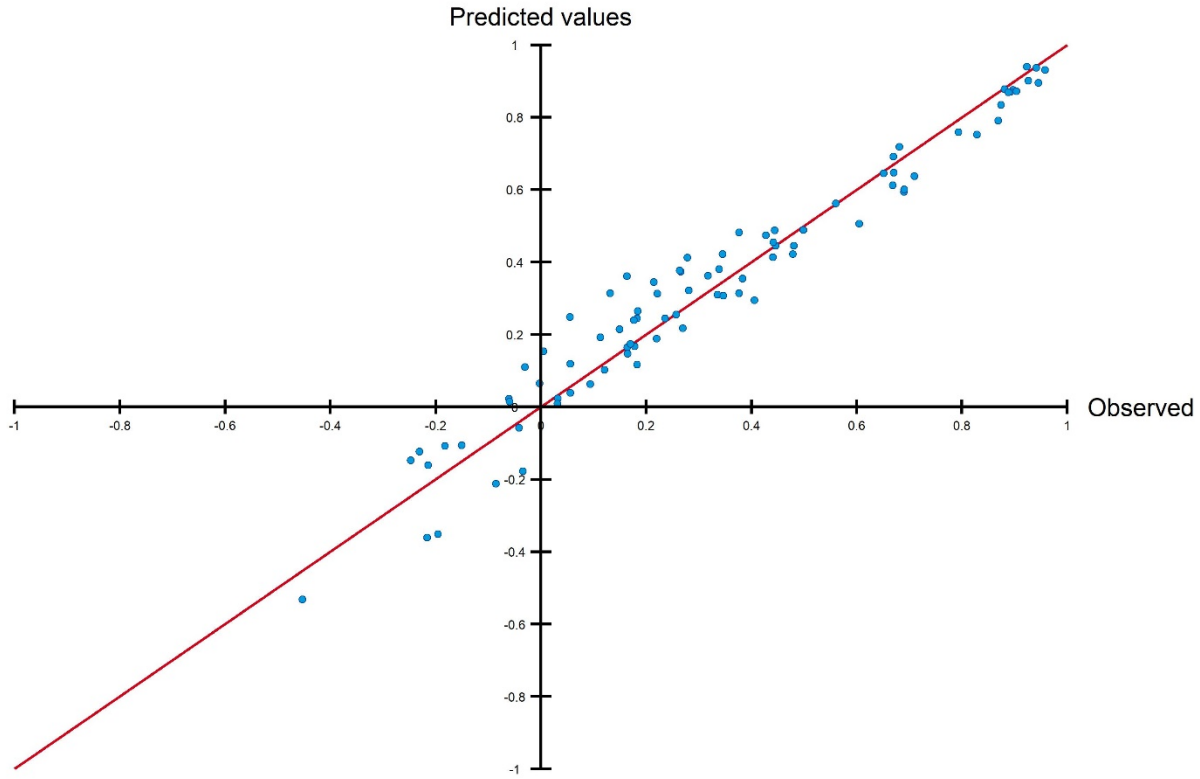
1. The coefficients are logically consistent. Logically consistency was problematic for the ADD20 specification.
2. The model corrects for heteroscedasticity
3. The MULT8 model with intercept, random effects, and correction for heteroscedasticity performs strongly in terms of goodness of fit. Overall it was ranked 6th, but it had the lowest BIC-value indicating that this model has the best trade-off between goodness of fit and parsimony.
4. The MULT8 model with intercept, random effects, and correction for heteroscedasticity outperformed all other models in the cross-validation analysis. This criterion weighs more heavily in the selection than goodness of fit given that the final value set is predicted based on only a limited number of observed health states.
5. The data are not treated as censored because the predictions from uncensored models give a better fit with the predicted values of the DCE model. Moreover, the number of negatively values states in the uncensored model is already substantial.
6. The high level of agreement between DCE and cTTO predicted values justifies the use of a hybrid model
7. The MULT8 model is preferred over the MULT9 model, given that the additional coefficient in the MULT9 model was not statistically different from the constrained value in the MULT8 model and the better predictive accuracy of the MULT8 model.

4.4.2 Preferred model and value set

The hybrid MULT8 model with intercept, random effects and correction for heteroscedasticity was selected as preferred model. The estimated coefficients and the bootstrapped standard errors are provided in Table 8. Figure 19 shows that there is a good fit between the predicted and observed values of the 86 health states used in the valuation exercise. Values (depicted by the blue dots) below [above] the diagonal (solid red line) imply that the utility value predicted by the preferred model is lower [higher] than the average utility value reported in the cTTO valuation.

The robustness of the point estimates was checked by fitting the model on three alternative datasets, one larger and two smaller datasets: (1) the full sample; (2) the modelling data restricted to interviews collected prior to the COVID-19 pandemic; and (3) the modelling dataset excluding all cTTO valuations flagged in the feedback module. The coefficients of all three robustness checks are presented in Table 8 and are broadly in line with the results of the preferred model, with an average absolute difference of only 0.004, 0.002 and 0.005 between the coefficients of the preferred model and the coefficients from, respectively, the full sample, the sample limited to pre-COVID interviews and the sample excluding flagged interviews. The largest deviation was observed in the pain dimension with the coefficient of the preferred model being 0.017 above the coefficient obtained from the full sample. In all sensitivity results, the preference ranking of the dimensions was identical. This suggest that the results from the preferred model are robust to the choices made with respect to the exclusion criteria and the extension of the interview period after the first lockdown to increase representativeness in certain regions (see section 2.4.3).

Figure 19 – Scatterplot of predicted (hybrid) and observed (cTTO) utility values for the 86 health states used in the valuation



Note: The diagonal is presented by the red line and indicates an equal predicted and observed utility value.

Table 8 – Coefficients and bootstrapped standard errors of the preferred model and results of the sensitivity analysis

	Preferred model				Sensitivity analysis		
	Coefficient value	Standard error	T-statistic	P value	Full sample	Restricted sample pre-COVID	Restricted sample feedback module
$Intercept$	0.038	0.0148	2.55	0.005	0.037	0.041	0.033
β_{MO}	0.227	0.0102	22.19	<0.0001	0.220	0.228	0.230
β_{SC}	0.166	0.0108	15.32	<0.0001	0.165	0.166	0.170
β_{UA}	0.181	0.0098	18.37	<0.0001	0.181	0.180	0.184
β_{PD}	0.482	0.0138	34.84	<0.0001	0.466	0.481	0.492
β_{AD}	0.439	0.0132	33.13	<0.0001	0.434	0.446	0.447
L_2	0.139	0.0154	9.02	<0.0001	0.144	0.136	0.139
L_3	0.258	0.0158	16.34	<0.0001	0.262	0.258	0.254
L_4	0.788	0.0157	50.26	<0.0001	0.787	0.789	0.781

Box 5 – Examples on the calculation of EQ-5D-5L utility values

The utility value of any of the 3 125 possible EQ-5D-5L health states can be calculated using the coefficients of the preferred regression model (in Table 8). The coefficients indicate the disutility associated with each dimension and level relative to a situation in full health. For each dimension, the coefficient value of the dimension is multiplied by the level coefficient. The coefficients for level 1 and 5 are not estimated but standardized at 0 and 1, respectively. The resulting disutilities in each dimension and the intercept are summed together and subtracted from 1, the utility value of perfect heath. Let us illustrate this with the following examples:

Utility associated with **health state 15342** can be calculated as follows:

- Disutility in dimension mobility for level 1 equals $\beta_{MO} \times 0 = 0.227 \times 0 = 0$
- Disutility in dimension self-care for level 5 equals $\beta_{SC} \times 1 = 0.166 \times 1 = 0.166$
- Disutility in dimension usual activities for level 3 equals $\beta_{UA} \times L_3 = 0.181 \times 0.258 = 0.047$
- Disutility in dimension pain/discomfort for level 4 equals $\beta_{PD} \times L_4 = 0.482 \times 0.788 = 0.380$
- Disutility in dimension anxiety/depression for level 2 equals $\beta_{AD} \times L_2 = 0.439 \times 0.139 = 0.061$
- Overall disutility $0.038 + 0 + 0.166 + 0.047 + 0.380 + 0.061 = 0.691$ (due to rounding, the disutility is 0.691 instead of 0.692)
- Utility value of health state 15342 is $1 - 0.691 = \mathbf{0.309}$

Utility of the **worst state 55555** equals $1 - (0.038 + 0.227 \times 1 + 0.166 \times 1 + 0.181 \times 1 + 0.482 \times 1 + 0.439 \times 1) = \mathbf{-0.532}$

Utility of **health state 54321** equals $1 - (0.038 + 0.227 \times 1 + 0.166 \times 0.788 + 0.181 \times 0.258 + 0.482 \times 0.139 + 0.439 \times 0) = \mathbf{0.491}$

The value set

The results of the preferred model were used to generate the EQ-5D-5L value set for Belgium with utility values for all 3 125 health state. A complete overview of the value set is given in Appendix 3 and can be downloaded in electronic format. Note that the utility value for the unconscious health state is not derived from the regression model, but equals the weighted average of the stated preferences. Box 5 gives some examples on how to calculate the EQ-5D-5L utility value for a specific health state using the estimates from the preferred model.

In Table 9 a comparison is made of some key characteristics of the value sets produced by the preferred model and by competing models that performed well in the selection procedure. The results indicate that the features of the value set produced by the preferred model do not strongly diverge from value sets of other potential models. The highest impact was related to the choice of a hybrid model and the choice not to treat the data as censored. First, the preference ranking of the dimensions self-care and usual activities was different between hybrid and cTTO models, with all hybrid models in Table 9 having a relatively higher utility loss related to the inability to do usual activities and the cTTO models having a higher utility loss related to the inability to wash or dress. The utility loss related to either dimension is, however, of similar magnitude. Also in other models not presented here, e.g. without random effects or correction for heteroscedasticity, the preference ranking is robust for the three dimensions with the highest utility loss while the position of the dimensions self-care and usual activities can diverge. Second, the censored models have a higher fraction of health states valued worse than dead as well as a more negative value for the worst state. This is not surprising and was already discussed in detail in section 4.3.4.

Table 9 – Comparison of key characteristics of the value sets produced by different high-performing models

	All models have intercept, random effects and correction for heteroscedasticity						
	Value set: MULT8 hybrid	**MULT8**	**MULT8, correction for censoring**	**MULT8 hybrid, correction for censoring**	**MULT9 hybrid**	**MULT11 hybrid**	**Adjusted-ADD20 hybrid**
% health states valued worse than dead	15.0%	15.9%	18.7%	18.0%	15.2%	15.1%	15.3%
Preference ranking of dimensions (ordered from highest to lowest utility loss at level 5)	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care	Pain/discomfort Anxiety/depression Mobility Self-care Usual activities	Pain/discomfort Anxiety/depression Mobility Self-care Usual activities	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care
Minimum value (state 55555)	-0.532	-0.505	-0.606	-0.630	-0.531	-0.524	-0.526
Maximum value	0.939 (state 12111)	0.929 (state 11211)	0.935 (state 11211)	0.950 (state 12111)	0.939 (state 12111)	0.929 (state 12111)	0.954 (states 11211, 11311)

4.4.3 Comparison with 3L value set and value sets of other countries

Panel A in Figure 20 presents Kernel density plots for the new EQ-5D-5L value set for Belgium, the previous value set based on the EQ-5D-3L and the related EQ-5D-5L crosswalk values. Table 10 provides a comparison of a number of key characteristics of the 3L and new 5L value set.

New value set covers a wider range of utility value

The comparison highlights a shift in utility values both to the left and right of the distribution in the new value set. The displacement becomes already clear when looking at the lowest and second highest valued health states (full health is by definition the highest), with utility at, respectively, -0.158 and 0.817 in the 3L value set and at -0.532 and 0.939 in the new value set.

The distribution of the Belgian EQ-5D-3L value set is characterized by two peaks (at about 0.20 and 0.65) with the majority of utility values concentrated between -0.15 and 0.80 (see Figure 20). About 6.6% of the health states were valued worse than dead. Second, the values in the EQ-5D-5L crosswalk are distributed similarly to the EQ-5D-3L value set. Roughly the same utility range is covered, but with a smoother transition between two peaks. Third, the new EQ-5D-5L value set covers a larger utility space with the majority of values (99.2%) between -0.4 and 0.9. The peak in density of the new value set is lower with relatively more states having a higher as well as lower utility value than previously, as indicated by the blue density line above the other lines from utility value 0.4 onwards and below utility value 0.05. There is a higher share (15.0%) of negatively values health states.

The ranking of dimensions has altered somewhat, although anxiety/depression and pain/discomfort still have the highest impact on utility (see Table 10). In the EQ-5D-3L value set, the dimension anxiety/depression is associated with the highest utility loss, followed by

pain/discomfort and self-care and then mobility and usual activities. In the new value set, the preference ranking indicates the highest utility loss in the dimension pain/discomfort, closely followed by anxiety/depression and then mobility, usual activities and self-care.

Distribution of Belgian value set comparable to German and Dutch value sets

In Panel B in Figure 20 the distribution of the new Belgian EQ-5D-5L value set is compared to the value set in neighbouring countries. The distribution of the Belgian EQ-5D-5L value set is most similar to the distribution of the value sets in Germany and the Netherlands. Also with respect to the share of negatively valued health states, these three datasets are comparable. Nonetheless, there are differences between the Belgian, Dutch and German value sets. The preference ranking of the dimensions differs between the countries, with the dimensions pain/discomfort having the highest utility loss at level 5 in Belgium and Germany as opposed to anxiety/depression in the Netherlands. Another difference is that the dimension self-care is deemed relatively more important in Germany than in Belgium or the Netherlands, while the dimension mobility is ranked relatively lower.

The differences between the Belgian, French and English value set are more outspoken. Utility values in the French value set are generally higher, while the English value set has a higher density of small positive utility values (between 0.20 and 0.60).

Overall, the preference rankings of the dimensions in the other countries is relatively comparable to what was found for the Belgian EQ-5D-5L value set. In nearly all neighbouring countries, the dimensions anxiety/depression and pain/discomfort are associated with the highest utility decrements (with the exception of France), while the dimension usual activities has the lowest impact on utility (with the exception of the Netherlands).

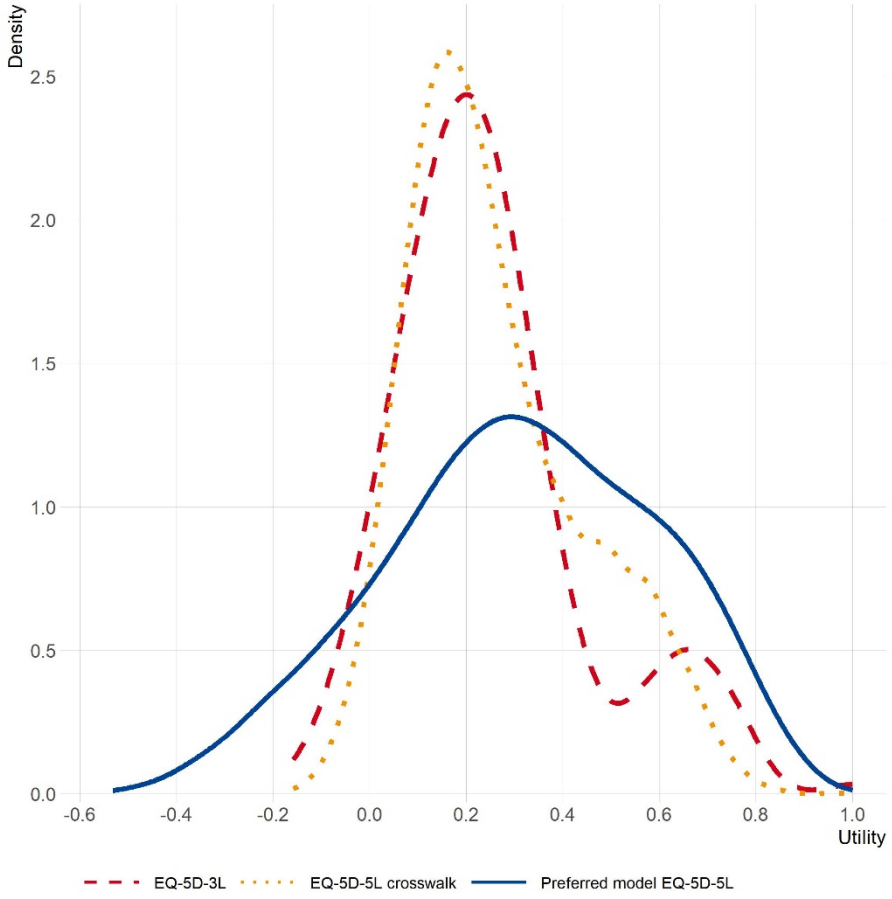
Table 10 – Comparison of key characteristics of the Belgian EQ-5D-5L value set, Belgian EQ-5D-3L value and EQ-5D-5L value sets of neighbouring countries

	EQ-5D-3L value set	New Belgian EQ-5D-5L value set	German EQ-5D-5L value set	Dutch EQ-5D-5L value set	French EQ-5D-5L value set	English EQ-5D-5L value set
% health states valued worse than dead	6.6%	15.0%	15.1%	15.5%	6.9%	5.1%
Preference ranking of dimensions (ordered from highest to lowest utility loss at level 5)	Anxiety/depression Pain/discomfort Self-care Mobility Usual activities	Pain/discomfort Anxiety/depression Mobility Usual activities Self-care	Pain/discomfort Anxiety/depression Self-care Mobility Usual activities	Anxiety/depression Pain/discomfort Mobility Usual activities Self-care	Pain/discomfort Mobility Self-care Anxiety/depression Usual activities	Pain/discomfort Anxiety/depression Mobility Self-care Usual activities
Minimum value	-0.158 (state 33333)	-0.532 (state 55555)	-0.661 (state 55555)	-0.446 (states 55555, 54455, 55455, 54555)	-0.525 (state 55555)	-0.285 (state 55555)
Maximum value (except state 11111)	0.817 (state 11211)	0.939 (state 12111)	0.974 (state 21111)	0.918 (state 21111)	0.980 (state 11121)	0.950 (states 11211, 12111)

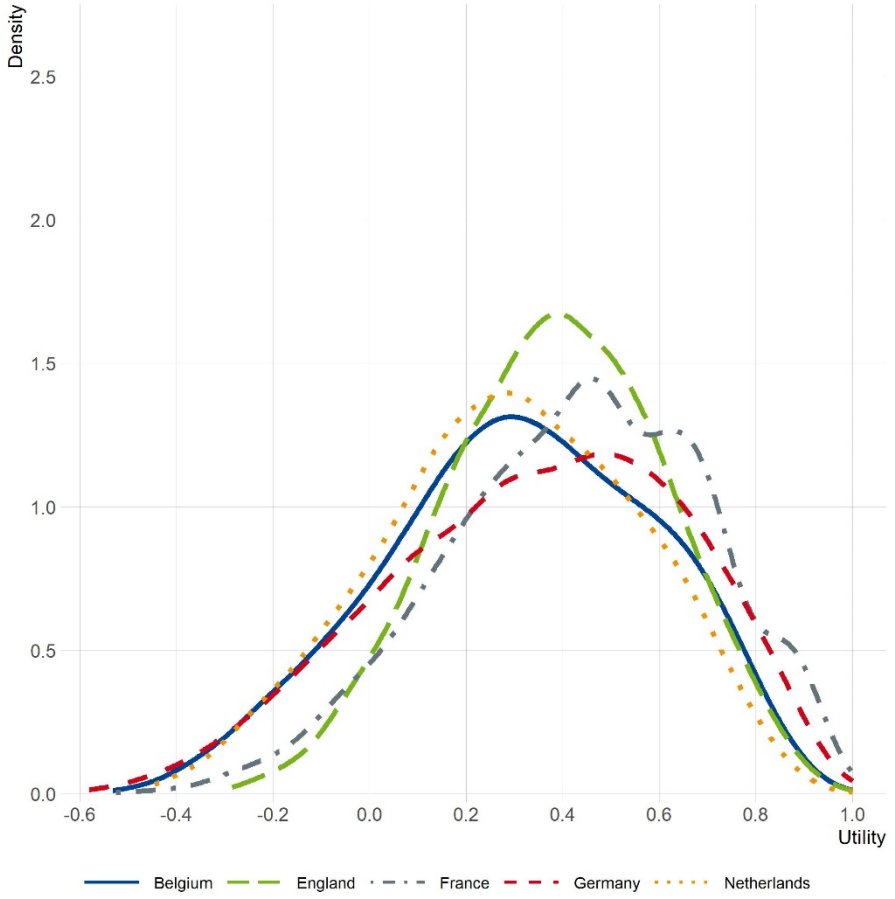
Sources: Andrade et al. (2020)[40], Devlin et al. (2018)[42], Ludwig et al. (2018)[39], Versteegh et al. (2016)[43]

Figure 20 – Kernel density plots of value sets

A. Belgian EQ-5D-5L value set compared with previous Belgian value sets based on EQ-5D-3L



B. Belgian EQ-5D-5L value set compared with EQ-5D-5L value sets of neighbouring countries



5 DISCUSSION AND CONCLUSION

The current study developed an EQ-5D-5L value set, based on the preferences of a random sample of the Belgian general public. The 5L version of the EQ-5D instrument – with five response categories reflecting the level of functioning per dimension –was created by the EuroQol Group to improve the discriminatory power and sensitivity of the previous 3L version of the instrument with three response categories per dimension–. This allows to better capture smaller but possibly still significant improvements in mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The number of health states that can be described with the EQ-5D increased from 243 in the 3L version to 3 125 in the 5L version. The EQ-5D is widely used in clinical research and economic evaluations and is important to inform health policy.

A valuation study for the EQ-5D-3L has been performed for Flanders in 2003.[47] Although it is possible to transform 3L utility values to 5L utility values by means of a mapping function (the crosswalk values), no value set specific to the EQ-5D-5L was available for Belgium.[48] In order to benefit from the full potential of the EQ-5D-5L, a Belgian EQ-5D-5L value set is indispensable. Also from a societal perspective, an update in utility values to better match the preferences of the general public was needed given the changes in demography as well as the composition of the population since 2003.

EQ-5D-5L value sets have already been determined for a large range of countries using a standardized protocol, including neighbouring countries France, England, the Netherlands and Germany.[39, 40, 42, 43] However, utility values from neighbouring countries cannot serve as a substitute for a Belgian value set. Differences between national value sets can be considerable and reflect both methodological differences and different underlying societal preferences.[101] A comparison of the Belgian EQ-5D-5L value set developed in the current study with those of other countries confirms such differences in utility values for the EQ-5D-5L health states. For example, Belgium has a much higher proportion of health states worse than dead compared to France. This underlines the importance of establishing and using national value sets in healthcare policy making.

5.1 Strengths and limitations in the development of the new value set

The EQ-5D-5L value set developed in this study is based on a strong and valid method. Compared to the study performed in 2003 that derived the preference values for the EQ-5D-3L for Flanders[47], this study has several major methodological improvements:

- This study followed the most recent EuroQol Valuation Technology (EQ-VT 2.1) protocol group for valuation studies, including extensive training of interviewers, data monitoring and quality control procedures with the aim of improving the accurateness and quality of the valuation data.[45, 46]
- The valuation techniques to obtain utility values included a standardized version of the composite Time Trade-Off (cTTO) and Discrete Choice Experiment (DCE) instead of the visual analogue scale (VAS). The cTTO and DCE are considered more robust methods for utility measurement than the VAS because they are embedded in economic random utility theory.
- Face-to-face interviews were conducted instead of postal surveys. Research has shown that cTTO valuation can be cognitive challenging and the one-to-one setting allows an interviewer to provide detailed instructions and feedback where appropriate.[45, 61]
- It measured utilities for Belgium (Brussels, Flanders and Wallonia, including representatives of the German-speaking community), while the previous value set was limited to Flanders. The extension to the entire Belgian territory is of course preferable if the values are used for national decision making.

Compared to similar studies in other countries, this study has one particular strength. Respondents were randomly selected from the National Register using a multistage, stratified, cluster sampling with unequal probability design. Using comparative information from administrative sources and reference surveys, we concluded that our final sample of 892 respondents was representative for the Belgian population in terms of age, sex, region,

health status, health-related quality of life, education and employment status. Although representativeness in these fields does not necessarily imply representativeness in terms of health preferences, it at least reduces possible impact of lack of representativeness on these factors on the utility values obtained.

A wide range of potential models were estimated and submitted to a thorough selection procedure. The range of models tested finally allowed to select the one that performed best on specific criteria. Taking all elements into consideration, the parsimonious hybrid MULT8 model with intercept, random effects and correction for heteroscedasticity has been selected as preferred model.

Several problems were, however, encountered during the data collection related to the study time frame and management of interviewers. One limitation of the study is that it was not possible to interview the targeted 1 000 people. The COVID-19 pandemic was one reason, but in general it proved very difficult to find people prepared for a face-to-face interview at their home from a list of people randomly selected from the National Register (more difficult than for example when issued from a market research database). The interview and cTTO instrument were also experienced as complex, as a consequence of which many interviews – in particular at the start of the study – were excluded from the dataset during the quality control process. In addition, existing interviewers had to be re-trained because of insufficient quality of the data, some interviewers dropped out or were removed, and new interviewers had to be recruited and trained. All of this created delays in the study and stretched the researchers and subcontractor. Although there is evidence to support the face validity of the data used to develop the value set, it seems that part of the variability in the data, and in particular in health states valued worse than dead in the cTTO valuation, can be attributed to differences between interviewers.

A lesson learnt is that performing a health state valuation study in a representative sample of the general population requires a huge amount of resources. More specifically, the use of a strict sampling procedure to ensure representativeness of the sample, i.e. a multistage, stratified, cluster sampling with unequal probability design combined with a strict interview protocol and thorough quality control made the task of interviewers complex, required close management of the interviewers and slowed down the study.

5.2 Implications of the new value set

The new EQ-5D-5L value set updates the currently used EQ-5D-3L value set from 2003 with 243 potential health states (and the related EQ-5D-5L crosswalk value set) to a more extensive value set with 3 125 potential health states. Compared to the 3L value set, the 5L value set has a higher fraction of health states valued worse than dead and covers a wider range of utility values, ranging from -0.532 (worst state 55555) to 0.939 (state 12111), when excluding state 11111 which is valued by definition at the maximum value of 1. The 3L value set ranged from -0.158 (worst state 33333) to 0.817 (state 12111). The higher valuation of mild health states reflects the very limited willingness to trade-off life-years to avoid mild health problems that was observed among survey respondents.

A wider range of utility values allows for a better discrimination between patients, both for mild and (very) severe health conditions. This might improve the applicability of the EQ-5D instrument, and is expected to lead to diverging results and conclusions when using the EQ-5D-5L or EQ-5D-3L as patient-reported outcome measure, in cost-effectiveness analyses or to assess population health (see section 5.3).

The preference ranking of the dimensions has altered somewhat. In the 3L value set, the dimension anxiety/depression is associated with the highest utility loss, followed by pain/discomfort and self-care and then mobility and usual activities. In the 5L value set, the dimension pain/discomfort is associated with the highest loss in health-related quality of life, closely followed by anxiety/depression and then mobility, usual activities and self-care. This suggests that policies/interventions related to pain relief and mental care have a high potential for gains in health-related quality of life.

5.3 Applicability of the Belgian EQ-5D-5L value set

The value set established in the current study will provide major opportunities for future clinical and economic evaluations in healthcare as well as population health assessments. Such evaluations are useful to inform healthcare policy decisions. The use of a single Belgian value set allows for consistent decision-making across patient populations and sociodemographic groups in the country. This is for instance important for resource allocation decisions and priority setting in the healthcare sector, where comparisons between different patient populations and health problems must inevitably be made.

For the **assessment of the impact of health interventions on the generic HRQoL** of patients, it suffices to collect data on the descriptive system of the EQ-5D-5L in (patient) populations who did and populations who did not receive the intervention to allow valid comparisons between the outcomes in the two groups. The EQ-5D can be considered as a patient-reported outcome measure (PROM).[52, 53] Besides for impact assessment of health interventions, PROMs are increasingly used to improve quality of care, as they facilitate the communication between healthcare providers and patients. While PROMs are usually applied in an individual patient-provider relationship, the routine collection of EQ-5D-5L data (as PROM) would not only facilitate this application in clinical practice but also offer many research and assessment opportunities. A previous KCE-report already recommended the more systematic collection of a generic PROM in clinical settings for this purpose.[102] The EQ-5D-5L could serve this purpose, especially given that an up-to-date EQ-5D-5L value set is available now. The standard set of health state utilities makes sure that the outcomes are comparable between different health conditions, treatments, care providers and people, as people who describe their health state in the same way are assigned the same utility value. It avoids bias due to, for instance, differences in general attitude towards health or patient's coping with chronic conditions which may influence utility values when they are derived directly from a patient and not from the general public.[11, 13]

For **cost-effectiveness analyses** of health interventions, or more specifically the calculation of QALYs frequently used in such analyses, the value set allows for the calculation of the quality-weights for the life years gained with a specific intervention.[3, 11, 13, 25] Cost-effectiveness analyses performed with the 3L value set and the new 5L value set are expected to produce different results.[103] For example, the higher utility values of mild health states in the 5L value set would impact the appraisal of technologies targeting very mild health problems, likely leading to lower incremental quality-adjusted life-years gained and lower estimates of cost effectiveness.[80, 103] As the distribution of utility values is substantially different between the 3L and 5L value sets, it is difficult to predict for a specific intervention whether QALY gains will be larger or smaller with the 3L or 5L value set, the effect will vary from one intervention to another. Given the rigour with which the current value set was established, it can be assumed that the cost-effectiveness estimates will be more accurate with the new value set. This may, however, raise issues of historical consistency between past (with 3L) and future (with 5L) economic evaluations and pricing decisions.

For the **assessment of population health**, as is done for instance in the Health Interview Survey (HIS), the value set will allow multiple quantitative analyses and comparisons between population groups. By assigning the utility scores to the health states described by the participants in the HIS, analyses can extend beyond the purely descriptive statistics. Relationships between population characteristics (including health behaviour (e.g. smoking), chronic illness, socioeconomic status, or any other variable collected in the HIS) and HRQoL can be assessed. Such analyses were already done with the previous value set but can now be updated with the new values, the advantage being that no crosswalk values need to be used but utility values from the EQ-5D-5L value set can be assigned directly to the EQ-5D-5L health state descriptions of the participants to the HIS.

Combining the EQ-5D-5L health states described by the participants in the HIS with the value set also allows the **establishment of population 'norms'** (or reference values).[49-51] Such norms allow more valid comparisons between patients' HRQoL and the HRQoL of people from the general population with similar characteristics (age, sex) as the patient. These comparisons could for example be used for assessing which patient populations are relatively worse off compared to others, or, in other words,

have the highest unmet needs, taking the "normal" level of HRQoL of comparable people from the general public into account. It is often not realistic to assume that the disutility of patients can be fully resolved by healthcare, as becoming older might have an impact on health that is unavoidable. Hence, a comparison with the "reference value" or "population norm" for HRQoL should be made when assessing the level of HRQoL of patient populations. Population norms have been established for Belgium using the crosswalk value set for the EQ-5D-5L based on the EQ-5D-3L value set by Sciensano (paper submitted).[50]

5.4 Future research and reflections

The EQ-5D-5L value set is based on the preferences of a random sample of the Belgian adult population and can be used to assign values for health-related quality of life to health states expressed by adults. It is not adjusted for children and adolescents, for which a separate instrument has been developed to describe the health and health-related quality of life, the EQ-5D-Y targeted at children and adolescents aged between 8 and 15.[27, 104, 105] A Belgian value set for the EQ-5D-Y is currently being developed. Adolescents between 16 and 18 can be assessed using either the EQ-5D-Y or one of the adult versions.[27] A separate value set for children was deemed necessary because evidence suggest that adults and children perceive and value health differently leading to different utility values for the same health state.[27] An EQ-5D-Y value set makes it possible to e.g. have PROMs adjusted to children and adolescents or perform a cost-effectiveness analysis of interventions for children and adolescents. The availability of a separate value set for children and adolescents raises, however, the question whether or not separate value sets for other age groups or population groups, e.g. the elderly, need to be developed. The data collected in the current study offers opportunities for future research in this sense, such as the analyses of differences between health state utilities according to respondent characteristics like age, region, current health status etc., and would offer opportunities to research the ethical implications if differentiated utility values would be used for different population subgroups. Unfortunately, time and resources did not allow to perform such analyses in the context of the current study.

In relation to the previous point, it is important to note that health care competences in Belgium are shared between the Federal state and the Federated entities (three regions based on territory and three communities based on language).[106] The Federated entities for example take decisions on and (partially) finance care for the elderly, home care, health promotion and disease prevention (e.g. screening and immunization). It might therefore be useful to have regional value sets as well. However, the number of collected interviews is insufficient to produce such regional value sets. Therefore the use of the national value set is recommended. The survey sample in the current study is representative of the population at the regional level in terms of age and sex.

Areas requiring further analysis and discussion are related to the observation that 15% of the health states are valued worse than dead and to the preference ranking among the dimensions. The former means that for 15% of the health states that can be described with the EQ-5D-5L people prefer being dead over and above these health states. Further research is needed to gain a deeper understanding of what drives the negative valuation of health states. In addition, this finding may also trigger a political debate about appropriate treatment for severe chronic illness, e.g. the desirability of continuous treatment in some circumstances versus providing palliative care. With regard to the preference ranking, it is of interest to understand why problems in the dimensions pain/discomfort and anxiety/depression have the highest negative impact on health-related quality of life according to the Belgian population and why this is the same or different in other countries.

In the longer future, it might be necessary to re-assess the utility values of health states, as it can be expected that health state valuations change with changing population health. Suppose, for instance, that health policy from now on heavily invests in pain relief treatments and mental care, leading to overall better population health on these dimensions. In that case, other dimensions might again become more important and hence utility decrements due to problems on these other dimensions might become larger. However, the cost-benefit of performing a similar study like the current one needs to be considered as well. Given that public resources are spend to it, it is not preferable to re-do the exercise every 4 years.

APPENDICES

APPENDIX 1. LIKELIHOOD FUNCTIONS IN REGRESSION ANALYSIS

All models were estimated using maximum likelihood estimation. The definition of the log-likelihood function is based on Ramos-Goñi et al. (2016)[94] and programmed using proc NLMIXED in SAS 9.4.[96]

Log-Likelihood function of the cTTO model

Based on the discussion in section 3.2.1, a cTTO model can be formulated as follows:

$$DU_i = I + \beta_{cTTO} X_j + \nu_i + \eta_{ij}$$

where DU_{ij} is the disutility related to health state j reported by respondent i; I is the intercept in the regression; $\beta_{cTTO} X_j$ represents one of the four core regression specifications with coefficients (β_{cTTO}) and variables (X_j) representing the valuation of the different dimensions and levels; ν_i the random effect associated with respondent i, ν_i is characterized by a normal distribution with zero mean and variance σ_r^2; η_{ij} is the error term. The error term is assumed to be normally distributed with zero mean and variance σ_η^2. In case heteroscedasticity is assumed, the natural logarithm of σ_η^2 is estimated in a separate regression using the same X_j variables and core regression specification. In absence of an intercept, random effects or heteroscedasticity, the specifications can be further simplified.

The log-likelihood $\ln(L)$ is defined as follows:

$$\ln(L) = -\frac{1}{2}\sum_i\sum_j\left(\ln\left(2\,\pi\,exp\left(ln(\sigma_\eta)\right)^2\right) + \left(\frac{DU_{ij} - \left(I + \beta_{cTTO}X_j + \nu_i\right)}{exp\left(ln(\sigma_\eta)\right)}\right)^2\right)$$

Part of the log-Likelihood function; for **all non-censored observations**

$$+\sum_i\sum_j\left(ln\left(\Phi\left(\frac{-\left(DU_{ij} - \left(I + \beta_{cTTO}X_j + \nu_i\right)\right)}{exp\left(ln(\sigma_\eta)\right)}\right)\right)\right)$$

Part of the log-Likelihood function; for **all censored observations**, i.e. right censored with respect to disutility, with $\Phi(.)$ the standard normal cumulative distribution function.

Log-Likelihood function the DCE model

Based on the discussion in section 3.2.2, a DCE model can be formulated as follows:

$$C_{ip} = \beta_{DCE}(X_{pA} - X_{pB}) + \varepsilon_{ip}$$

Where health states A and B in pair p are evaluated by respondent i; C_{ip} is a binary choice variable indicating whether disutility is highest for health state A ($C = 1$) or for health state B ($C = 0$); $\beta_{DCE}(X_{pA} - X_{pB})$ represents one of the four core regression specifications with coefficients (β_{DCE}) and variables ($X_{p.}$) indicating the valuation of the ifferent dimensions and levels for health states A and B; ε_{ip} is the error term. The error term is assumed to follow a logistic distribution. There is no intercept, random effect or heteroscedasticity.

The log-likelihood $\ln(L)$ is defined as follows:

$$\ln(L) = \sum_i\sum_p\left(\ln\left(\frac{1}{1+\exp\left(-\beta_{DCE}(X_{pA} - X_{pB})\right)}\right) * C_{ip} + \ln\left(\frac{\exp\left(-\beta_{DCE}(X_{pA} - X_{pB})\right)}{1+\exp\left(-\beta_{DCE}(X_{pA} - X_{pB})\right)}\right) * \left(1 - C_{ip}\right)\right)$$

Log-Likelihood function of the hybrid model

The hybrid model used in the current study uses one likelihood function to combine the DCE and cTTO part of the data in order to maximize the available information in the estimation. Both likelihood functions are multiplied or, equivalently, the log-likelihood functions are summed up.

As both model are anchored on a different scale, an assumption has to be made on how both scale relate to each other. The assumption in the hybrid model is that the coefficients in the DCE model (β_{DCE}) can be rescaled to match the coefficients from the cTTO model (β_{cTTO}). As discussed in section 3.2.3, proportionality between both coefficients is assumed with a single scaling parameter that applies to all coefficients:

The log-likelihood $\ln(L)$ is then defined as follows:

$$\ln(L) = -\frac{1}{2}\sum_i\sum_j\left(\ln\left(2\,\pi\,exp\left(ln(\sigma_\eta)\right)^2\right) + \left(\frac{DU_{ij} - \left(I + \beta_{HYB}X_j + \nu_i\right)}{exp\left(ln(\sigma_\eta)\right)}\right)^2\right)$$

Part of the log-Likelihood function; for **all non-censored cTTO observations**

$$+\sum_i\sum_j\left(ln\left(\Phi\left(\frac{-\left(DU_{ij} - \left(I + \beta_{HYB}X_j + \nu_i\right)\right)}{exp\left(ln(\sigma_\eta)\right)}\right)\right)\right)$$

Part of the log-Likelihood function; for **all censored cTTO observations**, i.e. right censored with respect to disutility, with $\Phi(.)$ the standard normal cumulative distribution function.

$$+\sum_i\sum_p\left(\begin{array}{l}\ln\left(\frac{1}{1+\exp\left(-\frac{\beta_{HYB}}{\theta}(X_{pA}-X_{pB})\right)}\right) * C_{ip} + \\ \ln\left(\frac{\exp\left(-\frac{\beta_{HYB}}{\theta}(X_{pA}-X_{pB})\right)}{1+\exp\left(-\frac{\beta_{HYB}}{\theta}(X_{pA}-X_{pB})\right)}\right) * \left(1-C_{ip}\right)\end{array}\right)$$

Part of the log-Likelihood function; for **all DCE observations**

APPENDIX 2. REPRESENTATIVENESS OF THE POPULATION AT REGIONAL LEVEL

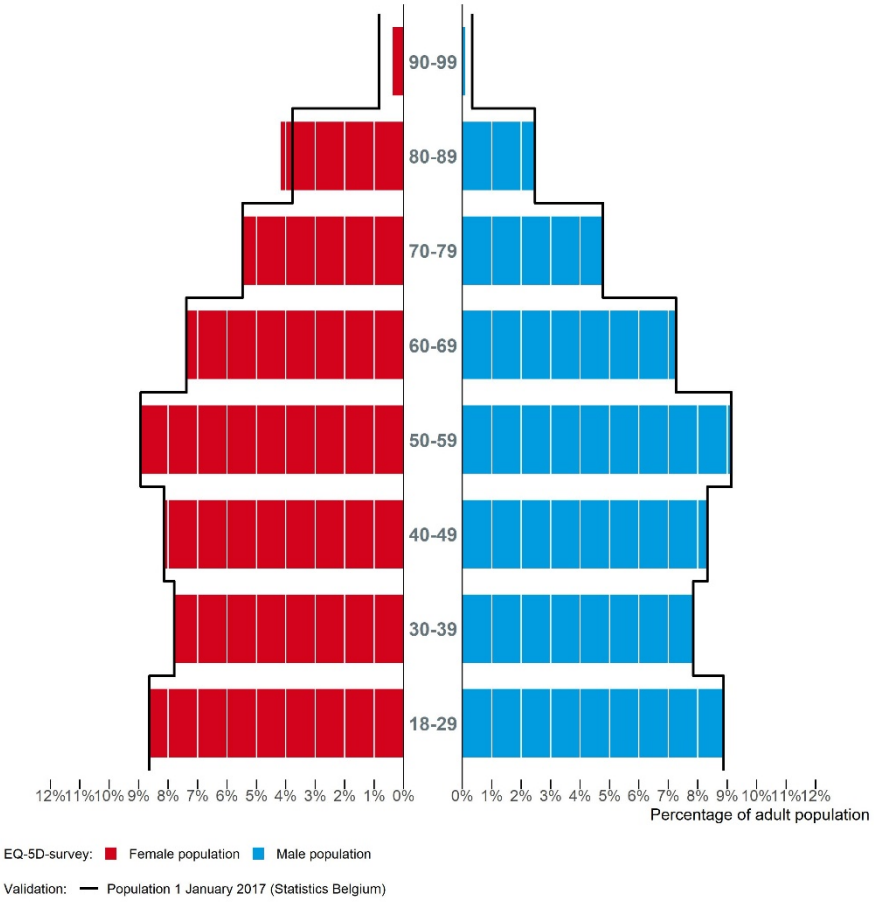
Panel A of Figure 10 presents a population pyramid to compare the age and sex composition of the survey sample with the makeup of the Belgian population on 1 January 2017. In this appendix, a similar comparison is made for the three regions: Flemish region, Walloon region and Brussels region. Figure 21 presents the regional population pyramids. Age refers to the age at time of sampling and not at the time of interview. Regional population statistics were obtained from Statistics Belgium.[71]

It is obvious from Figure 21 that the age profile of the population in Brussels diverges from the other two regions with in particular a higher fraction of adults aged between 18 and 39. In general, there is a very good match at the regional level between the makeup of the survey sample and the population. This was expected as the post-stratification weights not only account for age and sex, but also for province, which is a smaller geographic unit than the region.

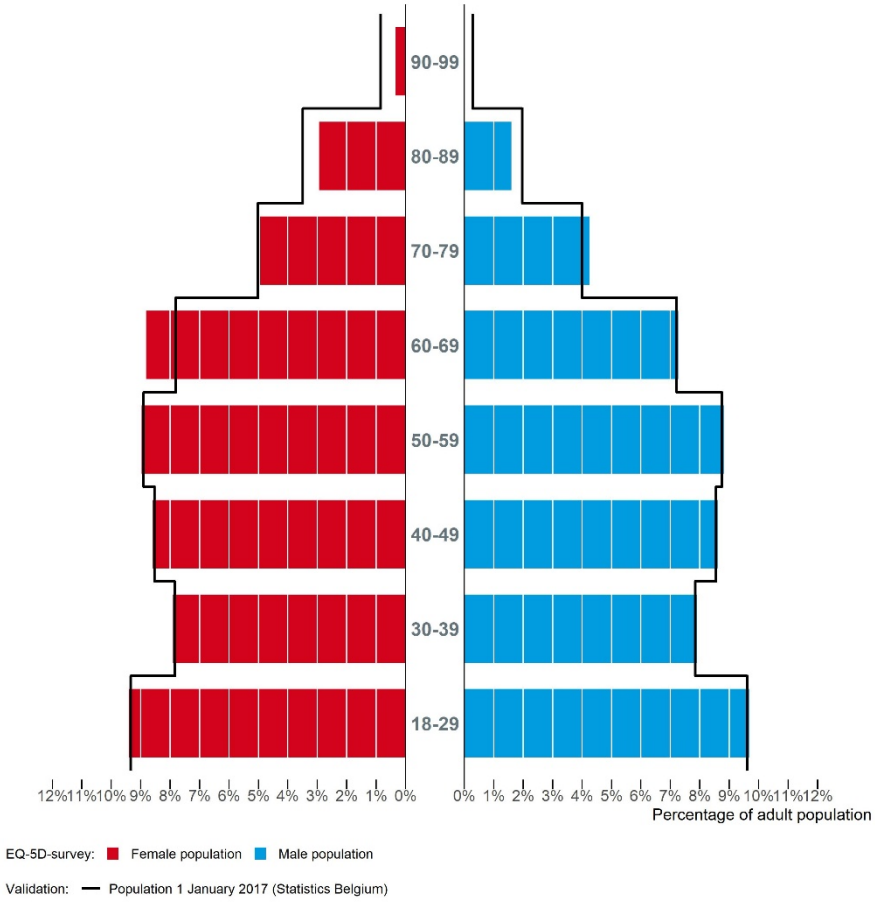
There is only a minor overrepresentation of women aged between 60-69 in the Walloon region and a deficit in the number of respondents aged between 90 and 99 for both sexes in the Flemish and Walloon region.

Figure 21 – Distribution of age and sex in the EQ-5D-5L survey sample relative to the population, by region

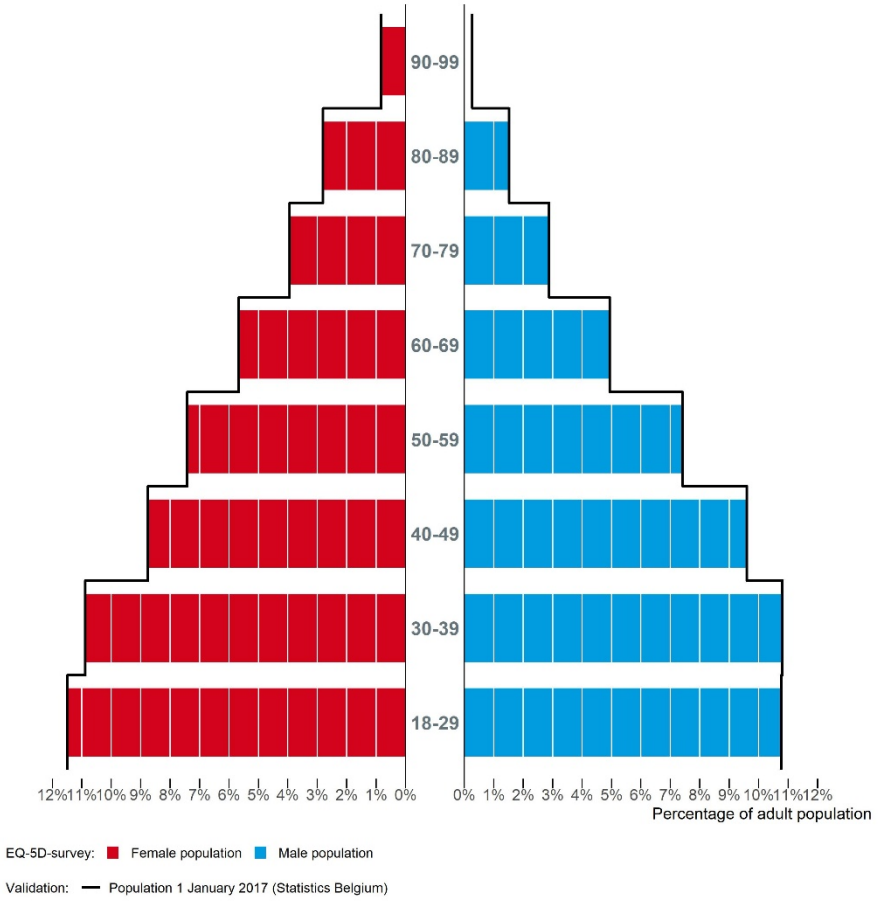
A. Population pyramid Flemish region



B. Population pyramid Walloon region



C. Population pyramid Brussels region



APPENDIX 3. FINAL VALUE SET

Table 11 – EQ-5D-5L value set for Belgium

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
11111	1.000	**11152**	0.419	**11243**	0.444	**11334**	0.445	**11425**	0.314	**11521**	0.714	**12112**	0.878	**12153**	0.344	**12244**	0.189
11112	0.901	**11153**	0.367	**11244**	0.212	**11335**	0.352	**11431**	0.695	**11522**	0.653	**12113**	0.826	**12154**	0.111	**12245**	0.096
11113	0.849	**11154**	0.134	**11245**	0.119	**11341**	0.536	**11432**	0.634	**11523**	0.601	**12114**	0.594	**12155**	0.018	**12251**	0.432
11114	0.617	**11155**	0.041	**11251**	0.455	**11342**	0.475	**11433**	0.582	**11524**	0.369	**12115**	0.501	**12211**	0.914	**12252**	0.371
11115	0.524	**11211**	0.937	**11252**	0.394	**11343**	0.422	**11434**	0.350	**11525**	0.276	**12121**	0.872	**12212**	0.853	**12253**	0.319
11121	0.895	**11212**	0.876	**11253**	0.342	**11344**	0.190	**11435**	0.257	**11531**	0.657	**12122**	0.811	**12213**	0.801	**12254**	0.086
11122	0.834	**11213**	0.824	**11254**	0.109	**11345**	0.097	**11441**	0.440	**11532**	0.596	**12123**	0.759	**12214**	0.569	**12255**	-0.007
11123	0.782	**11214**	0.592	**11255**	0.016	**11351**	0.433	**11442**	0.379	**11533**	0.544	**12124**	0.527	**12215**	0.475	**12311**	0.893
11124	0.550	**11215**	0.498	**11311**	0.916	**11352**	0.372	**11443**	0.327	**11534**	0.311	**12125**	0.434	**12221**	0.847	**12312**	0.832
11125	0.457	**11221**	0.870	**11312**	0.855	**11353**	0.320	**11444**	0.094	**11535**	0.218	**12131**	0.815	**12222**	0.786	**12313**	0.779
11131	0.838	**11222**	0.809	**11313**	0.802	**11354**	0.088	**11445**	0.001	**11541**	0.402	**12132**	0.754	**12223**	0.734	**12314**	0.547
11132	0.777	**11223**	0.757	**11314**	0.570	**11355**	-0.005	**11451**	0.338	**11542**	0.341	**12133**	0.701	**12224**	0.501	**12315**	0.454
11133	0.725	**11224**	0.524	**11315**	0.477	**11411**	0.820	**11452**	0.277	**11543**	0.288	**12134**	0.469	**12225**	0.408	**12321**	0.826
11134	0.492	**11225**	0.431	**11321**	0.849	**11412**	0.759	**11453**	0.224	**11544**	0.056	**12135**	0.376	**12231**	0.790	**12322**	0.765
11135	0.399	**11231**	0.813	**11322**	0.788	**11413**	0.707	**11454**	-0.008	**11545**	-0.037	**12141**	0.559	**12232**	0.729	**12323**	0.712
11141	0.582	**11232**	0.752	**11323**	0.735	**11414**	0.474	**11455**	-0.101	**11551**	0.299	**12142**	0.498	**12233**	0.676	**12324**	0.480
11142	0.521	**11233**	0.699	**11324**	0.503	**11415**	0.381	**11511**	0.782	**11552**	0.238	**12143**	0.446	**12234**	0.444	**12325**	0.387
11143	0.469	**11234**	0.467	**11325**	0.410	**11421**	0.753	**11512**	0.721	**11553**	0.186	**12144**	0.214	**12235**	0.351	**12331**	0.768
11144	0.237	**11235**	0.374	**11331**	0.791	**11422**	0.692	**11513**	0.668	**11554**	-0.046	**12145**	0.121	**12241**	0.534	**12332**	0.707
11145	0.144	**11241**	0.557	**11332**	0.730	**11423**	0.640	**11514**	0.436	**11555**	-0.139	**12151**	0.457	**12242**	0.473	**12333**	0.655
11151	0.480	**11242**	0.496	**11333**	0.678	**11424**	0.407	**11515**	0.343	**12111**	0.939	**12152**	0.396	**12243**	0.421	**12334**	0.422

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
12335	0.329	**12434**	0.327	**12533**	0.521	**13132**	0.734	**13231**	0.770	**13325**	0.367	**13424**	0.364	**13523**	0.558	**14122**	0.704
12341	0.513	**12435**	0.234	**12534**	0.288	**13133**	0.682	**13232**	0.709	**13331**	0.748	**13425**	0.271	**13524**	0.326	**14123**	0.652
12342	0.452	**12441**	0.417	**12535**	0.195	**13134**	0.449	**13233**	0.657	**13332**	0.687	**13431**	0.653	**13525**	0.233	**14124**	0.419
12343	0.399	**12442**	0.356	**12541**	0.379	**13135**	0.356	**13234**	0.424	**13333**	0.635	**13432**	0.592	**13531**	0.614	**14125**	0.326
12344	0.167	**12443**	0.304	**12542**	0.318	**13141**	0.540	**13235**	0.331	**13334**	0.403	**13433**	0.539	**13532**	0.553	**14131**	0.707
12345	0.074	**12444**	0.071	**12543**	0.265	**13142**	0.479	**13241**	0.515	**13335**	0.310	**13434**	0.307	**13533**	0.501	**14132**	0.646
12351	0.410	**12445**	-0.022	**12544**	0.033	**13143**	0.426	**13242**	0.453	**13341**	0.493	**13435**	0.214	**13534**	0.269	**14133**	0.594
12352	0.349	**12451**	0.315	**12545**	-0.060	**13144**	0.194	**13243**	0.401	**13342**	0.432	**13441**	0.397	**13535**	0.176	**14134**	0.362
12353	0.297	**12452**	0.254	**12551**	0.276	**13145**	0.101	**13244**	0.169	**13343**	0.380	**13442**	0.336	**13541**	0.359	**14135**	0.269
12354	0.065	**12453**	0.201	**12552**	0.215	**13151**	0.437	**13245**	0.076	**13344**	0.147	**13443**	0.284	**13542**	0.298	**14141**	0.452
12355	-0.028	**12454**	-0.031	**12553**	0.163	**13152**	0.376	**13251**	0.412	**13345**	0.054	**13444**	0.052	**13543**	0.246	**14142**	0.391
12411	0.797	**12455**	-0.124	**12554**	-0.069	**13153**	0.324	**13252**	0.351	**13351**	0.391	**13445**	-0.041	**13544**	0.013	**14143**	0.339
12412	0.736	**12511**	0.759	**12555**	-0.162	**13154**	0.092	**13253**	0.299	**13352**	0.330	**13451**	0.295	**13545**	-0.080	**14144**	0.106
12413	0.684	**12512**	0.698	**13111**	0.920	**13155**	-0.001	**13254**	0.067	**13353**	0.277	**13452**	0.234	**13551**	0.257	**14145**	0.013
12414	0.451	**12513**	0.645	**13112**	0.859	**13211**	0.894	**13255**	-0.027	**13354**	0.045	**13453**	0.182	**13552**	0.196	**14151**	0.350
12415	0.358	**12514**	0.413	**13113**	0.806	**13212**	0.833	**13311**	0.873	**13355**	-0.048	**13454**	-0.051	**13553**	0.143	**14152**	0.289
12421	0.730	**12515**	0.320	**13114**	0.574	**13213**	0.781	**13312**	0.812	**13411**	0.777	**13455**	-0.144	**13554**	-0.089	**14153**	0.236
12422	0.669	**12521**	0.691	**13115**	0.481	**13214**	0.549	**13313**	0.760	**13412**	0.716	**13511**	0.739	**13555**	-0.182	**14154**	0.004
12423	0.617	**12522**	0.630	**13121**	0.853	**13215**	0.456	**13314**	0.527	**13413**	0.664	**13512**	0.678	**14111**	0.832	**14155**	-0.089
12424	0.384	**12523**	0.578	**13122**	0.791	**13221**	0.827	**13315**	0.434	**13414**	0.432	**13513**	0.626	**14112**	0.771	**14211**	0.807
12425	0.291	**12524**	0.346	**13123**	0.739	**13222**	0.766	**13321**	0.806	**13415**	0.338	**13514**	0.393	**14113**	0.719	**14212**	0.746
12431	0.672	**12525**	0.253	**13124**	0.507	**13223**	0.714	**13322**	0.745	**13421**	0.710	**13515**	0.300	**14114**	0.486	**14213**	0.694
12432	0.611	**12531**	0.634	**13125**	0.414	**13224**	0.482	**13323**	0.693	**13422**	0.649	**13521**	0.672	**14115**	0.393	**14214**	0.461
12433	0.559	**12532**	0.573	**13131**	0.795	**13225**	0.389	**13324**	0.460	**13423**	0.597	**13522**	0.611	**14121**	0.765	**14215**	0.368

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
14221	0.740	**14315**	0.347	**14414**	0.344	**14513**	0.538	**15112**	0.736	**15211**	0.772	**15255**	-0.149	**15354**	-0.078	**15453**	0.059
14222	0.679	**14321**	0.718	**14415**	0.251	**14514**	0.306	**15113**	0.684	**15212**	0.711	**15311**	0.750	**15355**	-0.171	**15454**	-0.173
14223	0.626	**14322**	0.657	**14421**	0.622	**14515**	0.213	**15114**	0.451	**15213**	0.658	**15312**	0.689	**15411**	0.654	**15455**	-0.267
14224	0.394	**14323**	0.605	**14422**	0.561	**14521**	0.584	**15115**	0.358	**15214**	0.426	**15313**	0.637	**15412**	0.593	**15511**	0.616
14225	0.301	**14324**	0.373	**14423**	0.509	**14522**	0.523	**15121**	0.730	**15215**	0.333	**15314**	0.405	**15413**	0.541	**15512**	0.555
14231	0.682	**14325**	0.280	**14424**	0.277	**14523**	0.471	**15122**	0.669	**15221**	0.705	**15315**	0.311	**15414**	0.309	**15513**	0.503
14232	0.621	**14331**	0.661	**14425**	0.184	**14524**	0.239	**15123**	0.617	**15222**	0.644	**15321**	0.683	**15415**	0.216	**15514**	0.270
14233	0.569	**14332**	0.600	**14431**	0.565	**14525**	0.145	**15124**	0.384	**15223**	0.591	**15322**	0.622	**15421**	0.587	**15515**	0.177
14234	0.337	**14333**	0.547	**14432**	0.504	**14531**	0.527	**15125**	0.291	**15224**	0.359	**15323**	0.570	**15422**	0.526	**15521**	0.549
14235	0.244	**14334**	0.315	**14433**	0.452	**14532**	0.466	**15131**	0.672	**15225**	0.266	**15324**	0.337	**15423**	0.474	**15522**	0.488
14241	0.427	**14335**	0.222	**14434**	0.219	**14533**	0.413	**15132**	0.611	**15231**	0.647	**15325**	0.244	**15424**	0.242	**15523**	0.436
14242	0.366	**14341**	0.405	**14435**	0.126	**14534**	0.181	**15133**	0.559	**15232**	0.586	**15331**	0.626	**15425**	0.149	**15524**	0.203
14243	0.314	**14342**	0.344	**14441**	0.310	**14535**	0.088	**15134**	0.327	**15233**	0.534	**15332**	0.565	**15431**	0.530	**15525**	0.110
14244	0.081	**14343**	0.292	**14442**	0.249	**14541**	0.271	**15135**	0.234	**15234**	0.302	**15333**	0.512	**15432**	0.469	**15531**	0.492
14245	-0.012	**14344**	0.060	**14443**	0.196	**14542**	0.210	**15141**	0.417	**15235**	0.208	**15334**	0.280	**15433**	0.417	**15532**	0.431
14251	0.325	**14345**	-0.033	**14444**	-0.036	**14543**	0.158	**15142**	0.356	**15241**	0.392	**15335**	0.187	**15434**	0.184	**15533**	0.378
14252	0.264	**14351**	0.303	**14445**	-0.129	**14544**	-0.074	**15143**	0.304	**15242**	0.331	**15341**	0.370	**15435**	0.091	**15534**	0.146
14253	0.211	**14352**	0.242	**14451**	0.207	**14545**	-0.167	**15144**	0.071	**15243**	0.278	**15342**	0.309	**15441**	0.275	**15535**	0.053
14254	-0.021	**14353**	0.190	**14452**	0.146	**14551**	0.169	**15145**	-0.022	**15244**	0.046	**15343**	0.257	**15442**	0.213	**15541**	0.236
14255	-0.114	**14354**	-0.043	**14453**	0.094	**14552**	0.108	**15151**	0.315	**15245**	-0.047	**15344**	0.025	**15443**	0.161	**15542**	0.175
14311	0.785	**14355**	-0.136	**14454**	-0.138	**14553**	0.056	**15152**	0.254	**15251**	0.289	**15345**	-0.068	**15444**	-0.071	**15543**	0.123
14312	0.724	**14411**	0.690	**14455**	-0.231	**14554**	-0.177	**15153**	0.201	**15252**	0.228	**15351**	0.268	**15445**	-0.164	**15544**	-0.109
14313	0.672	**14412**	0.629	**14511**	0.651	**14555**	-0.270	**15154**	-0.031	**15253**	0.176	**15352**	0.207	**15451**	0.172	**15545**	-0.203
14314	0.440	**14413**	0.576	**14512**	0.590	**15111**	0.797	**15155**	-0.124	**15254**	-0.056	**15353**	0.155	**15452**	0.111	**15551**	0.134

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
15552	0.073	**21151**	0.449	**21245**	0.087	**21344**	0.159	**21443**	0.295	**21542**	0.309	**22141**	0.528	**22235**	0.319	**22334**	0.391
15553	0.021	**21152**	0.387	**21251**	0.423	**21345**	0.065	**21444**	0.063	**21543**	0.257	**22142**	0.467	**22241**	0.503	**22335**	0.298
15554	-0.212	**21153**	0.335	**21252**	0.362	**21351**	0.402	**21445**	-0.030	**21544**	0.024	**22143**	0.415	**22242**	0.442	**22341**	0.481
15555	-0.305	**21154**	0.103	**21253**	0.310	**21352**	0.341	**21451**	0.306	**21545**	-0.069	**22144**	0.182	**22243**	0.389	**22342**	0.420
21111	0.931	**21155**	0.010	**21254**	0.078	**21353**	0.289	**21452**	0.245	**21551**	0.268	**22145**	0.089	**22244**	0.157	**22343**	0.368
21112	0.870	**21211**	0.906	**21255**	-0.015	**21354**	0.056	**21453**	0.193	**21552**	0.207	**22151**	0.426	**22245**	0.064	**22344**	0.136
21113	0.817	**21212**	0.845	**21311**	0.884	**21355**	-0.037	**21454**	-0.040	**21553**	0.154	**22152**	0.364	**22251**	0.400	**22345**	0.042
21114	0.585	**21213**	0.792	**21312**	0.823	**21411**	0.788	**21455**	-0.133	**21554**	-0.078	**22153**	0.312	**22252**	0.339	**22351**	0.379
21115	0.492	**21214**	0.560	**21313**	0.771	**21412**	0.727	**21511**	0.750	**21555**	-0.171	**22154**	0.080	**22253**	0.287	**22352**	0.318
21121	0.864	**21215**	0.467	**21314**	0.538	**21413**	0.675	**21512**	0.689	**22111**	0.908	**22155**	-0.013	**22254**	0.055	**22353**	0.266
21122	0.803	**21221**	0.839	**21315**	0.445	**21414**	0.443	**21513**	0.637	**22112**	0.847	**22211**	0.883	**22255**	-0.038	**22354**	0.033
21123	0.750	**21222**	0.778	**21321**	0.817	**21415**	0.350	**21514**	0.404	**22113**	0.794	**22212**	0.822	**22311**	0.861	**22355**	-0.060
21124	0.518	**21223**	0.725	**21322**	0.756	**21421**	0.721	**21515**	0.311	**22114**	0.562	**22213**	0.769	**22312**	0.800	**22411**	0.765
21125	0.425	**21224**	0.493	**21323**	0.704	**21422**	0.660	**21521**	0.683	**22115**	0.469	**22214**	0.537	**22313**	0.748	**22412**	0.704
21131	0.806	**21225**	0.400	**21324**	0.471	**21423**	0.608	**21522**	0.622	**22121**	0.841	**22215**	0.444	**22314**	0.515	**22413**	0.652
21132	0.745	**21231**	0.781	**21325**	0.378	**21424**	0.376	**21523**	0.570	**22122**	0.780	**22221**	0.816	**22315**	0.422	**22414**	0.420
21133	0.693	**21232**	0.720	**21331**	0.760	**21425**	0.283	**21524**	0.337	**22123**	0.727	**22222**	0.755	**22321**	0.794	**22415**	0.327
21134	0.461	**21233**	0.668	**21332**	0.699	**21431**	0.664	**21525**	0.244	**22124**	0.495	**22223**	0.702	**22322**	0.733	**22421**	0.698
21135	0.368	**21234**	0.435	**21333**	0.646	**21432**	0.603	**21531**	0.625	**22125**	0.402	**22224**	0.470	**22323**	0.681	**22422**	0.637
21141	0.551	**21235**	0.342	**21334**	0.414	**21433**	0.551	**21532**	0.564	**22131**	0.783	**22225**	0.377	**22324**	0.448	**22423**	0.585
21142	0.490	**21241**	0.526	**21335**	0.321	**21434**	0.318	**21533**	0.512	**22132**	0.722	**22231**	0.758	**22325**	0.355	**22424**	0.353
21143	0.438	**21242**	0.465	**21341**	0.504	**21435**	0.225	**21534**	0.280	**22133**	0.670	**22232**	0.697	**22331**	0.737	**22425**	0.260
21144	0.205	**21243**	0.412	**21342**	0.443	**21441**	0.408	**21535**	0.187	**22134**	0.438	**22233**	0.645	**22332**	0.676	**22431**	0.641
21145	0.112	**21244**	0.180	**21343**	0.391	**21442**	0.347	**21541**	0.370	**22135**	0.345	**22234**	0.412	**22333**	0.623	**22432**	0.580

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
22433	0.528	**22532**	0.541	**23131**	0.764	**23225**	0.357	**23324**	0.429	**23423**	0.565	**23522**	0.579	**24121**	0.733	**24215**	0.337
22434	0.295	**22533**	0.489	**23132**	0.703	**23231**	0.738	**23325**	0.336	**23424**	0.333	**23523**	0.527	**24122**	0.672	**24221**	0.708
22435	0.202	**22534**	0.257	**23133**	0.650	**23232**	0.677	**23331**	0.717	**23425**	0.240	**23524**	0.295	**24123**	0.620	**24222**	0.647
22441	0.385	**22535**	0.164	**23134**	0.418	**23233**	0.625	**23332**	0.656	**23431**	0.621	**23525**	0.202	**24124**	0.388	**24223**	0.595
22442	0.324	**22541**	0.347	**23135**	0.325	**23234**	0.393	**23333**	0.604	**23432**	0.560	**23531**	0.583	**24125**	0.295	**24224**	0.363
22443	0.272	**22542**	0.286	**23141**	0.508	**23235**	0.300	**23334**	0.371	**23433**	0.508	**23532**	0.522	**24131**	0.676	**24225**	0.269
22444	0.040	**22543**	0.234	**23142**	0.447	**23241**	0.483	**23335**	0.278	**23434**	0.275	**23533**	0.469	**24132**	0.615	**24231**	0.651
22445	-0.053	**22544**	0.001	**23143**	0.395	**23242**	0.422	**23341**	0.461	**23435**	0.182	**23534**	0.237	**24133**	0.563	**24232**	0.590
22451	0.283	**22545**	-0.092	**23144**	0.162	**23243**	0.370	**23342**	0.400	**23441**	0.366	**23535**	0.144	**24134**	0.330	**24233**	0.537
22452	0.222	**22551**	0.245	**23145**	0.069	**23244**	0.137	**23343**	0.348	**23442**	0.305	**23541**	0.327	**24135**	0.237	**24234**	0.305
22453	0.170	**22552**	0.184	**23151**	0.406	**23245**	0.044	**23344**	0.116	**23443**	0.252	**23542**	0.266	**24141**	0.420	**24235**	0.212
22454	-0.063	**22553**	0.131	**23152**	0.345	**23251**	0.381	**23345**	0.023	**23444**	0.020	**23543**	0.214	**24142**	0.359	**24241**	0.395
22455	-0.156	**22554**	-0.101	**23153**	0.293	**23252**	0.320	**23351**	0.359	**23445**	-0.073	**23544**	-0.018	**24143**	0.307	**24242**	0.334
22511	0.727	**22555**	-0.194	**23154**	0.060	**23253**	0.267	**23352**	0.298	**23451**	0.263	**23545**	-0.111	**24144**	0.075	**24243**	0.282
22512	0.666	**23111**	0.888	**23155**	-0.033	**23254**	0.035	**23353**	0.246	**23452**	0.202	**23551**	0.225	**24145**	-0.018	**24244**	0.050
22513	0.614	**23112**	0.827	**23211**	0.863	**23255**	-0.058	**23354**	0.013	**23453**	0.150	**23552**	0.164	**24151**	0.318	**24245**	-0.043
22514	0.381	**23113**	0.775	**23212**	0.802	**23311**	0.841	**23355**	-0.080	**23454**	-0.082	**23553**	0.112	**24152**	0.257	**24251**	0.293
22515	0.288	**23114**	0.542	**23213**	0.750	**23312**	0.780	**23411**	0.746	**23455**	-0.175	**23554**	-0.121	**24153**	0.205	**24252**	0.232
22521	0.660	**23115**	0.449	**23214**	0.517	**23313**	0.728	**23412**	0.685	**23511**	0.707	**23555**	-0.214	**24154**	-0.027	**24253**	0.180
22522	0.599	**23121**	0.821	**23215**	0.424	**23314**	0.496	**23413**	0.632	**23512**	0.646	**24111**	0.800	**24155**	-0.121	**24254**	-0.053
22523	0.547	**23122**	0.760	**23221**	0.796	**23315**	0.403	**23414**	0.400	**23513**	0.594	**24112**	0.739	**24211**	0.775	**24255**	-0.146
22524	0.314	**23123**	0.708	**23222**	0.735	**23321**	0.774	**23415**	0.307	**23514**	0.362	**24113**	0.687	**24212**	0.714	**24311**	0.754
22525	0.221	**23124**	0.475	**23223**	0.683	**23322**	0.713	**23421**	0.679	**23515**	0.269	**24114**	0.455	**24213**	0.662	**24312**	0.693
22531	0.602	**23125**	0.382	**23224**	0.450	**23323**	0.661	**23422**	0.618	**23521**	0.640	**24115**	0.362	**24214**	0.430	**24313**	0.640

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
24314	0.408	**24413**	0.545	**24512**	0.559	**25111**	0.765	**25155**	-0.156	**25254**	-0.088	**25353**	0.123	**25452**	0.080	**25551**	0.102
24315	0.315	**24414**	0.312	**24513**	0.506	**25112**	0.704	**25211**	0.740	**25255**	-0.181	**25354**	-0.109	**25453**	0.027	**25552**	0.041
24321	0.687	**24415**	0.219	**24514**	0.274	**25113**	0.652	**25212**	0.679	**25311**	0.719	**25355**	-0.202	**25454**	-0.205	**25553**	-0.011
24322	0.626	**24421**	0.591	**24515**	0.181	**25114**	0.420	**25213**	0.627	**25312**	0.658	**25411**	0.623	**25455**	-0.298	**25554**	-0.243
24323	0.573	**24422**	0.530	**24521**	0.553	**25115**	0.327	**25214**	0.395	**25313**	0.605	**25412**	0.562	**25511**	0.585	**25555**	-0.336
24324	0.341	**24423**	0.478	**24522**	0.492	**25121**	0.698	**25215**	0.301	**25314**	0.373	**25413**	0.510	**25512**	0.524	**31111**	0.904
24325	0.248	**24424**	0.245	**24523**	0.439	**25122**	0.637	**25221**	0.673	**25315**	0.280	**25414**	0.277	**25513**	0.471	**31112**	0.843
24331	0.629	**24425**	0.152	**24524**	0.207	**25123**	0.585	**25222**	0.612	**25321**	0.652	**25415**	0.184	**25514**	0.239	**31113**	0.790
24332	0.568	**24431**	0.533	**24525**	0.114	**25124**	0.353	**25223**	0.560	**25322**	0.591	**25421**	0.556	**25515**	0.146	**31114**	0.558
24333	0.516	**24432**	0.472	**24531**	0.495	**25125**	0.260	**25224**	0.327	**25323**	0.538	**25422**	0.495	**25521**	0.517	**31115**	0.465
24334	0.284	**24433**	0.420	**24532**	0.434	**25131**	0.641	**25225**	0.234	**25324**	0.306	**25423**	0.443	**25522**	0.456	**31121**	0.837
24335	0.191	**24434**	0.188	**24533**	0.382	**25132**	0.580	**25231**	0.616	**25325**	0.213	**25424**	0.210	**25523**	0.404	**31122**	0.776
24341	0.374	**24435**	0.095	**24534**	0.150	**25133**	0.528	**25232**	0.555	**25331**	0.594	**25425**	0.117	**25524**	0.172	**31123**	0.723
24342	0.313	**24441**	0.278	**24535**	0.056	**25134**	0.295	**25233**	0.502	**25332**	0.533	**25431**	0.498	**25525**	0.079	**31124**	0.491
24343	0.261	**24442**	0.217	**24541**	0.240	**25135**	0.202	**25234**	0.270	**25333**	0.481	**25432**	0.437	**25531**	0.460	**31125**	0.398
24344	0.028	**24443**	0.165	**24542**	0.179	**25141**	0.385	**25235**	0.177	**25334**	0.248	**25433**	0.385	**25532**	0.399	**31131**	0.779
24345	-0.065	**24444**	-0.068	**24543**	0.126	**25142**	0.324	**25241**	0.360	**25335**	0.155	**25434**	0.153	**25533**	0.347	**31132**	0.718
24351	0.271	**24445**	-0.161	**24544**	-0.106	**25143**	0.272	**25242**	0.299	**25341**	0.339	**25435**	0.060	**25534**	0.114	**31133**	0.666
24352	0.210	**24451**	0.176	**24545**	-0.199	**25144**	0.040	**25243**	0.247	**25342**	0.278	**25441**	0.243	**25535**	0.021	**31134**	0.434
24353	0.158	**24452**	0.115	**24551**	0.137	**25145**	-0.053	**25244**	0.015	**25343**	0.225	**25442**	0.182	**25541**	0.205	**31135**	0.341
24354	-0.074	**24453**	0.062	**24552**	0.076	**25151**	0.283	**25245**	-0.078	**25344**	-0.007	**25443**	0.130	**25542**	0.144	**31141**	0.524
24355	-0.167	**24454**	-0.170	**24553**	0.024	**25152**	0.222	**25251**	0.258	**25345**	-0.100	**25444**	-0.103	**25543**	0.091	**31142**	0.463
24411	0.658	**24455**	-0.263	**24554**	-0.208	**25153**	0.170	**25252**	0.197	**25351**	0.236	**25445**	-0.196	**25544**	-0.141	**31143**	0.411
24412	0.597	**24511**	0.620	**24555**	-0.301	**25154**	-0.063	**25253**	0.145	**25352**	0.175	**25451**	0.141	**25545**	-0.234	**31144**	0.178

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
31145	0.085	**31244**	0.153	**31343**	0.364	**31442**	0.320	**31541**	0.343	**32135**	0.318	**32234**	0.385	**32333**	0.596	**32432**	0.553
31151	0.421	**31245**	0.060	**31344**	0.132	**31443**	0.268	**31542**	0.282	**32141**	0.501	**32235**	0.292	**32334**	0.364	**32433**	0.501
31152	0.360	**31251**	0.396	**31345**	0.038	**31444**	0.036	**31543**	0.230	**32142**	0.440	**32241**	0.476	**32335**	0.271	**32434**	0.268
31153	0.308	**31252**	0.335	**31351**	0.375	**31445**	-0.057	**31544**	-0.003	**32143**	0.388	**32242**	0.415	**32341**	0.454	**32435**	0.175
31154	0.076	**31253**	0.283	**31352**	0.314	**31451**	0.279	**31545**	-0.096	**32144**	0.155	**32243**	0.362	**32342**	0.393	**32441**	0.358
31155	-0.017	**31254**	0.051	**31353**	0.262	**31452**	0.218	**31551**	0.241	**32145**	0.062	**32244**	0.130	**32343**	0.341	**32442**	0.297
31211	0.879	**31255**	-0.042	**31354**	0.029	**31453**	0.166	**31552**	0.180	**32151**	0.398	**32245**	0.037	**32344**	0.109	**32443**	0.245
31212	0.818	**31311**	0.857	**31355**	-0.064	**31454**	-0.067	**31553**	0.127	**32152**	0.337	**32251**	0.373	**32345**	0.015	**32444**	0.013
31213	0.765	**31312**	0.796	**31411**	0.761	**31455**	-0.160	**31554**	-0.105	**32153**	0.285	**32252**	0.312	**32351**	0.352	**32445**	-0.080
31214	0.533	**31313**	0.744	**31412**	0.700	**31511**	0.723	**31555**	-0.198	**32154**	0.053	**32253**	0.260	**32352**	0.291	**32451**	0.256
31215	0.440	**31314**	0.511	**31413**	0.648	**31512**	0.662	**32111**	0.881	**32155**	-0.040	**32254**	0.028	**32353**	0.239	**32452**	0.195
31221	0.812	**31315**	0.418	**31414**	0.416	**31513**	0.610	**32112**	0.820	**32211**	0.856	**32255**	-0.065	**32354**	0.006	**32453**	0.143
31222	0.751	**31321**	0.790	**31415**	0.323	**31514**	0.377	**32113**	0.767	**32212**	0.795	**32311**	0.834	**32355**	-0.087	**32454**	-0.090
31223	0.698	**31322**	0.729	**31421**	0.694	**31515**	0.284	**32114**	0.535	**32213**	0.742	**32312**	0.773	**32411**	0.738	**32455**	-0.183
31224	0.466	**31323**	0.677	**31422**	0.633	**31521**	0.656	**32115**	0.442	**32214**	0.510	**32313**	0.721	**32412**	0.677	**32511**	0.700
31225	0.373	**31324**	0.444	**31423**	0.581	**31522**	0.595	**32121**	0.814	**32215**	0.417	**32314**	0.488	**32413**	0.625	**32512**	0.639
31231	0.754	**31325**	0.351	**31424**	0.349	**31523**	0.543	**32122**	0.753	**32221**	0.789	**32315**	0.395	**32414**	0.393	**32513**	0.587
31232	0.693	**31331**	0.733	**31425**	0.256	**31524**	0.310	**32123**	0.700	**32222**	0.728	**32321**	0.767	**32415**	0.300	**32514**	0.354
31233	0.641	**31332**	0.672	**31431**	0.637	**31525**	0.217	**32124**	0.468	**32223**	0.675	**32322**	0.706	**32421**	0.671	**32515**	0.261
31234	0.408	**31333**	0.619	**31432**	0.576	**31531**	0.598	**32125**	0.375	**32224**	0.443	**32323**	0.654	**32422**	0.610	**32521**	0.633
31235	0.315	**31334**	0.387	**31433**	0.524	**31532**	0.537	**32131**	0.756	**32225**	0.350	**32324**	0.421	**32423**	0.558	**32522**	0.572
31241	0.499	**31335**	0.294	**31434**	0.291	**31533**	0.485	**32132**	0.695	**32231**	0.731	**32325**	0.328	**32424**	0.326	**32523**	0.520
31242	0.438	**31341**	0.477	**31435**	0.198	**31534**	0.253	**32133**	0.643	**32232**	0.670	**32331**	0.710	**32425**	0.233	**32524**	0.287
31243	0.385	**31342**	0.416	**31441**	0.381	**31535**	0.160	**32134**	0.411	**32233**	0.618	**32332**	0.649	**32431**	0.614	**32525**	0.194

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
32531	0.575	**33125**	0.355	**33224**	0.423	**33323**	0.634	**33422**	0.591	**33521**	0.613	**34115**	0.335	**34214**	0.403	**34313**	0.613
32532	0.514	**33131**	0.737	**33225**	0.330	**33324**	0.402	**33423**	0.538	**33522**	0.552	**34121**	0.706	**34215**	0.310	**34314**	0.381
32533	0.462	**33132**	0.675	**33231**	0.711	**33325**	0.309	**33424**	0.306	**33523**	0.500	**34122**	0.645	**34221**	0.681	**34315**	0.288
32534	0.230	**33133**	0.623	**33232**	0.650	**33331**	0.690	**33425**	0.213	**33524**	0.268	**34123**	0.593	**34222**	0.620	**34321**	0.660
32535	0.137	**33134**	0.391	**33233**	0.598	**33332**	0.629	**33431**	0.594	**33525**	0.175	**34124**	0.361	**34223**	0.568	**34322**	0.599
32541	0.320	**33135**	0.298	**33234**	0.366	**33333**	0.577	**33432**	0.533	**33531**	0.556	**34125**	0.268	**34224**	0.336	**34323**	0.546
32542	0.259	**33141**	0.481	**33235**	0.273	**33334**	0.344	**33433**	0.481	**33532**	0.495	**34131**	0.649	**34225**	0.242	**34324**	0.314
32543	0.207	**33142**	0.420	**33241**	0.456	**33335**	0.251	**33434**	0.248	**33533**	0.442	**34132**	0.588	**34231**	0.624	**34325**	0.221
32544	-0.026	**33143**	0.368	**33242**	0.395	**33341**	0.434	**33435**	0.155	**33534**	0.210	**34133**	0.536	**34232**	0.563	**34331**	0.602
32545	-0.119	**33144**	0.135	**33243**	0.343	**33342**	0.373	**33441**	0.339	**33535**	0.117	**34134**	0.303	**34233**	0.510	**34332**	0.541
32551	0.218	**33145**	0.042	**33244**	0.110	**33343**	0.321	**33442**	0.278	**33541**	0.300	**34135**	0.210	**34234**	0.278	**34333**	0.489
32552	0.157	**33151**	0.379	**33245**	0.017	**33344**	0.089	**33443**	0.225	**33542**	0.239	**34141**	0.393	**34235**	0.185	**34334**	0.257
32553	0.104	**33152**	0.318	**33251**	0.354	**33345**	-0.004	**33444**	-0.007	**33543**	0.187	**34142**	0.332	**34241**	0.368	**34335**	0.163
32554	-0.128	**33153**	0.265	**33252**	0.293	**33351**	0.332	**33445**	-0.100	**33544**	-0.045	**34143**	0.280	**34242**	0.307	**34341**	0.347
32555	-0.221	**33154**	0.033	**33253**	0.240	**33352**	0.271	**33451**	0.236	**33545**	-0.138	**34144**	0.048	**34243**	0.255	**34342**	0.286
33111	0.861	**33155**	-0.060	**33254**	0.008	**33353**	0.219	**33452**	0.175	**33551**	0.198	**34145**	-0.045	**34244**	0.023	**34343**	0.234
33112	0.800	**33211**	0.836	**33255**	-0.085	**33354**	-0.014	**33453**	0.123	**33552**	0.137	**34151**	0.291	**34245**	-0.070	**34344**	0.001
33113	0.748	**33212**	0.775	**33311**	0.814	**33355**	-0.107	**33454**	-0.109	**33553**	0.085	**34152**	0.230	**34251**	0.266	**34345**	-0.092
33114	0.515	**33213**	0.723	**33312**	0.753	**33411**	0.719	**33455**	-0.202	**33554**	-0.148	**34153**	0.178	**34252**	0.205	**34351**	0.244
33115	0.422	**33214**	0.490	**33313**	0.701	**33412**	0.658	**33511**	0.680	**33555**	-0.241	**34154**	-0.054	**34253**	0.153	**34352**	0.183
33121	0.794	**33215**	0.397	**33314**	0.469	**33413**	0.605	**33512**	0.619	**34111**	0.773	**34155**	-0.148	**34254**	-0.080	**34353**	0.131
33122	0.733	**33221**	0.769	**33315**	0.376	**33414**	0.373	**33513**	0.567	**34112**	0.712	**34211**	0.748	**34255**	-0.173	**34354**	-0.101
33123	0.681	**33222**	0.708	**33321**	0.747	**33415**	0.280	**33514**	0.335	**34113**	0.660	**34212**	0.687	**34311**	0.727	**34355**	-0.194
33124	0.448	**33223**	0.656	**33322**	0.686	**33421**	0.652	**33515**	0.242	**34114**	0.428	**34213**	0.635	**34312**	0.666	**34411**	0.631

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
34412	0.570	**34511**	0.593	**34555**	-0.328	**35154**	-0.090	**35253**	0.118	**35352**	0.148	**35451**	0.114	**35545**	-0.261	**41144**	0.058
34413	0.518	**34512**	0.532	**35111**	0.738	**35155**	-0.183	**35254**	-0.115	**35353**	0.096	**35452**	0.053	**35551**	0.075	**41145**	-0.035
34414	0.285	**34513**	0.479	**35112**	0.677	**35211**	0.713	**35255**	-0.208	**35354**	-0.136	**35453**	0.000	**35552**	0.014	**41151**	0.301
34415	0.192	**34514**	0.247	**35113**	0.625	**35212**	0.652	**35311**	0.692	**35355**	-0.229	**35454**	-0.232	**35553**	-0.038	**41152**	0.240
34421	0.564	**34515**	0.154	**35114**	0.393	**35213**	0.600	**35312**	0.631	**35411**	0.596	**35455**	-0.325	**35554**	-0.270	**41153**	0.188
34422	0.503	**34521**	0.526	**35115**	0.300	**35214**	0.368	**35313**	0.578	**35412**	0.535	**35511**	0.558	**35555**	-0.363	**41154**	-0.044
34423	0.451	**34522**	0.465	**35121**	0.671	**35215**	0.274	**35314**	0.346	**35413**	0.483	**35512**	0.497	**41111**	0.784	**41155**	-0.137
34424	0.218	**34523**	0.412	**35122**	0.610	**35221**	0.646	**35315**	0.253	**35414**	0.250	**35513**	0.444	**41112**	0.723	**41211**	0.759
34425	0.125	**34524**	0.180	**35123**	0.558	**35222**	0.585	**35321**	0.625	**35415**	0.157	**35514**	0.212	**41113**	0.670	**41212**	0.697
34431	0.506	**34525**	0.087	**35124**	0.326	**35223**	0.533	**35322**	0.564	**35421**	0.529	**35515**	0.119	**41114**	0.438	**41213**	0.645
34432	0.445	**34531**	0.468	**35125**	0.233	**35224**	0.300	**35323**	0.511	**35422**	0.468	**35521**	0.490	**41115**	0.345	**41214**	0.413
34433	0.393	**34532**	0.407	**35131**	0.614	**35225**	0.207	**35324**	0.279	**35423**	0.416	**35522**	0.429	**41121**	0.717	**41215**	0.320
34434	0.161	**34533**	0.355	**35132**	0.553	**35231**	0.589	**35325**	0.186	**35424**	0.183	**35523**	0.377	**41122**	0.656	**41221**	0.691
34435	0.068	**34534**	0.122	**35133**	0.500	**35232**	0.528	**35331**	0.567	**35425**	0.090	**35524**	0.145	**41123**	0.603	**41222**	0.630
34441	0.251	**34535**	0.029	**35134**	0.268	**35233**	0.475	**35332**	0.506	**35431**	0.471	**35525**	0.052	**41124**	0.371	**41223**	0.578
34442	0.190	**34541**	0.213	**35135**	0.175	**35234**	0.243	**35333**	0.454	**35432**	0.410	**35531**	0.433	**41125**	0.278	**41224**	0.346
34443	0.138	**34542**	0.152	**35141**	0.358	**35235**	0.150	**35334**	0.221	**35433**	0.358	**35532**	0.372	**41131**	0.659	**41225**	0.253
34444	-0.095	**34543**	0.099	**35142**	0.297	**35241**	0.333	**35335**	0.128	**35434**	0.126	**35533**	0.320	**41132**	0.598	**41231**	0.634
34445	-0.188	**34544**	-0.133	**35143**	0.245	**35242**	0.272	**35341**	0.312	**35435**	0.033	**35534**	0.087	**41133**	0.546	**41232**	0.573
34451	0.149	**34545**	-0.226	**35144**	0.013	**35243**	0.220	**35342**	0.251	**35441**	0.216	**35535**	-0.006	**41134**	0.313	**41233**	0.521
34452	0.088	**34551**	0.110	**35145**	-0.080	**35244**	-0.012	**35343**	0.198	**35442**	0.155	**35541**	0.178	**41135**	0.220	**41234**	0.288
34453	0.035	**34552**	0.049	**35151**	0.256	**35245**	-0.105	**35344**	-0.034	**35443**	0.103	**35542**	0.117	**41141**	0.404	**41235**	0.195
34454	-0.197	**34553**	-0.003	**35152**	0.195	**35251**	0.231	**35345**	-0.127	**35444**	-0.130	**35543**	0.064	**41142**	0.343	**41241**	0.379
34455	-0.290	**34554**	-0.235	**35153**	0.143	**35252**	0.170	**35351**	0.209	**35445**	-0.223	**35544**	-0.168	**41143**	0.290	**41242**	0.318

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
41243	0.265	**41342**	0.296	**41441**	0.261	**41535**	0.040	**42134**	0.290	**42233**	0.498	**42332**	0.528	**42431**	0.494	**42525**	0.074
41244	0.033	**41343**	0.244	**41442**	0.200	**41541**	0.223	**42135**	0.197	**42234**	0.265	**42333**	0.476	**42432**	0.433	**42531**	0.455
41245	-0.060	**41344**	0.011	**41443**	0.148	**41542**	0.162	**42141**	0.381	**42235**	0.172	**42334**	0.244	**42433**	0.380	**42532**	0.394
41251	0.276	**41345**	-0.082	**41444**	-0.084	**41543**	0.110	**42142**	0.320	**42241**	0.356	**42335**	0.151	**42434**	0.148	**42533**	0.342
41252	0.215	**41351**	0.255	**41445**	-0.177	**41544**	-0.123	**42143**	0.267	**42242**	0.295	**42341**	0.334	**42435**	0.055	**42534**	0.110
41253	0.163	**41352**	0.194	**41451**	0.159	**41545**	-0.216	**42144**	0.035	**42243**	0.242	**42342**	0.273	**42441**	0.238	**42535**	0.017
41254	-0.069	**41353**	0.141	**41452**	0.098	**41551**	0.121	**42145**	-0.058	**42244**	0.010	**42343**	0.221	**42442**	0.177	**42541**	0.200
41255	-0.162	**41354**	-0.091	**41453**	0.046	**41552**	0.060	**42151**	0.278	**42245**	-0.083	**42344**	-0.012	**42443**	0.125	**42542**	0.139
41311	0.737	**41355**	-0.184	**41454**	-0.187	**41553**	0.007	**42152**	0.217	**42251**	0.253	**42345**	-0.105	**42444**	-0.107	**42543**	0.087
41312	0.676	**41411**	0.641	**41455**	-0.280	**41554**	-0.225	**42153**	0.165	**42252**	0.192	**42351**	0.232	**42445**	-0.200	**42544**	-0.146
41313	0.624	**41412**	0.580	**41511**	0.603	**41555**	-0.318	**42154**	-0.067	**42253**	0.140	**42352**	0.171	**42451**	0.136	**42545**	-0.239
41314	0.391	**41413**	0.528	**41512**	0.542	**42111**	0.761	**42155**	-0.160	**42254**	-0.092	**42353**	0.118	**42452**	0.075	**42551**	0.098
41315	0.298	**41414**	0.296	**41513**	0.490	**42112**	0.700	**42211**	0.735	**42255**	-0.185	**42354**	-0.114	**42453**	0.023	**42552**	0.037
41321	0.670	**41415**	0.203	**41514**	0.257	**42113**	0.647	**42212**	0.674	**42311**	0.714	**42355**	-0.207	**42454**	-0.210	**42553**	-0.016
41322	0.609	**41421**	0.574	**41515**	0.164	**42114**	0.415	**42213**	0.622	**42312**	0.653	**42411**	0.618	**42455**	-0.303	**42554**	-0.248
41323	0.557	**41422**	0.513	**41521**	0.536	**42115**	0.322	**42214**	0.390	**42313**	0.601	**42412**	0.557	**42511**	0.580	**42555**	-0.341
41324	0.324	**41423**	0.461	**41522**	0.475	**42121**	0.694	**42215**	0.297	**42314**	0.368	**42413**	0.505	**42512**	0.519	**43111**	0.741
41325	0.231	**41424**	0.229	**41523**	0.423	**42122**	0.633	**42221**	0.668	**42315**	0.275	**42414**	0.273	**42513**	0.467	**43112**	0.680
41331	0.612	**41425**	0.135	**41524**	0.190	**42123**	0.580	**42222**	0.607	**42321**	0.647	**42415**	0.180	**42514**	0.234	**43113**	0.628
41332	0.551	**41431**	0.517	**41525**	0.097	**42124**	0.348	**42223**	0.555	**42322**	0.586	**42421**	0.551	**42515**	0.141	**43114**	0.395
41333	0.499	**41432**	0.456	**41531**	0.478	**42125**	0.255	**42224**	0.323	**42323**	0.534	**42422**	0.490	**42521**	0.513	**43115**	0.302
41334	0.267	**41433**	0.403	**41532**	0.417	**42131**	0.636	**42225**	0.230	**42324**	0.301	**42423**	0.438	**42522**	0.452	**43121**	0.674
41335	0.174	**41434**	0.171	**41533**	0.365	**42132**	0.575	**42231**	0.611	**42325**	0.208	**42424**	0.206	**42523**	0.400	**43122**	0.613
41341	0.357	**41435**	0.078	**41534**	0.133	**42133**	0.523	**42232**	0.550	**42331**	0.589	**42425**	0.112	**42524**	0.167	**43123**	0.561

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
43124	0.328	**43223**	0.535	**43322**	0.566	**43421**	0.531	**43515**	0.121	**44114**	0.308	**44213**	0.515	**44312**	0.546	**44411**	0.511
43125	0.235	**43224**	0.303	**43323**	0.514	**43422**	0.470	**43521**	0.493	**44115**	0.215	**44214**	0.283	**44313**	0.493	**44412**	0.450
43131	0.616	**43225**	0.210	**43324**	0.282	**43423**	0.418	**43522**	0.432	**44121**	0.586	**44215**	0.189	**44314**	0.261	**44413**	0.398
43132	0.555	**43231**	0.591	**43325**	0.188	**43424**	0.186	**43523**	0.380	**44122**	0.525	**44221**	0.561	**44315**	0.168	**44414**	0.165
43133	0.503	**43232**	0.530	**43331**	0.570	**43425**	0.093	**43524**	0.147	**44123**	0.473	**44222**	0.500	**44321**	0.540	**44415**	0.072
43134	0.271	**43233**	0.478	**43332**	0.509	**43431**	0.474	**43525**	0.054	**44124**	0.241	**44223**	0.448	**44322**	0.479	**44421**	0.444
43135	0.178	**43234**	0.246	**43333**	0.456	**43432**	0.413	**43531**	0.436	**44125**	0.148	**44224**	0.215	**44323**	0.426	**44422**	0.383
43141	0.361	**43235**	0.153	**43334**	0.224	**43433**	0.361	**43532**	0.375	**44131**	0.529	**44225**	0.122	**44324**	0.194	**44423**	0.331
43142	0.300	**43241**	0.336	**43335**	0.131	**43434**	0.128	**43533**	0.322	**44132**	0.468	**44231**	0.504	**44325**	0.101	**44424**	0.098
43143	0.248	**43242**	0.275	**43341**	0.314	**43435**	0.035	**43534**	0.090	**44133**	0.415	**44232**	0.443	**44331**	0.482	**44425**	0.005
43144	0.015	**43243**	0.223	**43342**	0.253	**43441**	0.219	**43535**	-0.003	**44134**	0.183	**44233**	0.390	**44332**	0.421	**44431**	0.386
43145	-0.078	**43244**	-0.010	**43343**	0.201	**43442**	0.158	**43541**	0.180	**44135**	0.090	**44234**	0.158	**44333**	0.369	**44432**	0.325
43151	0.259	**43245**	-0.103	**43344**	-0.031	**43443**	0.105	**43542**	0.119	**44141**	0.273	**44235**	0.065	**44334**	0.136	**44433**	0.273
43152	0.198	**43251**	0.234	**43345**	-0.124	**43444**	-0.127	**43543**	0.067	**44142**	0.212	**44241**	0.248	**44335**	0.043	**44434**	0.041
43153	0.145	**43252**	0.172	**43351**	0.212	**43445**	-0.220	**43544**	-0.165	**44143**	0.160	**44242**	0.187	**44341**	0.227	**44435**	-0.052
43154	-0.087	**43253**	0.120	**43352**	0.151	**43451**	0.116	**43545**	-0.258	**44144**	-0.072	**44243**	0.135	**44342**	0.166	**44441**	0.131
43155	-0.180	**43254**	-0.112	**43353**	0.099	**43452**	0.055	**43551**	0.078	**44145**	-0.165	**44244**	-0.097	**44343**	0.113	**44442**	0.070
43211	0.716	**43255**	-0.205	**43354**	-0.134	**43453**	0.003	**43552**	0.017	**44151**	0.171	**44245**	-0.190	**44344**	-0.119	**44443**	0.018
43212	0.655	**43311**	0.694	**43355**	-0.227	**43454**	-0.229	**43553**	-0.035	**44152**	0.110	**44251**	0.146	**44345**	-0.212	**44444**	-0.215
43213	0.602	**43312**	0.633	**43411**	0.599	**43455**	-0.322	**43554**	-0.268	**44153**	0.058	**44252**	0.085	**44351**	0.124	**44445**	-0.308
43214	0.370	**43313**	0.581	**43412**	0.538	**43511**	0.560	**43555**	-0.361	**44154**	-0.175	**44253**	0.033	**44352**	0.063	**44451**	0.029
43215	0.277	**43314**	0.349	**43413**	0.485	**43512**	0.499	**44111**	0.653	**44155**	-0.268	**44254**	-0.200	**44353**	0.011	**44452**	-0.032
43221	0.649	**43315**	0.256	**43414**	0.253	**43513**	0.447	**44112**	0.592	**44211**	0.628	**44255**	-0.293	**44354**	-0.221	**44453**	-0.085
43222	0.588	**43321**	0.627	**43415**	0.160	**43514**	0.215	**44113**	0.540	**44212**	0.567	**44311**	0.607	**44355**	-0.314	**44454**	-0.317

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
44455	-0.410	**44554**	-0.355	**45153**	0.023	**45252**	0.050	**45351**	0.089	**45445**	-0.343	**45544**	-0.288	**51143**	0.242	**51242**	0.269
44511	0.473	**44555**	-0.448	**45154**	-0.210	**45253**	-0.002	**45352**	0.028	**45451**	-0.006	**45545**	-0.381	**51144**	0.010	**51243**	0.217
44512	0.412	**45111**	0.618	**45155**	-0.303	**45254**	-0.235	**45353**	-0.024	**45452**	-0.067	**45551**	-0.045	**51145**	-0.083	**51244**	-0.015
44513	0.359	**45112**	0.557	**45211**	0.593	**45255**	-0.328	**45354**	-0.256	**45453**	-0.120	**45552**	-0.106	**51151**	0.253	**51245**	-0.108
44514	0.127	**45113**	0.505	**45212**	0.532	**45311**	0.572	**45355**	-0.349	**45454**	-0.352	**45553**	-0.158	**51152**	0.192	**51251**	0.228
44515	0.034	**45114**	0.273	**45213**	0.480	**45312**	0.510	**45411**	0.476	**45455**	-0.445	**45554**	-0.390	**51153**	0.140	**51252**	0.167
44521	0.405	**45115**	0.179	**45214**	0.247	**45313**	0.458	**45412**	0.415	**45511**	0.437	**45555**	-0.484	**51154**	-0.092	**51253**	0.115
44522	0.344	**45121**	0.551	**45215**	0.154	**45314**	0.226	**45413**	0.363	**45512**	0.376	**51111**	0.736	**51155**	-0.185	**51254**	-0.118
44523	0.292	**45122**	0.490	**45221**	0.526	**45315**	0.133	**45414**	0.130	**45513**	0.324	**51112**	0.675	**51211**	0.710	**51255**	-0.211
44524	0.060	**45123**	0.438	**45222**	0.465	**45321**	0.504	**45415**	0.037	**45514**	0.092	**51113**	0.622	**51212**	0.649	**51311**	0.689
44525	-0.033	**45124**	0.205	**45223**	0.413	**45322**	0.443	**45421**	0.409	**45515**	-0.001	**51114**	0.390	**51213**	0.597	**51312**	0.628
44531	0.348	**45125**	0.112	**45224**	0.180	**45323**	0.391	**45422**	0.348	**45521**	0.370	**51115**	0.297	**51214**	0.365	**51313**	0.576
44532	0.287	**45131**	0.494	**45225**	0.087	**45324**	0.159	**45423**	0.295	**45522**	0.309	**51121**	0.668	**51215**	0.272	**51314**	0.343
44533	0.235	**45132**	0.433	**45231**	0.469	**45325**	0.066	**45424**	0.063	**45523**	0.257	**51122**	0.607	**51221**	0.643	**51315**	0.250
44534	0.002	**45133**	0.380	**45232**	0.408	**45331**	0.447	**45425**	-0.030	**45524**	0.025	**51123**	0.555	**51222**	0.582	**51321**	0.622
44535	-0.091	**45134**	0.148	**45233**	0.355	**45332**	0.386	**45431**	0.351	**45525**	-0.068	**51124**	0.323	**51223**	0.530	**51322**	0.561
44541	0.093	**45135**	0.055	**45234**	0.123	**45333**	0.334	**45432**	0.290	**45531**	0.313	**51125**	0.230	**51224**	0.298	**51323**	0.508
44542	0.032	**45141**	0.238	**45235**	0.030	**45334**	0.101	**45433**	0.238	**45532**	0.252	**51131**	0.611	**51225**	0.205	**51324**	0.276
44543	-0.021	**45142**	0.177	**45241**	0.213	**45335**	0.008	**45434**	0.006	**45533**	0.200	**51132**	0.550	**51231**	0.586	**51325**	0.183
44544	-0.253	**45143**	0.125	**45242**	0.152	**45341**	0.192	**45435**	-0.087	**45534**	-0.033	**51133**	0.498	**51232**	0.525	**51331**	0.564
44545	-0.346	**45144**	-0.107	**45243**	0.100	**45342**	0.131	**45441**	0.096	**45535**	-0.126	**51134**	0.265	**51233**	0.473	**51332**	0.503
44551	-0.010	**45145**	-0.200	**45244**	-0.133	**45343**	0.078	**45442**	0.035	**45541**	0.057	**51135**	0.172	**51234**	0.240	**51333**	0.451
44552	-0.071	**45151**	0.136	**45245**	-0.226	**45344**	-0.154	**45443**	-0.017	**45542**	-0.004	**51141**	0.356	**51235**	0.147	**51334**	0.219
44553	-0.123	**45152**	0.075	**45251**	0.111	**45345**	-0.247	**45444**	-0.250	**45543**	-0.056	**51142**	0.295	**51241**	0.330	**51335**	0.126

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
51341	0.309	**51435**	0.030	**51534**	0.085	**52133**	0.475	**52232**	0.502	**52331**	0.541	**52425**	0.064	**52524**	0.119	**53123**	0.512
51342	0.248	**51441**	0.213	**51535**	-0.008	**52134**	0.242	**52233**	0.450	**52332**	0.480	**52431**	0.446	**52525**	0.026	**53124**	0.280
51343	0.196	**51442**	0.152	**51541**	0.175	**52135**	0.149	**52234**	0.217	**52333**	0.428	**52432**	0.385	**52531**	0.407	**53125**	0.187
51344	-0.037	**51443**	0.100	**51542**	0.114	**52141**	0.333	**52235**	0.124	**52334**	0.196	**52433**	0.332	**52532**	0.346	**53131**	0.568
51345	-0.130	**51444**	-0.132	**51543**	0.062	**52142**	0.272	**52241**	0.307	**52335**	0.103	**52434**	0.100	**52533**	0.294	**53132**	0.507
51351	0.207	**51445**	-0.226	**51544**	-0.171	**52143**	0.219	**52242**	0.246	**52341**	0.286	**52435**	0.007	**52534**	0.062	**53133**	0.455
51352	0.146	**51451**	0.111	**51545**	-0.264	**52144**	-0.013	**52243**	0.194	**52342**	0.225	**52441**	0.190	**52535**	-0.031	**53134**	0.223
51353	0.093	**51452**	0.050	**51551**	0.073	**52145**	-0.106	**52244**	-0.038	**52343**	0.173	**52442**	0.129	**52541**	0.152	**53135**	0.130
51354	-0.139	**51453**	-0.002	**51552**	0.011	**52151**	0.230	**52245**	-0.131	**52344**	-0.060	**52443**	0.077	**52542**	0.091	**53141**	0.313
51355	-0.232	**51454**	-0.235	**51553**	-0.041	**52152**	0.169	**52251**	0.205	**52345**	-0.153	**52444**	-0.155	**52543**	0.039	**53142**	0.252
51411	0.593	**51455**	-0.328	**51554**	-0.273	**52153**	0.117	**52252**	0.144	**52351**	0.184	**52445**	-0.249	**52544**	-0.194	**53143**	0.200
51412	0.532	**51511**	0.555	**51555**	-0.366	**52154**	-0.115	**52253**	0.092	**52352**	0.123	**52451**	0.088	**52545**	-0.287	**53144**	-0.033
51413	0.480	**51512**	0.494	**52111**	0.713	**52155**	-0.208	**52254**	-0.141	**52353**	0.070	**52452**	0.027	**52551**	0.050	**53145**	-0.126
51414	0.247	**51513**	0.441	**52112**	0.651	**52211**	0.687	**52255**	-0.234	**52354**	-0.162	**52453**	-0.025	**52552**	-0.012	**53151**	0.211
51415	0.154	**51514**	0.209	**52113**	0.599	**52212**	0.626	**52311**	0.666	**52355**	-0.255	**52454**	-0.258	**52553**	-0.064	**53152**	0.150
51421	0.526	**51515**	0.116	**52114**	0.367	**52213**	0.574	**52312**	0.605	**52411**	0.570	**52455**	-0.351	**52554**	-0.296	**53153**	0.097
51422	0.465	**51521**	0.488	**52115**	0.274	**52214**	0.342	**52313**	0.553	**52412**	0.509	**52511**	0.532	**52555**	-0.389	**53154**	-0.135
51423	0.413	**51522**	0.427	**52121**	0.645	**52215**	0.249	**52314**	0.320	**52413**	0.457	**52512**	0.471	**53111**	0.693	**53155**	-0.228
51424	0.180	**51523**	0.374	**52122**	0.584	**52221**	0.620	**52315**	0.227	**52414**	0.224	**52513**	0.418	**53112**	0.632	**53211**	0.668
51425	0.087	**51524**	0.142	**52123**	0.532	**52222**	0.559	**52321**	0.599	**52415**	0.131	**52514**	0.186	**53113**	0.580	**53212**	0.607
51431	0.469	**51525**	0.049	**52124**	0.300	**52223**	0.507	**52322**	0.538	**52421**	0.503	**52515**	0.093	**53114**	0.347	**53213**	0.554
51432	0.408	**51531**	0.430	**52125**	0.207	**52224**	0.275	**52323**	0.485	**52422**	0.442	**52521**	0.465	**53115**	0.254	**53214**	0.322
51433	0.355	**51532**	0.369	**52131**	0.588	**52225**	0.182	**52324**	0.253	**52423**	0.390	**52522**	0.404	**53121**	0.626	**53215**	0.229
51434	0.123	**51533**	0.317	**52132**	0.527	**52231**	0.563	**52325**	0.160	**52424**	0.157	**52523**	0.351	**53122**	0.565	**53221**	0.601

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
53222	0.540	**53321**	0.579	**53415**	0.112	**53514**	0.166	**54113**	0.492	**54212**	0.519	**54311**	0.559	**54355**	-0.362	**54454**	-0.365
53223	0.487	**53322**	0.518	**53421**	0.483	**53515**	0.073	**54114**	0.260	**54213**	0.467	**54312**	0.497	**54411**	0.463	**54455**	-0.458
53224	0.255	**53323**	0.466	**53422**	0.422	**53521**	0.445	**54115**	0.166	**54214**	0.234	**54313**	0.445	**54412**	0.402	**54511**	0.424
53225	0.162	**53324**	0.233	**53423**	0.370	**53522**	0.384	**54121**	0.538	**54215**	0.141	**54314**	0.213	**54413**	0.349	**54512**	0.363
53231	0.543	**53325**	0.140	**53424**	0.138	**53523**	0.332	**54122**	0.477	**54221**	0.513	**54315**	0.120	**54414**	0.117	**54513**	0.311
53232	0.482	**53331**	0.522	**53425**	0.045	**53524**	0.099	**54123**	0.425	**54222**	0.452	**54321**	0.491	**54415**	0.024	**54514**	0.079
53233	0.430	**53332**	0.461	**53431**	0.426	**53525**	0.006	**54124**	0.192	**54223**	0.400	**54322**	0.430	**54421**	0.396	**54515**	-0.014
53234	0.198	**53333**	0.408	**53432**	0.365	**53531**	0.388	**54125**	0.099	**54224**	0.167	**54323**	0.378	**54422**	0.335	**54521**	0.357
53235	0.104	**53334**	0.176	**53433**	0.313	**53532**	0.327	**54131**	0.481	**54225**	0.074	**54324**	0.146	**54423**	0.282	**54522**	0.296
53241	0.288	**53335**	0.083	**53434**	0.080	**53533**	0.274	**54132**	0.420	**54231**	0.456	**54325**	0.053	**54424**	0.050	**54523**	0.244
53242	0.227	**53341**	0.266	**53435**	-0.013	**53534**	0.042	**54133**	0.367	**54232**	0.394	**54331**	0.434	**54425**	-0.043	**54524**	0.012
53243	0.174	**53342**	0.205	**53441**	0.170	**53535**	-0.051	**54134**	0.135	**54233**	0.342	**54332**	0.373	**54431**	0.338	**54525**	-0.081
53244	-0.058	**53343**	0.153	**53442**	0.109	**53541**	0.132	**54135**	0.042	**54234**	0.110	**54333**	0.321	**54432**	0.277	**54531**	0.300
53245	-0.151	**53344**	-0.079	**53443**	0.057	**53542**	0.071	**54141**	0.225	**54235**	0.017	**54334**	0.088	**54433**	0.225	**54532**	0.239
53251	0.185	**53345**	-0.173	**53444**	-0.175	**53543**	0.019	**54142**	0.164	**54241**	0.200	**54335**	-0.005	**54434**	-0.007	**54533**	0.187
53252	0.124	**53351**	0.164	**53445**	-0.268	**53544**	-0.214	**54143**	0.112	**54242**	0.139	**54341**	0.179	**54435**	-0.100	**54534**	-0.046
53253	0.072	**53352**	0.103	**53451**	0.068	**53545**	-0.307	**54144**	-0.120	**54243**	0.087	**54342**	0.118	**54441**	0.083	**54535**	-0.139
53254	-0.160	**53353**	0.051	**53452**	0.007	**53551**	0.030	**54145**	-0.213	**54244**	-0.146	**54343**	0.065	**54442**	0.022	**54541**	0.044
53255	-0.253	**53354**	-0.182	**53453**	-0.045	**53552**	-0.031	**54151**	0.123	**54245**	-0.239	**54344**	-0.167	**54443**	-0.030	**54542**	-0.017
53311	0.646	**53355**	-0.275	**53454**	-0.277	**53553**	-0.083	**54152**	0.062	**54251**	0.098	**54345**	-0.260	**54444**	-0.263	**54543**	-0.069
53312	0.585	**53411**	0.550	**53455**	-0.371	**53554**	-0.316	**54153**	0.010	**54252**	0.037	**54351**	0.076	**54445**	-0.356	**54544**	-0.301
53313	0.533	**53412**	0.489	**53511**	0.512	**53555**	-0.409	**54154**	-0.223	**54253**	-0.016	**54352**	0.015	**54451**	-0.019	**54545**	-0.394
53314	0.301	**53413**	0.437	**53512**	0.451	**54111**	0.605	**54155**	-0.316	**54254**	-0.248	**54353**	-0.037	**54452**	-0.081	**54551**	-0.058
53315	0.207	**53414**	0.205	**53513**	0.399	**54112**	0.544	**54211**	0.580	**54255**	-0.341	**54354**	-0.269	**54453**	-0.133	**54552**	-0.119

Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values	Profile	Utility values
54553	-0.171	**55133**	0.332	**55213**	0.432	**55243**	0.052	**55323**	0.343	**55353**	-0.072	**55433**	0.190	**55513**	0.276	**55543**	-0.104
54554	-0.403	**55134**	0.100	**55214**	0.199	**55244**	-0.181	**55324**	0.111	**55354**	-0.304	**55434**	-0.042	**55514**	0.044	**55544**	-0.336
54555	-0.497	**55135**	0.007	**55215**	0.106	**55245**	-0.274	**55325**	0.018	**55355**	-0.398	**55435**	-0.136	**55515**	-0.049	**55545**	-0.429
55111	0.570	**55141**	0.190	**55221**	0.478	**55251**	0.063	**55331**	0.399	**55411**	0.428	**55441**	0.048	**55521**	0.322	**55551**	-0.093
55112	0.509	**55142**	0.129	**55222**	0.417	**55252**	0.002	**55332**	0.338	**55412**	0.367	**55442**	-0.013	**55522**	0.261	**55552**	-0.154
55113	0.457	**55143**	0.077	**55223**	0.365	**55253**	-0.051	**55333**	0.286	**55413**	0.314	**55443**	-0.066	**55523**	0.209	**55553**	-0.206
55114	0.224	**55144**	-0.155	**55224**	0.132	**55254**	-0.283	**55334**	0.053	**55414**	0.082	**55444**	-0.298	**55524**	-0.023	**55554**	-0.439
55115	0.131	**55145**	-0.249	**55225**	0.039	**55255**	-0.376	**55335**	-0.040	**55415**	-0.011	**55445**	-0.391	**55525**	-0.116	**55555**	-0.532
55121	0.503	**55151**	0.088	**55231**	0.420	**55311**	0.523	**55341**	0.143	**55421**	0.361	**55451**	-0.055	**55531**	0.265	**Uncon-scious**	-0.454
55122	0.442	**55152**	0.027	**55232**	0.359	**55312**	0.462	**55342**	0.082	**55422**	0.300	**55452**	-0.116	**55532**	0.204		
55123	0.390	**55153**	-0.025	**55233**	0.307	**55313**	0.410	**55343**	0.030	**55423**	0.247	**55453**	-0.168	**55533**	0.152		
55124	0.157	**55154**	-0.258	**55234**	0.075	**55314**	0.178	**55344**	-0.202	**55424**	0.015	**55454**	-0.400	**55534**	-0.081		
55125	0.064	**55155**	-0.351	**55235**	-0.018	**55315**	0.085	**55345**	-0.295	**55425**	-0.078	**55455**	-0.493	**55535**	-0.174		
55131	0.446	**55211**	0.545	**55241**	0.165	**55321**	0.456	**55351**	0.041	**55431**	0.303	**55511**	0.389	**55541**	0.009		
55132	0.385	**55212**	0.484	**55242**	0.104	**55322**	0.395	**55352**	-0.020	**55432**	0.242	**55512**	0.328	**55542**	-0.052		

REFERENCES

1. Cleemput I, Devriese S, Christiaens W, Kohn L. Multi-criteria decision analysis for the appraisal of medical needs: a pilot study. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2016. KCE Reports (272) Available from: https://kce.fgov.be/sites/default/files/atoms/files/KCE_272_Unmet_needs_Report2.pdf
2. Cleemput I, Devriese S, Kohn L, Devos C, van Til J, Groothuis-Oudshoorn K, et al. Incorporating societal preferences in reimbursement decisions – Relative importance of decision criteria according to Belgian citizens. Health Services Research (HSR). Brussel: Belgian Health Care Knowledge Centre (KCE); 2014. KCE Reports (234) Available from: https://kce.fgov.be/sites/default/files/atoms/files/KCE_234_reimbursement_decisions_Report_0_1.pdf
3. EUnetHTA. Guideline: Endpoints used for relative effectiveness assessment of pharmaceuticals, health-related quality of life and utility measures. European network for Health Technology Assessment; 2013. Available from: https://eunethta.eu/wp-content/uploads/2013/01/Health-related-quality-of-life.pdf
4. Karimi M, Brazier J. Health, health-related quality of life, and quality of life: what is the difference? PharmacoEconomics. 2016;34(7):645-9.
5. European Medicines Agency. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. London: 2006. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-regulatory-guidance-use-healthrelated-quality-life-hrql-measures-evaluation_en.pdf
6. Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. Available from: https://www.fda.gov/media/77832/download

7. Cleemput I, Neyt M. Which quality of life measures fit your relative effectiveness assessment? International Journal of Technology Assessment in Health Care. 2015;31(3):147-53.

8. Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials: a review. Health Technology Assessment. 1998;2(14):1-74.

9. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. Health Technology Assessment. 1999;3(9):i-iv, 1-164.

10. Calvert MJ, Freemantle N. Use of health-related quality of life in prescribing research. Part 2: methodological considerations for the assessment of health-related quality of life in clinical trials. Journal of Clinical Pharmacy and Therapeutics. 2004;29(1):85-94.

11. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. British Medical Bulletin. 2010;96(1):5-21.

12. Bleichrodt H, Johannesson M. Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. Journal of Health Economics. 1997;16(2):155-75.

13. Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Second edition. Oxford university press; 2017.

14. Rowen D, Brazier J, van Hout B. A comparison of methods for converting DCE values onto the full health-dead QALY scale. Medical Decision Making. 2015;35(3):328-40.

15. Lugnér AK, Krabbe PFM. An overview of the time trade-off method: concept, foundation, and the evaluation of distorting factors in putting a value on health. Expert Review of Pharmacoeconomics & Outcomes Research. 2020;20(4):331-42.

16. Green C, Brazier J, Deverill M. Valuing health-related quality of life. PharmacoEconomics. 2000;17(2):151-65.

17. Krabbe PFM, Devlin NJ, Stolk EA, Shah KK, Oppe M, van Hout B, et al. Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-state values. Medical care. 2014;52(11):935-43.

18. Norman R, Cronin P, Viney R. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. Applied Health Economics and Health Policy. 2013;11(3):287-98.

19. Stolk EA, Oppe M, Scalone L, Krabbe PFM. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. Value in Health. 2010;13(8):1005-13.

20. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. Journal of Health Economics. 2012;31(1):306-18.

21. Petrou S, Morrell J, Spiby H. Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. Health and Quality of Life Outcomes. 2009;7:40.

22. Pullenayegum E, Xie F. Scoring the 5-level EQ-5D: can latent utilities derived from a discrete choice model be transformed to health utilities derived from time tradeoff tasks? Medical Decision Making. 2013;33(4):567-78.

23. Shiroiwa T, Ikeda S, Noto S, Igarashi A, Fukuda T, Saito S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. Value in Health. 2016;19(5):648-54.

24. Mulhern B, Norman R, Shah K, Bansback N, Longworth L, Viney R. How should discrete choice experiments with duration choice sets be presented for the valuation of health states? Medical Decision Making. 2018;38(3):306-18.

25. Cleemput I, Neyt M, Van de Sande S, Thiry N. Belgian guidelines for economic evaluations and budget impact analyses: second edition. Health Technology Assessment (HTA). Brussels: Belgian Health Care Knowledge Centre (KCE); 2012. KCE Reports (183) Available

from: https://kce.fgov.be/sites/default/files/page_documents/KCE_183C_economic_evaluations_second_edition_0.pdf

26. Brazier J, Akehurst R, Brennan A, Dolan P, Claxton K, McCabe C, et al. Should patients have a greater role in valuing health states? Applied Health Economics and Health Policy. 2005;4(4):201-8.
27. Kreimeier S, Greiner W. EQ-5D-Y as a health-related quality of life instrument for children and adolescents: the instrument's characteristics, development, current use, and challenges of developing its value set. Value in Health. 2019;22(1):31-7.
28. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goñi JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. PharmacoEconomics. 2016;34(10):993-1004.
29. EuroQol. EuroQol | About Us [Web page]. 2021 [cited 20/05/2021]. Available from: https://euroqol.org/euroqol/
30. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. The European Journal of Health Economics. 2020;21(8):1245-57.
31. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
32. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research. 2011;20(10):1727-36.
33. EuroQol Research Foundation. EQ-5D-5L User Guide (version 3.0). Rotterdam: 2019. Available from: https://euroqol.org/publications/user-guides
34. Harrison MJ, Davies LM, Bansback NJ, McCoy MJ, Verstappen SM, Watson K, et al. The comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. Quality of Life Research. 2009;18(9):1195-205.
35. Perneger TV, Combescure C, Courvoisier DS. General population reference values for the French version of the EuroQol EQ-5D health utility instrument. Value in Health. 2010;13(5):631-5.
36. Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. Value in Health. 2008;11(2):275-84.
37. Bharmal M, Thomas J, 3rd. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. Value in Health. 2006;9(4):262-71.
38. Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Medical Care. 2007;45(3):259-63.
39. Ludwig K, Graf von der Schulenburg JM, Greiner W. German value set for the EQ-5D-5L. PharmacoEconomics. 2018;36(6):663-74.
40. Andrade LF, Ludwig K, Ramos-Goñi JM, Oppe M, de Pouvourville G. A French value set for the EQ-5D-5L. PharmacoEconomics. 2020;38(4):413-25.
41. Hobbins A, Barry L, Kelleher D, Shah K, Devlin N, Ramos-Goñi JM, et al. Utility values for health states in Ireland: a value set for the EQ-5D-5L. PharmacoEconomics. 2018;36(11):1345-53.
42. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health economics. 2018;27(1):7-22.
43. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. Value in Health. 2016;19(4):343-52.
44. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabases JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L

health states using a hybrid approach. Medical Care. 2017;55(7):e51-e8.

45. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goñi JM. Overview, update, and lessons learned from the international EQ-5D-5L valuation work: version 2 of the EQ-5D-5L valuation protocol. Value in Health. 2019;22(1):23-30.

46. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value in Health. 2014;17(4):445-53.

47. Cleemput I. A social preference valuations set for EQ-5D health states in Flanders, Belgium. The European Journal of Health Economics. 2010;11(2):205-13.

48. van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15(5):708-15.

49. Bilcke J, Hens N, Beutels P. Quality-of-life: a many-splendored thing? Belgian population norms and 34 potential determinants explored by beta regression. Quality of Life Research. 2017;26(8):2011-23.

50. Van Wilder L, Charafeddine R, Beutels P, Bruyndonckx R, Cleemput I, Demarest S, et al. Belgian population norms for the EQ-5D-5L, 2013 and 2018. Under Review. 2021.

51. Szende A, Janssen B, Cabases J. Self-reported population health: an international perspective based on EQ-5D. 2014.

52. Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. Health Economics. 2010;19(8):886-905.

53. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present and future. Applied Health Economics and Health Policy. 2017;15(2):127-37.

54. WIV-ISP. Health Interview Survey, Belgium, 2013. Auto (self completed) questionnaire [Web page].Brussels: Scientific Institute of Public Health; 2017. Available from: https://his.wiv-isp.be/Shared%20Documents/qauto_2013.pdf

55. Sciensano. Health Interview Survey [Web page]. 2021 [cited 20/05/2021]. Available from: https://his.wiv-isp.be/SitePages/Home.aspx

56. Charafeddine R. Qualité de vie liée à la santé. In: Van der Heyden J, Charafeddine R, editors. Enquête de santé 2013. Rapport 1: santé et bien-être. Brussels: WIV-ISP; 2014. Available from: https://his.wiv-isp.be/fr/Documents%20partages/QL_FR_2013.pdf

57. Charafeddine R. Gezondheidsgerelateerde kwaliteit van leven. In: Van der Heyden J, Charafeddine R, editors. Gezondheidsenquête 2013. Rapport 1: gezondheid en welzijn. Brussels: WIV-ISP; 2014. Available from: https://his.wiv-isp.be/nl/Gedeelde%20%20documenten/QL_NL_2013.pdf

58. Charafeddine R, Braekman E, Van der Heyden J. Enquête de santé 2018: qualité de vie liée à la santé. Brussels: 2019. Available from: https://his.wiv-isp.be/fr/Documents%20partages/QL_FR_2018.pdf

59. Charafeddine R, Braekman E, Van der Heyden J. Gezondheidsenquête 2018: gezondheidsgerelateerde kwaliteit van leven. Brussels: 2019. Available from: https://his.wiv-isp.be/nl/Gedeelde%20%20documenten/QL_NL_2018.pdf

60. Janssen BMF, Oppe M, Versteegh MM, Stolk EA. Introducing the composite time trade-off: a test of feasibility and face validity. The European Journal of Health Economics. 2013;14(1):5-13.

61. Shah KK, Lloyd A, Oppe M, Devlin NJ. One-to-one versus group setting for conducting computer-assisted TTO studies: findings from pilot studies in England and the Netherlands. The European Journal of Health Economics. 2013;14(1):65-73.

62. Rencz F, Brodszky V, Gulácsi L, Golicki D, Ruzsa G, Pickard AS, et al. Parallel valuation of the EQ-5D-3L and EQ-5D-5L by time trade-off in Hungary. Value in Health. 2020;23(9):1235-45.

63. Oppe M, van Hout B. The “power” of eliciting EQ-5D-5L values: the experimental design of the EQ-VT. Rotterdam: EuroQol Group; 2017. EuroQol Working Paper Series 17003 Available from: https://euroqol.org/wp-content/uploads/2016/10/EuroQol-Working-Paper-Series-Manuscript-17003-Mark-Oppe.pdf

64. van Hout B, Mulhern B, Feng Y, Shah K, Devlin N. The EQ-5D-5L value set for England: response to the "Quality Assurance". Value in Health. 2020;23(5):649-55.

65. Shah K, Rand-Hendriksen K, Ramos JM, Prause AJ, Stolk E. Improving the quality of data collected in EQ-5D-5L valuation studies: a summary of the EQ-VT research methodology programme. In: Proceedings of 31st Scientific Plenary Meeting of the EuroQol Group; 2014 September 25-27; Stockholm, Sweden. Available from: https://eq-5dpublications.euroqol.org/download?id=0_53918&fileId=54332

66. Wong ELY, Ramos-Goñi JM, Cheung AWL, Wong AYK, Rivero-Arias O. Assessing the use of a feedback module to model EQ-5D-5L health states values in Hong Kong. The patient. 2018;11(2):235-47.

67. Wong ELY, Shah K, Cheung AWL, Wong AYK, Visser M, Stolk E. Evaluation of split version and feedback module on the improvement of time trade-off data. Value in Health. 2018;21(6):732-41.

68. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJ, Stolk E. Quality control process for EQ-5D-5L valuation studies. Value in Health. 2017;20(3):466-73.

69. Heeringa SG, West BT, Berglund PA. Applied survey data analysis (2nd ed.). Boca Raton: Chapman and Hall/CRC; 2017.

70. Charafeddine R, Demarest S, Drieskens S, Gisle L, Tafforeau J, Van der Heyden J. Health interview survey 2013. Research protocol. Brussels: WIV-ISP; 2013. Available from: https://his.wiv-isp.be/Shared%20Documents/Protocol2013.pdf

71. Statistics Belgium. Population by place of residence, nationality, marital status, age and sex [Web page]. 2017 [cited 1/03/2021]. Available from: https://statbel.fgov.be/en/open-data/population-place-residence-nationality-marital-status-age-and-sex-0

72. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. Medical Care. 2016;54(1):98-105.

73. Augustovski F, Rey-Ares L, Irazola V, Garay OU, Gianneo O, Fernandez G, et al. An EQ-5D-5L value set based on Uruguayan population preferences. Quality of Life Research. 2016;25(2):323-33.

74. Kim S-H, Ahn J, Ock M, Shin S, Park J, Luo N, et al. The EQ-5D-5L valuation study in Korea. Quality of Life Research. 2016;25(7):1845-52.

75. Ferreira PL, Antunes P, Ferreira LN, Pereira LN, Ramos-Goñi JM. A hybrid modelling approach for eliciting health state preferences: the Portuguese EQ-5D-5L value set. Quality of Life Research. 2019;28(12):3163-75.

76. Golicki D, Jakubczyk M, Graczyk K, Niewada M. Valuation of EQ-5D-5L health states in Poland: the first EQ-VT-based study in central and eastern Europe. PharmacoEconomics. 2019;37(9):1165-76.

77. Eurostat. EU statistics on income and living conditions [Web page]. 2021 [cited 20/05/2021]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:EU_statistics_on_income_and_living_conditions_(EU-SILC)

78. Statistics Belgium. Quality report Belgian SILC 2019 [Web page].Brussels; 2020 [cited 1/09/2020]. Available from: https://statbel.fgov.be/sites/default/files/files/documents/Huishouden

s/10.7%20Inkomen%20en%20levensomstandigheden/10.7.1%20Armoederisico/Plus/FR/BE_QualityReport_2019.pdf

79. Rand-Hendriksen K, Ramos-Goñi JM, Augestad LA, Luo N. Less is more: cross-validation testing of simplified nonlinear regression model specifications for EQ-5D-5L health state values. Value in Health. 2017;20(7):945-52.

80. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value in Health. 2017;20(4):662-9.

81. Feng Y, Devlin N, Shah K, Mulhern B, van Hout B. New methods for modelling EQ-5D-5L value sets: an application to English data. London: Office of Health Economics; 2016. Research Paper 16/02 Available from: https://www.ohe.org/publications/new-methods-modelling-eq-5d-5l-value-sets-application-english-data

82. Oppe M, Ramos-Goñi JM, van Hout B. Modeling EQ-5D-5L valuation data. In: Busschbach JJ, editor. Proceedings of 29th Scientific Plenary Meeting of the EuroQol Group; 2012 September 13-15 2012; Rotterdam; p. 61-91. Available from: https://eq-5dpublications.euroqol.org/download?id=0_53782&fileId=54196

83. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: past, present and future. PharmacoEconomics. 2019;37(2):201-26.

84. OECD. Discrete choice experiments. In: Cost-benefit analysis and the environment: further developments and policy use. Paris: OECD Publishing; 2018. Available from: https://www.oecd-ilibrary.org/content/component/9789264085169-8-en

85. Dolan P. Modeling Valuations for EuroQol Health States. Medical Care. 1997;35(11):1095-108.

86. Purba FD, Hunfeld JAM, Iskandarsyah A, Fitriana TS, Sadarjoen SS, Ramos-Goñi JM, et al. The Indonesian EQ-5D-5L value set. PharmacoEconomics. 2017.

87. Al Shabasy SA, Abbassi MM, Finch AP, Baines D, Farid SF. The EQ-5D-5L valuation study in Egypt. PharmacoEconomics. 2021;39(5):549-61.

88. Lin H-W, Li C-I, Lin F-J, Chang J-Y, Gau C-S, Luo N, et al. Valuation of the EQ-5D-5L in Taiwan. PLOS ONE. 2018;13(12):e0209344.

89. Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, et al. United States valuation of EQ-5D-5L health states using an international protocol. Value in Health. 2019;22(8):931-41.

90. Welie AG, Gebretekle GB, Stolk E, Mukuria C, Krahn MD, Enquoselassie F, et al. Valuing health state: an EQ-5D-5L value set for Ethiopians. Value in Health Regional Issues. 2020;22:7-14.

91. Pattanaphesaj J, Thavorncharoensap M, Ramos-Goñi JM, Tongsiri S, Ingsrisawang L, Teerawattananon Y. The EQ-5D-5L valuation study in Thailand. Expert Review of Pharmacoeconomics & Outcomes Research. 2018;18(5):551-8.

92. Mai VQ, Sun S, Minh HV, Luo N, Giang KB, Lindholm L, et al. An EQ-5D-5L Value Set for Vietnam. Quality of Life Research. 2020;29(7):1923-33.

93. Shafie AA, Vasan Thakumar A, Lim CJ, Luo N, Rand-Hendriksen K, Md Yusof FA. EQ-5D-5L valuation for the Malaysian population. PharmacoEconomics. 2019;37(5):715-25.

94. Ramos-Goñi JM, Craig B, Oppe M, van Hout B. Combining continuous and dichotomous responses in a hybrid model. Rotterdam: EuroQol Research Foundation; 2016. EuroQol Working Paper Series 16002 Available from: https://www.euroqol.org/wp-content/uploads/working_paper_series/EuroQol_Working_Paper_Series_Manuscript_16002_-_Juan_Ramos-Goni.pdf

95. Ramos-Goñi JM, Craig BM, Oppe M, Ramallo-Fariña Y, Pinto-Prades JL, Luo N, et al. Handling data quality Issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. Value in Health. 2018;21(5):596-604.

96. SAS Institute Inc. SAS/STAT®14.1 user's guide: the NLMIXED procedure. Cary, NC: SAS Institute Inc.; 2015.

97. Van de Voorde C, Van den Heede K, Beguin C, Bouckaert N, Camberlin C, de Bekker P, et al. Required hospital capacity in 2025 and criteria for rationalisation of complex cancer surgery, radiotherapy and maternity services. Health Services Research (HSR). Brussel: Belgian Health Care Knowledge Centre (KCE); 2017. KCE Reports (289) Available from: https://kce.fgov.be/sites/default/files/atoms/files/Download%20het%20rapport%20in%20het%20Engels%20%28550%20p.%29.pdf

98. Berete F, Braekman E, Charafeddine R, Demarest S, Drieskens S, Gisle L, et al. Vijfde COVID-19-Gezondheidsenquête. Eerste resultaten. Brussel: Sciensano; 2020. Available from: https://www.sciensano.be/nl/biblio/cinquieme-enquete-de-sante-coivd-19-resultats-preliminaires

99. Federal Public Service Social Security. Analysis of the evolution of the social situation and social protection in Belgium 2020: Persistent challenges. Brussels: Federal Public Service Social Security; 2021. Available from: https://socialsecurity.belgium.be/sites/default/files/content/docs/en/publications/silc/silc-analysis-social-situation-and-protection-belgium-2020-en.pdf

100. SAS Institute Inc. SAS/STAT®14.1 user's guide: the MODEL procedure. Cary, NC: SAS Institute Inc.; 2015.

101. Knies S, Evers SMAA, Candel MJJM, Severens JL, Ament AJHA. Utilities of the EQ-5D. PharmacoEconomics. 2009;27(9):767-79.

102. Desomer A, Van den Heede K, Triemstra M, Paget J, De Boer D, Kohn L, et al. Use of patient-reported outcome and experience measures in patient care and policy – Short report. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2018. KCE Reports (303Cs) Available from: https://kce.fgov.be/sites/default/files/atoms/files/KCE_303C_Patient_reported_outcomes_Short_Report_0.pdf

103. Hernandez Alava M, Wailoo A, Grimm S, Pudney S, Gomes M, Sadique Z, et al. EQ-5D-5L versus EQ-5D-3L: the impact on cost effectiveness in the United Kingdom. Value in Health. 2018;21(1):49-56.

104. Ramos-Goñi JM, Oppe M, Stolk E, Shah K, Kreimeier S, Rivero-Arias O, et al. International valuation protocol for the EQ-5D-Y-3L. PharmacoEconomics. 2020;38(7):653-63.

105. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of Life Research. 2010;19(6):875-86.

106. Gerkens S, Merkur S. Belgium: Health system review. Health Systems in Transition. 2020;22(5):237.