



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

The Danish EQ-5D-5L Value Set

A Hybrid Model Using cTTO and DCE Data

Jensen, Cathrine Elgaard; Sørensen, Sabrina Storgaard; Gudex, Claire; Jensen, Morten Berg; Pedersen, Kjeld Møller; Ehlers, Lars Holger

Published in:
Applied Health Economics and Health Policy

DOI (link to publication from Publisher):
[10.1007/s40258-021-00639-3](https://doi.org/10.1007/s40258-021-00639-3)

Creative Commons License
CC BY-NC 4.0

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Jensen, C. E., Sørensen, S. S., Gudex, C., Jensen, M. B., Pedersen, K. M., & Ehlers, L. H. (2021). The Danish EQ-5D-5L Value Set: A Hybrid Model Using cTTO and DCE Data. *Applied Health Economics and Health Policy*, 19(4), 579-591. Advance online publication. <https://doi.org/10.1007/s40258-021-00639-3>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



The Danish EQ-5D-5L Value Set: A Hybrid Model Using cTTO and DCE Data

Cathrine Elgaard Jensen¹ · Sabrina Storgaard Sørensen¹ · Claire Gudex² · Morten Berg Jensen³ · Kjeld Møller Pedersen⁴ · Lars Holger Ehlers¹

Accepted: 7 January 2021 / Published online: 2 February 2021
© The Author(s) 2021

Abstract

Objectives Quality-adjusted life-years (QALYs) are expected to be used for priority setting of hospital-dispensed medicines in Denmark from 2021. The aim of this study was to develop the first Danish value set for the EQ-5D-5L based on interviews with a representative sample of the Danish adult population.

Methods A nationally representative sample based on age (> 18 years), gender, education, and geographical region was recruited using data provided by Statistics Denmark. Computer-assisted personal interviews were carried out using the EQ-VT 2.1. Respondents each valued ten health states using composite time trade-off (cTTO) and seven health states using discrete-choice experiment (DCE). Different predictive models were explored using cTTO and DCE data alone or in combination as hybrid models. Model performance was assessed using logical consistency.

Results A total of 1014 interviews were included in the analyses. The sample was representative of the Danish adult population, though the sample contained slightly more respondents with higher education than in the general population. Only the heteroscedastic censored hybrid model combining cTTO and DCE data yielded consistent results, and hence was chosen for modelling the final Danish value set. The predicted values ranged from – 0.757 to 1, and anxiety/depression was the dimension assigned most value by respondents.

Conclusions This study established the Danish EQ-5D-5L value set, which represents the preferences of the Danish general population, and is expected to provide key input for healthcare decision-making in a Danish context.

Key Points for Decision Makers

This paper presents the Danish EQ-5D-5L value set based on a representative sample of the Danish adult population. The study is characterized by high-quality data according to EQ-VT quality indicators, which is believed to be due to the use of a standard and detailed interview protocol, extensive interviewer training, and quality control during data collection.

The recruitment strategy enabled a continuous monitoring of the representativeness of the sample and targeted recruitment of under-represented groups.

The paper adds to the existing literature by demonstrating the heteroscedastic hybrid model combining composite time trade-off (cTTO) and discrete-choice experiment (DCE) data as an applicable approach to obtain an EQ-5D-5L value set for healthcare prioritization.

✉ Cathrine Elgaard Jensen
cathelgaard@gmail.com

¹ Department of Clinical Medicine, Danish Center for Healthcare Improvements, Aalborg University, Aalborg, Denmark

² Department of Clinical Research, University of Southern Denmark and OPEN - Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark

³ Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark

⁴ Department of Management and Economics, University of Southern Denmark, Odense, Denmark

1 Introduction

The EQ-5D is the most commonly used generic measure to elicit patient-reported health-related quality of life (HRQoL) for estimation of quality-adjusted life-years (QALYs) [1, 2]. In Denmark, the need for relevant HRQoL weights for calculating QALYs is greater than ever before as the Danish Medicine Council will initiate use of cost-utility analyses to assess new and existing medicines across hospitals and regions by 2021 [3, 4]. A Danish value set is available for the EQ-5D-3L [5, 6], in which the five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) have three levels of severity [7]. Although an interim “crosswalk” value set is available for the newer EQ-5D-5L [8, 9], in which the five dimensions have five levels of severity [10, 11], a Danish EQ-5D-5L valuation study has not yet been conducted. In the new Danish guidelines for economic evaluation of new pharmaceuticals, EQ-5D-5L is described as the “reference case” that should be used as first choice for estimating QALYs [12].

While EQ-5D-3L and EQ-5D-5L value sets should show similar trends, the extra levels in the EQ-5D-5L generate a larger number of health states, and the wording of the severity levels differs between the versions. Thus, the most severe mobility level has been changed from “confined to bed” (3L) to “unable to walk about” (5L), and the middle levels in mobility, self-care, and usual activities dimensions have been changed from “some problems” (3L) to “moderate problems” (5L) as in the pain/discomfort and anxiety/depression dimensions [10, 13]. Furthermore, while preferences for EQ-5D-3L health states were elicited using conventional time trade-off (TTO), EQ-5D-5L valuation studies use composite TTO (cTTO), i.e. conventional TTO to value health states considered better than dead combined with lead-time TTO to value health states considered worse than dead [14, 15]. The EuroQol Valuation Technology (EQ-VT) also includes discrete-choice experiment (DCE) to value EQ-5D-5L health states [16]. DCE values may follow a similar pattern to TTO values [17], but DCE values lie on an arbitrary scale rather than one anchored at 0 (death) and 1 (full health) as required by the QALY model. Recently, a large volume of work has been carried out with the purpose of addressing this anchoring problem. See for instance [18–23] for examples of this work. Until this anchoring problem is properly resolved, DCE cannot be an alternative to the TTO approach, but DCE data may add extra information to produce a better model for valuation data. So-called hybrid models combining cTTO and DCE data have thus been used for several recent EQ-5D-5L value sets [24–28].

The aim of the present study was to generate a Danish value set for the EQ-5D-5L based on interviews with a

representative sample of the adult Danish general population using the standardized EQ-VT. The use of the standardized EQ-VT could potentially also allow for comparisons on a more equal footing across populations. An important aspect was to identify the best modelling approach for the final value set, given the choice of cTTO data alone or in combination with DCE data.

2 Methods

The reporting of the Danish valuation study follows the CREATE checklist for reporting valuation studies of multi-attribute utility-based instruments [29].

2.1 Participant Recruitment

Target sample size was 1200 interviews to achieve a minimum of 1000 high-quality interviews as stated in the EQ-VT to ensure consistent models for analyses [30]. To reach the target sample size, Statistics Denmark provided contact information on a randomly chosen representative sample of the Danish population with regards to age (> 18 years), gender, education, and geographical region. Statistics Denmark collects comprehensive statistical information on all Danes based on the use of the unique personal registration number and registers on the use of health and social services. Statistics Denmark also provides services for public administration and research [31]. Information on personal registration number, age, gender, education, and geographical region was provided on 4585 individuals divided into blocks of approximately 500, where each block met the requirements for representativeness. Using the personal registration number, individuals were sent a personal letter of invitation to their secure national digital mailbox linked to the personal registration number [32]. Initially invitations were sent to five randomly chosen blocks, i.e. approximately 2500 individuals, followed by invitations sent to the next block chosen randomly until all 4585 individuals had been invited. Statistics Denmark also provided information on residence, and individuals could choose to be interviewed at their own residence or a nearby public institution. To boost participation, reminders were sent to non-responders via their national digital mailbox or home address and/or they were contacted by phone. To speed up the recruitment and achieve the target sample size, a Danish market research company was included halfway through the study period. At this stage respondents were selected according to the same principles as used by Statistics Denmark to ensure the overall representativeness of the final sample. The market research company sent email invitations to their panel of survey respondents. Respondents who had not participated in a health survey within the last 6 months were asked to

answer further questions regarding age, gender, education, and geographical region to ensure representativeness of the final sample. To encourage participation, individuals were offered entry to a lottery for prizes.

2.2 The Valuation Interview

The EQ-VT version 2.1 software developed for EQ-5D-5L valuation studies was used and administered as a computer-assisted personal interview [16]. The interview comprised: (i) self-reported health using the EQ-5D-5L descriptive system and EQ VAS, (ii) questions on age, gender, and experience of serious illness, (iii) instructions and example of cTTO task, (iv) three practice cTTO tasks (mild, severe, and difficult to imagine) followed by cTTO valuation of ten EQ-5D-5L health states, (v) cTTO feedback module allowing respondents to identify states not ranked in the desired order and cTTO debriefing, (vi) DCE instructions, (vii) DCE valuation of seven pairs of EQ-5D-5L health states, (viii) DCE debriefing, (ix) experimental DCE valuation task (reported separately), and (x) questions on attitudes towards prioritisation in the Danish healthcare system (reported separately).

2.2.1 Techniques for Eliciting Preferences

In the TTO exercise, the respondent was first asked to choose between living 10 years in full health (x) or 10 years in the EQ-5D-5L health state being valued ($t = 10$). Time in full health was then altered until the respondent considered the two options to be the same, thus establishing the value of the health state being valued (x/t) between 0 (death) and 1 (full health). If a respondent was unwilling to trade off any time in full health to avoid living in the EQ-5D-5L health state (non-trader), the value for that state was 1. When a respondent traded off all the time in perfect health and was indifferent between this immediate death and living 10 years in the EQ-5D-5L health state being valued, that state was valued 0 (equivalent to death). If a respondent, on the other hand, considered the EQ-5D-5L health state being valued to be worse than dead, meaning they would prefer immediate death, a shift was made from conventional TTO to lead-time TTO. Shifting to lead-time TTO implies that the respondent was given an additional 10 years for trading. The respondent was again asked to trade off time in full health until the point of indifference, but the two options were now 10 years in full health (x) or 10 years in full health followed by 10 years in the health state being valued. The value for the health state was ($x - 10/10$), i.e. between -1 and 0 . The values for the TTO could thus range from -1 to 1 with a 0.05 increment as the smallest tradeable time was 6 months.

In the DCE task, the respondent read two EQ-5D-5L health states shown next to each other and indicated which

state was preferred. In these pairwise comparisons, neither of the health states was logically better than the other and no information was given about the duration of the states.

2.2.2 Health States Valued

In the EQ-VT, a standardized blocked design was implemented to select the health states to be valued by the respondents, where the severity of the states included in each block was balanced [15]. In the cTTO, 86 health states were valued divided into blocks of ten health states. Each group included one of the five “mild” EQ-5D-5L health states (four dimensions at level 1 and one dimension at level 2, e.g., 11112), eight “moderate” health states, and the most severe health state (i.e., 55555). Respondents were randomly assigned by the EQ-VT to one block of health states, and the order of the health states being valued was likewise randomized. In the DCE, 196 pairs of health states were valued, divided into 28 blocks of seven pairs that were similar in terms of level sum score. Respondents were randomly assigned to one of the 28 blocks by the EQ-VT. The order in which pairs were valued was randomized, as was the left-right positioning.

2.3 Data Quality

The interviewers had a master’s degree in either public health or medical market access and underwent an intensive 2.5 days of training prior to data collection. The EQ-VT quality control (QC) tool was used to monitor the quality of the collected data and to identify any interviewers performing poorly [33]. The QC tool focuses on protocol compliance of the interviewers and face validity of the collected data. An interview was flagged as being of potential poor quality if any of four cTTO indicators were observed: (i) no explanation of “worse than dead” task in the example, (ii) under 3 min spent on the cTTO example, (iii) logical inconsistency (state 55555 valued at least 0.5 higher than the lowest rated health state), and (iv) under 5 min spent on the ten cTTO tasks [16]. If four or more of an interviewer’s first ten interviews were flagged, the interviewer was asked to repeat training and the interviews were dropped. If the interviewer continued to perform poorly, they and all their interviews were dropped from the study.

Fortnightly quality reports on the protocol compliance and face validity were created from the QC tool and discussed with the EQ-VT support team to provide individual feedback to the interviewers [34].

2.4 Ethics

The Danish EQ-5D-5L valuation study is registered under Aalborg University with the Danish Data Protection Agency

(case number: 2017-899/10-0164). According to the Danish National Committee on Health Research Ethics, interview studies do not require approval (Committee Act §14, Sect. 2). Respondents received written and oral information about the study, including that it was voluntary to participate and that they could withdraw their consent at any time.

2.5 Statistical Analyses

Descriptive statistics were used to compare characteristics of the final sample with those of the adult Danish general population and to summarize self-reported health. cTTO valuations are reported as means and standard deviations (SDs).

Respondents not contributing with both cTTO and DCE data were dropped. Prior to the main modelling analysis, cTTO data for health states identified by respondents in the feedback module as not being ranked appropriately were dropped. No exclusions were made due to logical inconsistencies between EQ-5D-5L health states or non-trading. Analyses were conducted in Stata version 16.1.

2.5.1 Data Modelling

As only 86 EQ-5D-5L health states were valued directly, modelling was used to estimate values for all possible 3,125 health states. Modelling was conducted for cTTO data alone, DCE data alone, and a combination of cTTO and DCE data. As the EQ-VT was designed for maximum power to identify main effects, no interaction effects were included or investigated [30].

Two models were tested for the cTTO data: (i) a generalized least squares (GLS) random intercept model without censoring, and (ii) a random-effects Tobit model. The Tobit model takes explicit account of the censoring feature of the cTTO data that is due to the construction of the EQ-VT, where the observed values are censored at -1 [35]. Thus, from a conceptual point of view, the Tobit model is preferred to the GLS model and the Tobit model is the preferred choice in the most recent literature [28, 36, 37]. Both the Tobit model and the random effects part of the GLS model deal with another main feature of cTTO data, namely heteroscedasticity. Heteroscedasticity refers to the substantial variation among respondents regarding the valuation of health states, which tends to be more prominent for moderate and severe health states [24].

The McFadden conditional logit model is typically the preferred choice for DCE data [38]. However, parameter estimates from DCE data are not directly comparable to those from cTTO data as DCE data are not anchored on a 0–1 scale. Therefore, a conditional logit model was used to model DCE data with the scaling issue addressed by using the multiplicative constant from the hybrid model [39]. As a robustness check, a heteroscedastic conditional

logit model with heteroscedasticity being a function of observables was also estimated [40].

To explore whether modelling was improved by combining cTTO and DCE data in a hybrid model, a DCE conditional logit model was used as a building block with (i) the GLS random intercept model (=hybrid GLS and heteroscedasticity model), and (ii) the random-effects Tobit model (=heteroscedastic censored Tobit hybrid model) [39]. The key assumption behind the hybrid model is that the parameter vector from the analysis of cTTO data, β , equals the parameter vector from the analysis of DCE data, β' , up to a multiplicative constant, $\beta = \beta' \cdot \theta$. This assumption is assessed via plots of the predicted values of the health states from the cTTO data based on the estimated random-effects Tobit parameters and the conditional logit parameters. If the plots show a straight line, this supports the key assumption. Heterogeneity is accounted for in the hybrid model by letting the scale parameter, θ , be a function of the explanatory variables. Further details of the hybrid model are available in Ramos-Goñi et al. [39, 41].

Model performance was evaluated by (i) logical consistency where the absolute value of parameters associated with logically worse dimension levels must be higher than those associated with logically better levels, and (ii) goodness of fit for comparable model types if required. Traditional methods for comparing statistical models, i.e. Akaike information criterion (AIC) and Bayesian information criterion (BIC), were not viable as the log-likelihood of the hybrid model was larger than its constituent parts from the random-effects Tobit model and conditional logit model. Furthermore, use of recently popularized methods such as mean squared error or mean absolute error is not warranted for the hybrid model due to lack of supporting evidence [24].

2.5.2 Sensitivity Analyses

In line with recent reporting practices [24, 36], the robustness of the results was tested by repeating the modelling analyses after reintroducing the cTTO data for states that respondents had identified in the cTTO feedback module as inappropriately ranked.

2.6 Comparison of Value Sets

The characteristics of the Danish EQ-5D-5L value set were compared with those of the Danish EQ-5D-3L value set [5] and those of the Danish crosswalk value set, which was derived from a mapping procedure on pooled 3L and 5L data from six countries including Denmark [8, 9].

3 Results

Between October 2018 and November 2019, 1052 interviews were carried out. None of 13 interviewers performed poorly, but two were asked to leave as they were not sufficiently available for interviewing and their interviews were dropped ($n = 5$). Twelve further interviews were dropped due to software issues or respondents withdrawing consent or having cognitive/emotional issues. Participants not contributing both cTTO and DCE data ($n = 21$) were dropped, leaving 1014 interviews for inclusion.

3.1 The Sample

The sample was similar to the adult Danish general population on gender, age (slight under-representation of 18- to 24-year-olds and over-representation of 65- to 74-year-olds), marital status, and geographical region (Table 1). The sample had slightly more respondents with higher education than in the general population. Most respondents rated their own health as very good or excellent (64%), and under 10% had less good or poor health (Online Supplementary Material (OSM) 1). About half (49%) reported pain or discomfort, and 25–27% had problems with mobility or usual activities. Mean self-reported EQ VAS was 82.4 (SD 15.9) (OSM 1).

3.2 cTTO Data and Models

Each of the 1014 respondents valued ten health states with cTTO, providing 10,140 observations in total. All respondents assessed the most severe EQ-5D-5L state (55555), while the “mild” states had 195–214 evaluations (average 202.8), and the 80 “moderate” states had 96–111 evaluations (average 101.4). Descriptive statistics for the values for the 86 health states are given in OSM 2.

Figure 1 shows that mean cTTO values decreased with increasing health state severity as expected, with data heteroscedasticity reflected in higher standard deviations with greater severity. Observed cTTO values ranged from 1 to –1, and 22% of states were considered worse than death (Fig. 2).

Removal of health states identified by respondents as being incorrectly ranked ($n = 712$) gave 9428 observations for cTTO modelling. The GLS model (OSM 3) and random-effects Tobit-based model (Table 2) gave comparable results, but the Tobit model generally produced parameter estimates with slightly lower variance. In the Tobit model, the parameter estimate for mobility level 3 was inconsistent but not significantly different from the estimate for mobility level 2. The parameter estimates for self-care level 2, self-care level 3, and usual activities level 5 were not significantly different from the preceding level.

Table 1 Characteristics of the study sample compared to the Danish adult general population

| Characteristics | Study sample | | Danish adult general population ^a |
|--|--------------|------|--|
| | $(n = 1014)$ | | |
| | <i>N</i> | % | % |
| Gender | | | |
| Female | 523 | 51.6 | 50.6 |
| Male | 491 | 48.4 | 49.4 |
| Age group (years) | | | |
| 18–24 | 40 | 3.9 | 11.4 |
| 25–34 | 136 | 13.4 | 15.8 |
| 35–44 | 135 | 13.3 | 15.3 |
| 45–54 | 198 | 19.5 | 17.6 |
| 55–64 | 187 | 18.4 | 15.4 |
| 65–74 | 219 | 21.6 | 14.1 |
| 75+ | 99 | 9.8 | 10.5 |
| Marital status ($n = 1013$) | | | |
| Widowed | 48 | 4.7 | 5.9 |
| Divorced | 107 | 10.6 | 12.4 |
| Married | 509 | 50.2 | 48.0 |
| Unmarried | 349 | 34.4 | 33.7 |
| Highest education ($n = 1010$) | | | |
| Secondary school | 82 | 8.1 | 25.9 |
| High school/other | 67 | 6.6 | 12.0 |
| Skilled worker | 277 | 27.4 | 29.9 |
| Short-cycle higher education | 126 | 12.5 | 4.9 |
| Medium-cycle higher education | 279 | 27.6 | 16.9 |
| Long-cycle higher education | 179 | 17.7 | 10.3 |
| Work status ($n = 1010$) | | | |
| Employed/self-employed | 546 | 54.0 | 60.1 |
| Unemployed (able to work) | 44 | 4.4 | 2.1 |
| Outside the working force (e.g., retired, student) | 420 | 33.5 | 37.8 |
| Annual income ($n = 1010$) | | | |
| Under DKK 299,999 | 367 | 36.3 | 54.6 |
| DKK 300,000–499,999 | 426 | 42.2 | 30.9 |
| Over DKK 500,000 | 155 | 15.3 | 14.4 |
| Declined to answer | 28 | 2.8 | ^b |
| Did not know | 34 | 3.4 | ^b |
| Geographical region | | | |
| The North Denmark Region | 152 | 15.0 | 11.1 |
| Central Denmark Region | 251 | 24.7 | 24.2 |
| The Region of Southern Denmark | 197 | 19.4 | 20.4 |
| The Capital Region of Denmark | 282 | 27.8 | 31.2 |
| Region Zealand | 132 | 13.0 | 13.0 |

IQR interquartile range, *SD* standard deviation

^aFrom Statistics Denmark 2018

^bFull information provided by Statistics Denmark

To additionally assess the issues regarding heteroscedasticity, an interval regression was fitted to the data. This model took censoring into account and allowed for heteroscedasticity to be specified as a function of observables. The result was several inconsistencies that were significant. The random-effects Tobit model was therefore taken forward.

3.3 DCE Data and Models

Each of the 1014 respondents valued seven choice pairs resulting in 7098 observations. No additional exclusion criteria were applied for the DCE data. The conditional logit model gave inconsistent parameters for self-care level 3 and usual activity level 3 that were not significantly different from the preceding level (OSM 3).

The heteroscedastic conditional logit model gave similar results, with inconsistent parameters not significantly different from the preceding level (data not shown). As this model did not add more information, it was not taken further, and the simpler conditional logit model was preferred.

Scatter plots showed strong correlations between the predicted values for the 86 health states from the random-effects Tobit model and the conditional logit model (Fig. 3) indicating similar rank orderings of health states and supporting investigation into a hybrid model.

3.4 cTTO and DCE Hybrid Models

The combination of cTTO and DCE data in the heteroscedastic censored hybrid model removed the inconsistent parameter estimates present in the individual models (Table 2). Thus, all the parameter estimates were consistent, although the estimates for mobility level 3, self-care level 3, and usual activities level 3 were not significantly different from the estimate for the preceding level.

The hybrid model without censoring gave similar results to the heteroscedastic censored hybrid model, but the parameter estimates had slightly higher variance and a further level (usual activities level 5) had an insignificant parameter estimate (OSM 3). The heteroscedastic censored hybrid model

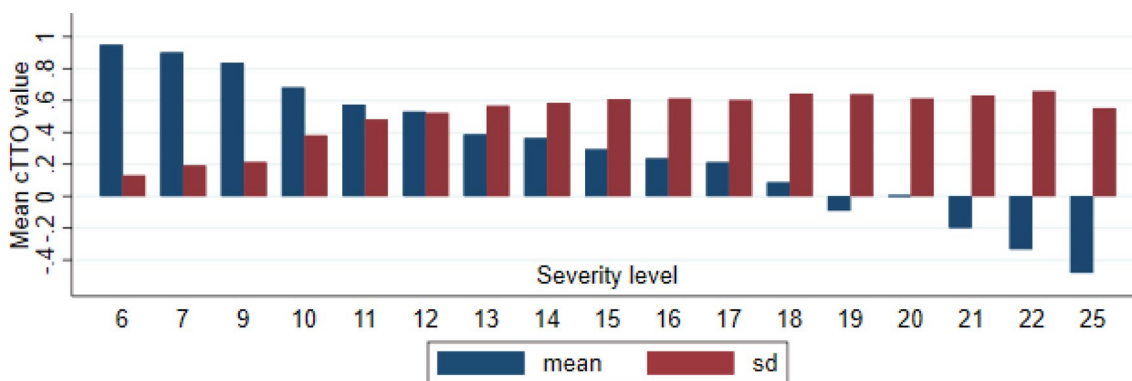


Fig. 1 Distribution of mean observed cTTO value (n = 10,140) by severity level of the health state. Severity level is calculated as the sum score of the dimension levels (e.g., health state 11112 gives a

severity level of 1 + 1 + 1 + 1 + 2 = 6). cTTO composite time trade-off, sd standard deviation

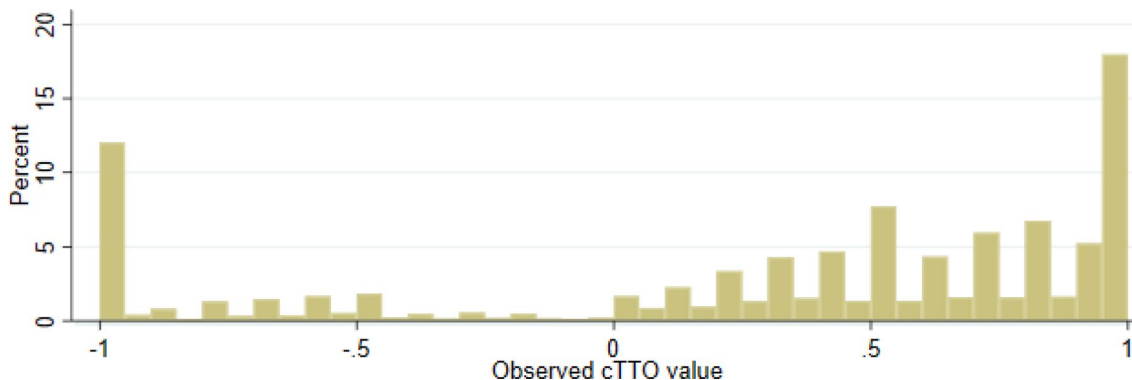


Fig. 2 Distribution (%) of observed cTTO values (n = 10,140) ranging from 1 (representing full health) to 0 (dead) and - 1 (representing states considered worse than dead). cTTO composite time trade-off

Table 2 Results for the random-effects Tobit model for composite time trade-off (cTTO) data and for the hybrid model based on cTTO plus discrete choice experiment (DCE) data. Beta coefficients should be read as utility decrements in the calculation of health-related quality of life

| Dimension level | Random-effects Tobit model | | Heteroscedastic censored hybrid model | | | | | |
|-----------------|----------------------------|---------------------|---------------------------------------|---------------------|---------|---------------------|-------|-----------------|
| | β | [95 % CI] | β | Incremental dummies | P value | Incremental dummies | SE | [95 % CI] |
| mo2 | 0.080 | [0.0499-0.1095] | 0.041 | 0.000 | 0.000 | 0.041 | 0.007 | [0.0265-0.0553] |
| mo3 | 0.061 | [0.0294-0.0921] | 0.013 | 0.281 | 0.281 | 0.054 | 0.012 | [0.0311-0.0765] |
| mo4 | 0.166 | [0.1315-0.2000] | 0.103 | 0.000 | 0.000 | 0.157 | 0.011 | [0.1347-0.1788] |
| mo5 | 0.205 | [0.1740-0.2368] | 0.063 | 0.000 | 0.000 | 0.220 | 0.011 | [0.1986-0.2415] |
| sc2 | 0.030 | [-0.0001-0.0591] | 0.035 | 0.000 | 0.000 | 0.035 | 0.007 | [0.0216-0.0479] |
| sc3 | 0.045 | [0.0113-0.0783] | 0.015 | 0.179 | 0.179 | 0.050 | 0.011 | [0.0288-0.0704] |
| sc4 | 0.111 | [0.0776-0.1448] | 0.094 | 0.000 | 0.000 | 0.144 | 0.011 | [0.1215-0.1656] |
| sc5 | 0.172 | [0.1415-0.2020] | 0.065 | 0.000 | 0.000 | 0.209 | 0.010 | [0.1886-0.2291] |
| ua2 | 0.045 | [0.0148-0.0762] | 0.033 | 0.000 | 0.000 | 0.033 | 0.007 | [0.0194-0.0471] |
| ua3 | 0.085 | [0.0516-0.1179] | 0.007 | 0.531 | 0.531 | 0.040 | 0.011 | [0.0195-0.0608] |
| ua4 | 0.153 | [0.1204-0.1860] | 0.099 | 0.000 | 0.000 | 0.139 | 0.010 | [0.1187-0.1592] |
| ua5 | 0.172 | [0.1410-0.2021] | 0.035 | 0.007 | 0.007 | 0.174 | 0.011 | [0.1526-0.1951] |
| pd2 | 0.038 | [0.0106-0.0661] | 0.048 | 0.000 | 0.000 | 0.048 | 0.006 | [0.0361-0.0606] |
| pd3 | 0.077 | [0.0430-0.1106] | 0.045 | 0.000 | 0.000 | 0.094 | 0.011 | [0.0718-0.1154] |
| pd4 | 0.363 | [0.3330-0.3934] | 0.287 | 0.000 | 0.000 | 0.381 | 0.012 | [0.3576-0.4044] |
| pd5 | 0.527 | [0.4947-0.5597] | 0.156 | 0.000 | 0.000 | 0.537 | 0.013 | [0.5109-0.5635] |
| ad2 | 0.084 | [0.0520-0.1158] | 0.072 | 0.000 | 0.000 | 0.072 | 0.007 | [0.0589-0.0851] |
| ad3 | 0.209 | [0.1727-0.2449] | 0.119 | 0.000 | 0.000 | 0.191 | 0.011 | [0.1699-0.2121] |
| ad4 | 0.430 | [0.3969-0.4630] | 0.239 | 0.000 | 0.000 | 0.430 | 0.011 | [0.4081-0.4512] |
| ad5 | 0.566 | [0.5359-0.5970] | 0.188 | 0.000 | 0.000 | 0.618 | 0.013 | [0.5935-0.6428] |
| Constant | -0.004 | [-0.0451 to 0.0364] | Not used | | | Not used | | |

ad anxiety/depression, CI confidence interval, mo mobility, pd pain/discomfort, sc self-care, SE standard error, ua usual activity

^aPreferred model

was thus the best model, with no logical inconsistencies. A scatter plot showed strong correlation with the cTTO and DCE models (Fig. 3).

3.5 Sensitivity Analyses

When the 712 states that were identified by respondents as being incorrectly ranked on cTTO were re-introduced, the model estimates were qualitatively unchanged (data not shown). As it was preferred to accept respondents' judgments about health states that were incorrectly ranked, the final model did not include the data for these states.

3.6 The Final Model for the Danish 5L Value Set

The heteroscedastic censored hybrid model combining cTTO and DCE data was chosen for modelling the final Danish value set (Table 2). The parameter estimates represent the utility decrements associated with each EQ-5D-5L dimension level and allow a value (utility) to be assigned to each of the 3125 health states, for example state 13224 has the utility, $U = 1 - 0 - 0.050 - 0.033 - 0.048 - 0.430 = 0.439$. The lowest score in the Danish value set is for health state 55555

at -0.757 . Respondents placed most weight on EQ-5D dimensions of anxiety/depression and pain/discomfort when expressing their preferences for the different health states.

3.7 Comparison of Value Sets

The 5L value set had a lower value for the worst possible health state (55555) compared to EQ-5D-3L and crosswalk value sets [5, 9] (Table 3). The 3L and 5L value sets had similar proportions of states worse than death (20–22%) compared to the crosswalk value set with 11%. The largest utility decrement in the 3L value set was for mobility followed by pain/discomfort and anxiety/depression, whereas the largest decrement in the 5L value set was for anxiety/depression followed by pain/discomfort and then mobility. In the 3L value set, the utility decrement of 0.411 for mobility level 3 was only slightly higher than the 0.396 for pain/discomfort level 3 [5]. In comparison, the 5L value set showed a substantial preference difference between these two dimensions where the decrement for pain/discomfort level 5 was 0.537 compared to 0.220 for mobility level 5.

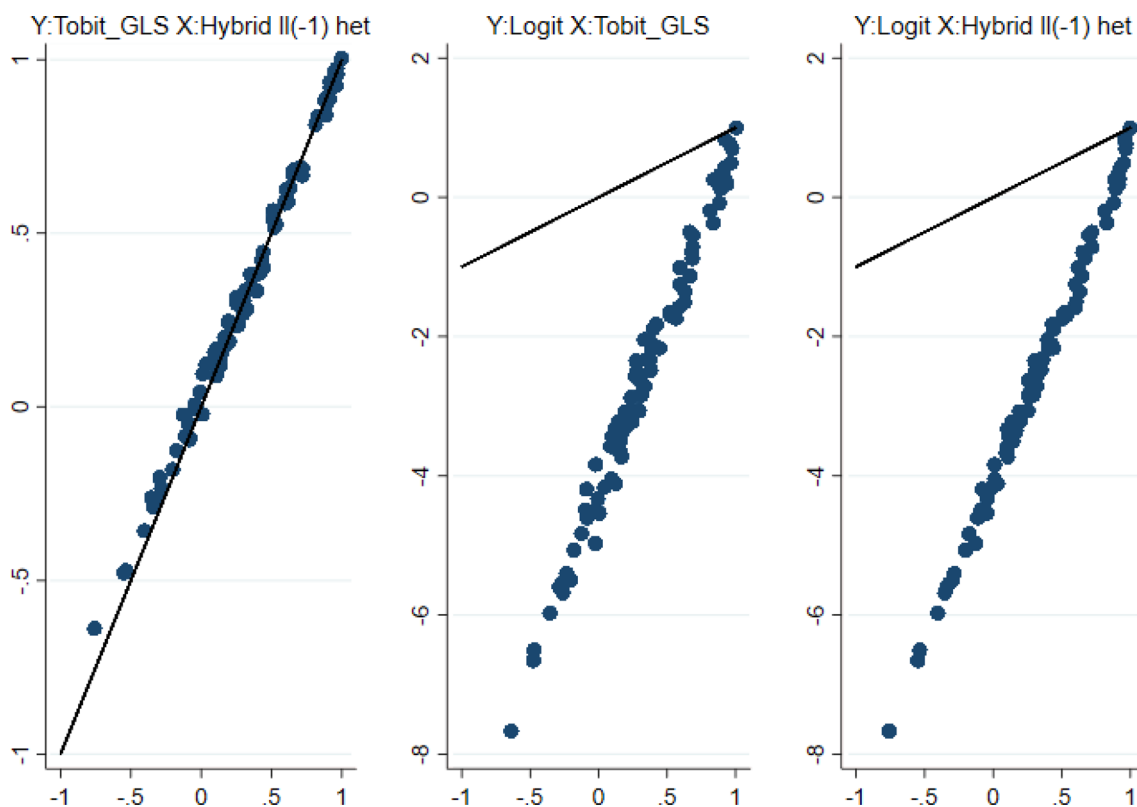


Fig. 3 Scatter plot of the predicted values for the 86 health states valued in cTTO using a random-effects Tobit model (cTTO data), the logit model (DCE data), and heteroscedastic hybrid model

(cTTO and DCE data). *cTTO* composite time trade-off, *DCE* discrete choice experiment

4 Discussion

This study reports the development of the Danish EQ-5D-5L value set based on preferences from the adult Danish general population using cTTO and DCE. A heteroscedastic censored hybrid model using both cTTO and DCE data was found to be the best approach for generating the Danish EQ-5D-5L value set.

Particular strengths of this study were the rigid adherence to the updated EQ-VT protocol version 2.1 [16] and the collaboration with Statistics Denmark. Compared to earlier valuation studies, this collaboration improved the sample representativity and represents a novel approach to sample selection and evaluation of representativity. Statistics Denmark provided precise knowledge of the distribution of age, gender, marital status, geographical region, and education needed for a representative sample of the Danish population above 18 years of age, and this was used to guide the recruitment of participants. A limitation of the study was the need to change recruitment source (i.e., from Statistics Denmark to a Danish market research company) for achieving the final sample numbers because recruitment was going too slowly. However, both strategies were governed by the statistical information from Statistics Denmark on the requirements for a representative sample. We were aware of potential differences between respondents randomly chosen by Statistics Denmark and those who were in the market research company's panel. To ensure that no "professional" respondents were recruited from the panel, respondents from the market research company were not eligible for participation if they had participated in a health survey within the last 6 months.

It was expected that it would be difficult to recruit participants for the study as the interview was of 1.5–2 h duration (due to extra questions on DCE and prioritisation, which is reported elsewhere) and there was no direct payment to participants, only an opportunity to enter a lottery for prizes. Furthermore, interviews were to be carried out across the country to ensure representativeness of the final sample. As the interviewers all lived in a rather small area in and around the North Denmark Region, the number of days available for interviews in each of the five different regions of Denmark was limited. It might have eased recruitment if an interview team had been available in each of the five regions, giving more time slots for interviews in each region. However, it was prioritized to have only one interview team working closely together to limit any interviewer effects.

The final sample of individuals showed a good representation of the general Danish population except for slight under-representation of individuals aged 18–24 years and of individuals with the lowest educational level. Under-representation of individuals with lower educational level has been demonstrated in other 5L valuation studies [26, 42–44] and occurred despite our best efforts during data collection.

The present study is characterized by high-quality cTTO data according to EQ-VT indicators as witnessed by the QC tool. The collected data showed high protocol compliance, for example with regards to the duration of interviews both within and across interviewers, no interviewers flagged as performing poorly, and high face validity of the data. The high quality can be traced to several sources. First, the interviewers had strong theoretical and methodological competence within the field and underwent extensive training prior to data collection. This is believed to have resulted in high protocol compliance that was evident in the quality reports.

Table 3 Comparison of key characteristics of the three Danish value sets for EQ-5D-3L, the crosswalk, and EQ-5D-5L

| Characteristics | EQ-5D-3L | Crosswalk | EQ-5D-5L |
|---|---------------------------------|---------------------------------|---------------------------------------|
| Valuation method | TTO | Crosswalk (TTO) | Hybrid (cTTO/DCE) |
| Maximum value | 1 | 1 | 1 |
| Minimum value | – 0.624 ^b | – 0.624 ^c | – 0.757 ^c |
| Number of health states | 243 | 3125 | 3125 |
| Health states WTD (%) | 19.75% | 11.01% | 21.7% |
| Dimensions ordered by largest to smallest utility decrements ^a | Mobility Pain/Discomfort | Mobility Pain/Discomfort | Anxiety/Depression Pain/Discomfort |
| | Anxiety/Depression Self-care | Anxiety/Depression Self-care | Mobility Self-care |
| | Usual Activities | Usual Activities | Usual Activities |

cTTO composite time trade-off, DCE discrete choice experiments, TTO time trade-off, WTD worse than dead

^aBased on the utility decrement for the most severe level in each dimension (level 3 for EQ-5D-3L, level 5 for the crosswalk, and EQ-5D-5L)

^bHealth state 33333

^cHealth state 55555

Second, the study used the most recent EQ-VT, version 2.1 [16], which includes a “dynamic” question during the cTTO example to ensure that respondents are introduced to valuation of health states both better than dead and worse than dead. The use of the QC tool itself enables continual monitoring of the data collected and facilitated individual feedback to interviewers to ensure high performance [16, 33, 45].

A heteroscedastic censored hybrid model combining cTTO and DCE data was established as the preferred approach. As the EQ-5D-5L has been demonstrated to have improved measurement properties over the EQ-5D-3L [46], we recommend the EQ-5D-5L with this newly developed value set to be used by Danish decision-makers when estimating QALYs.

A comparison of the three Danish value sets shows that the percentage of health states valued as being worse than dead was similar in the 3L [5] (20%) and 5L (22%) value sets, which is reassuring considering the addition of DCE data and the 20-year interval between the collection of data for the 3L study and the present study. The percentage of states worse than dead was noticeably lower in the crosswalk value set [9] (11%), possibly due to this being based on a mapping algorithm [8], whereas the 3L and 5L value sets are based on directly elicited preferences from the general population.

The percentage of health states valued as worse than dead in the Danish 5L value set was comparable to that in the US [47] 5L value set (20%), but lower than that in the Indonesian [48] (35%) and Irish [26] (36%) 5L value sets and higher than that in the French [37] (13%) and Polish [25] (4.4%) 5L value sets. It is difficult to ascertain the reasons for these differences. The general difficulty of interpreting states worse than death should be kept in mind as Gandhi et al. found little association between health-state severity and negative values, and questioned the usefulness of asking people to value health states considered worse than dead [49]. Other factors may play a part, however. Purba et al. [48] suggested that the high level of collectivism in Indonesia could make people want to avoid being a burden for their family and friends, thus they would rather die than be in severe health states for any length of time. Certainly, Danish society scores much higher on individualism (score of 74) and is similar to Ireland, France, and Poland (scores of 60–71) though lower than the USA (score of 91) compared to Indonesia (score of 14) [50]. Secondly, religious beliefs seem to influence people’s preferences for health states. In a Polish study [51], respondents who believed in the afterlife tended to be non-traders (i.e., unwilling to give up any life to avoid poor health states) and were less likely to consider a state worse than death, and the later Polish 5L valuation study included a parameter that scaled down the disutilities given by religious respondents [25]. Although the French valuation study had 13% of states considered worse than

death, nearly 14% of respondents were non-traders [37]. In comparison, only seven respondents (0.7%) were non-traders in the Danish sample and approximately 22% of health states considered worse than death.

An important difference between the Danish 5L and 3L value sets was the change in the ranking of the EQ-5D dimensions based on the utility decrements. While the anxiety/depression dimension showed the largest utility decrement in the 5L value set, the mobility dimension showed the largest decrement in the 3L value set [5]. This apparent change in health preferences may affect future prioritization in the Danish healthcare sector. Reasons for the change in Danish preferences are unclear, but similar changes were noted in Poland and some other (high-income) countries by Golicki et al. [25], and may be in part due to the change in the wording of the most severe mobility level. It is possible that Danes perceive mobility issues to be less problematic than earlier due to reforms to the Disability Pension Scheme in 2003 and 2012 [52] and the introduction of “Everyday rehabilitation” services in 2015 (section 83a of the Service Act) [53], which require the Danish municipalities to offer rehabilitation and assistance to people with disabilities to allow them to lead as normal and independent a life as possible. More weight might be placed on the anxiety/depression dimension as mental disorders have received more attention in Denmark, among others politically. Treatment of mental illness, including anxiety and depression, has likewise increased in Denmark from 188 health service contacts per 1,000 people in 2009 to 246 contacts per 1,000 people in 2017 [54], reflecting an increased incidence, more treatment opportunities, and/or greater recognition and acceptance of mental illness among the general population.

Comparing 5L value sets from other countries shows that the ranking in the Danish value set (anxiety/depression, pain/discomfort, mobility, self-care, usual activities) is identical to that in Ireland [26], and the first three dimensions are identical to those in the Ethiopian [55] 5L value set. In Poland [25], Portugal [24], the USA [47], and France [37], pain/discomfort has the greatest utility decrement, followed by mobility and then anxiety/depression as the third or fourth. The usual activities dimension has the smallest utility decrement in six of eight 5L value sets compared here including the Danish set; exceptions are the Ethiopian [55] and the Indonesian [48] 5L value sets, where self-care and pain/discomfort have the smallest utility decrement, respectively. Cross-national differences in the ranking of dimensions might limit the transferability of value sets across countries, suggesting the relevance of country-specific valuation studies. Again, it is difficult to be certain about the reasons for these differences. The 5L valuation studies that used the EQ-VT approach are less likely to have methodological differences as suggested for 3L studies, despite some similarities among Northern European countries including

Denmark [56]. Purba et al. [48] noted possible translation effects where the Indonesian word used for “depression” could be interpreted more as “sadness”, and this is possibly perceived as less severe than the Danish word, which lies further along the continuum towards clinical depression. In addition, mental health has received increased attention in recent years in Denmark with the launching in 2018 of a nationwide programme to provide earlier treatment of psychiatric disorders, free psychological help for mild mental illness, and individually tailored outpatient treatments [57].

The current study builds on recent interest in utilising different types of data for valuation studies and, in line with several recently derived value sets, a hybrid model combining cTTO and DCE data was identified as most appropriate for the Danish value set [24, 27, 28, 36]. DCE thus appears to contribute a different type of information than TTO, but it may not be easier to understand and answer than TTO [58], and the problem of 0–1 anchoring of DCE values is still to be resolved. One approach may be to incorporate duration into the DCE choices [59, 60].

A hybrid model was chosen here from a statistical point of view, but it is also important to review the utility theory models for TTO and DCE and ask how well the hybrid model reflects the underlying theoretical foundations. McFadden’s conditional logit model is typically the preferred choice for DCE random utility models [38], but the parameter estimates from DCE and cTTO are not directly comparable as DCE data are not anchored on a 0–1 scale. A conditional logit model was thus used to model DCE data with the scaling issue addressed by using the multiplicative constant from the hybrid model. This is an example of statistical convenience. It should be investigated how well the underlying theoretical model, for example the Hicksian utility model for TTO [61, 62], corresponds to the econometric specification and vice versa. One of the few attempts to provide a utility theoretical basis for the hybrid model is the episodic random utility model that unifies TTO and DCE approaches [63]. A variety of econometric modelling approaches are available for modelling preference data, and choice of the “right” model should be based on both statistical and theoretical properties. Devlin has noted that the choice of valuation method has a non-trivial impact on quality-of-life utilities and cannot be determined with recourse to statistical properties alone. Or, in other words, theory matters a lot. This holds even more when TTO and DCE are used together [64]. More research should be done on the utility theoretical foundations – even before new statistical are being introduced.

5 Conclusions

A heteroscedastic censored hybrid model using both cTTO and DCE data was identified as the best approach for generating the Danish EQ-5D-5L value set. A high-quality data set was achieved from a representative sample of the adult Danish general population, which is important for real-world use in a priority-setting context for, among other things, hospital-dispensed medicines. The study results emphasize the importance of a standard and detailed interview protocol, extensive interviewer training, and quality control during data collection.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40258-021-00639-3>.

Acknowledgements The authors thank the interviewers and respondents for making this research possible. We are also grateful for the contributions from the EQ-VT support team as well as Nancy Raakjær and Anita Egholm Jensen at Aalborg University.

Declarations

Funding The Danish EQ-5D-5L valuation study has received funding from the EuroQol Research Foundation, The Danish Health Foundation (Helsefonden), Aase og Ejnar Danielsens Fond, the North Denmark Region, and the Department of Business and Management at Aalborg University. Funding for Open Access was provided by the Department of Clinical Medicine, Aalborg University.

Conflicts of interest There are no conflicts of interests. Claire Gudex is a member of the EuroQol Research Foundation (the copyright holders of the EQ-5D-5L).

Ethics approval Ethics approval was not required for the interview study according to the Danish National Committee on Health Research Ethics (Committee Act §14, section 2). The study was registered under Aalborg University with the Danish Data Protection Agency (case number: 2017-899/10-0164).

Consent to participate Informed consent was obtained from all participants included in this study.

Consent for publication Not applicable.

Availability of data and material/Code availability Available from the corresponding author on reasonable request. Coding is likewise available on reasonable request.

Author contributions All authors contributed to the study conception and design as well as material preparation. CEJ was the project manager and contributed to the Danish translation of the training material, trained the interviewers, managed the data collection, was responsible for the bi-monthly quality control reports, and contributed to the data analyses. SSS and CG tested the training material and trained interviewers. SSS also contributed to data collection. MBJ performed the data analyses and discussed these with the EQ-VT support team. All authors wrote and revised all versions of the manuscript and all authors read and approved the final version of the manuscript.

Disclaimer statement The views expressed are those of the authors and not necessarily the views of the EuroQol Group.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- 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. *Eur J Heal Econ.* 2020;21:1245–57.
- Wisløff T, Hagen G, Hamidi V, Movik E, Klemp M, Olsen JA. Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *Pharmacoeconomics.* 2014;32(4):367–75.
- Medicinerådet. Danish Medicines Council. <https://medicinraadet.dk/om-os/in-english>. Accessed 30 Nov 2020.
- Medicinerådet. [In Danish] Nye metoder er på plads: Fra 1. januar 2021 skal Medicinerådet bruge QALY. 2020. <https://medicinraadet.dk/nyheder/2020/nye-metoder-er-pa-plads-fra-1-januar-2021-skal-medicinraadet-bruge-qaly>. Accessed 30 Nov 2020.
- Wittrup-Jensen K, Lauridsen J, Gudex C, Pedersen K. Generation of a Danish TTO value set for EQ-5D health states. *Scand J Public Heal.* 2009;37:459–66.
- Pedersen KM, Wittrup-Jensen K, Brooks R, Gudex C. Værdisætning af sundhed. Teorien om kvalitetsjusterede leveår og en dansk anvendelse [in Danish]. 2nd ed. Odense: Syddansk Universitetsforlag; 2006.
- EuroQol Research Foundation. EQ-5D-3L User Guide, 2018. <https://euroqol.org/publications/user-guides>.
- Van Hout B, Janssen MF, Feng YS, 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 Health.* 2012;15:708–15.
- EuroQol Research Foundation. EQ-5D-5L | Valuation | Crosswalk Index Value Calculator. <https://euroqol.org/eq-5d-instrument/s/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/>. Accessed Jun 29 2020.
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20:1727–36.
- EuroQol Research Foundation. EQ-5D-5L User Guide, 2019. <https://euroqol.org/publications/user-guides/>. Accessed 30 Nov 2020.
- Medicinerådet. [In Danish] Medicinerådets metodevejledning for vurdering af nye lægemidler. Version 1. 2020. https://medicinraadet.dk/media/5nvpk03/efter-1-januar-2021_medicinradets-metodevejledning-for-vurdering-af-nye-laegemidler-vers-1-0_adlegacy.pdf. Accessed 30 Nov 2020.
- Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res.* 2013;22:1717–27.
- Janssen BMF, Oppe M, Versteegh MM, Stolk EA. Introducing the composite time trade-off: a test of feasibility and face validity. *Eur J Heal Econ.* 2013;14:S5–13.
- Oppe M, Devlin NJ, Van Hout B, Krabbe PFM, De Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Heal.* 2014;17:445–53.
- 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 Heal.* 2019;22:23–30.
- Stolk EA, Oppe M, Scalone L, Krabbe PFM. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value Heal.* 2010;13(8):1005–13.
- Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. *J Health Econ.* 2012;31:306–18.
- Norman R, Mulhern B, Viney R. The impact of different DCE-based approaches when anchoring utility scores. *Pharmacoeconomics.* 2016;34:805–14.
- Mulhern B, Bansback N, Hole AR, Tsuchiya A. using discrete choice experiments with duration to model EQ-5D-5L health state preferences: testing experimental design strategies. *Med Decis Mak.* 2017;37:285–97.
- Rowen D, Brazier J, Van Hout B. A comparison of methods for converting DCE values onto the full health-dead QALY scale. *Med Decis Mak.* 2015;35:328–40.
- Shah KK, Ramos-Goñi JM, Kreimeier S, Devlin NJ. An exploration of methods for obtaining 0 = dead anchors for latent scale EQ-5D-Y values. *Eur J Heal Econ.* 2020;21:1091–103.
- Webb EJD, O'Dwyer J, Meads D, Kind P, Wright P. Transforming discrete choice experiment latent scale values for EQ-5D-3L using the visual analogue scale. *Eur J Heal Econ.* 2020;21:787–800.
- 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. *Qual Life Res.* 2019;28:3163–75.
- 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;37:1165–76.
- Hobbins A, Barry L, Kelleher D, Shah K, Devlin N, Goni JMR, et al. Utility values for health states in Ireland: a value set for the EQ-5D-5L. *Pharmacoeconomics.* 2018;36:1345–53.
- Pattanaphesaj J, Thavorncharoensap M, Ramos-Goñi JM, Tongsiri S, Ingsrisawang L, Teerawattananon Y. The EQ-5D-5L Valuation study in Thailand. *Expert Rev Pharmacoecon Outcomes Res.* 2018;18:551–8.
- Lin HW, Li CI, Lin FJ, Chang JY, Gau CS, Luo N, et al. Valuation of the EQ-5D-5L in Taiwan. *PLoS ONE.* 2018;13:1–16.
- Xie F, Pickard AS, Krabbe PFM, Revicki D, Viney R, Devlin N, et al. A checklist for reporting valuation studies of multi-attribute utility-based instruments (CREATE). *Pharmacoeconomics.* 2015;33:867–77.
- Oppe M, van Hout B. The, “power” of eliciting EQ-5D-5L values: the experimental design of the EQ-VT. *EuroQol Work Paper Ser.* 2017;17003:1–17.
- Thygesen LC, Daasnes C, Thaulow I, Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: Structure, access, legislation, and archiving. *Scand J Public Health.* 2011;39:12–6.

32. e-Boks. What is e-Boks? <https://www.e-boks.com/danmark/en/what-is-e-boks/>. Accessed 26 Mar 2020.
33. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJV, Stolk E. Quality control process for EQ-5D-5L valuation studies. *Value Heal*. 2017;20:466–73.
34. Purba FD, Hunfeld JAM, Iskandarsyah A, Fitriana TS, Sadarjoen SS, Passchier J, et al. Employing quality control and feedback to the EQ-5D-5L valuation protocol to improve the quality of data collection. *Qual Life Res*. 2017;26:1197–208.
35. Devlin N, Buckingham K, Shah K, Tsuchiya A, Tilling C, Wilkinson G, et al. A comparison of alternative variants of the lead and lag time TTO. *Health Econ*. 2013;22:517–32.
36. Ludwig K, von der Schulenburg GJM, Greiner W. German value set for the EQ-5D-5L. *Pharmacoeconomics*. 2018;36:663–74.
37. Andrade LF, Ludwig K, Goni JMR, Oppe M, de Pouvourville G. A French value set for the EQ-5D-5L. *Pharmacoeconomics*. 2020;38:413–25.
38. McFadden D. Conditional logit analysis of qualitative choice behaviour. In: Zarembka P, editor. *Frontiers in econometrics*. New York: Academic Press; 1974. p. 105–42.
39. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care*. 2017;55:e51–8.
40. Hole AR. Small-sample properties of tests for heteroscedasticity in the conditional logit model. *Econ Bull*. 2006;3:1–14.
41. Ramos-Goñi JM, Craig B, Oppe M, Van Hout B. Combining continuous and dichotomous responses in a hybrid model. *EuroQol Work Pap Ser*. 2016;16002:2–15.
42. Augustovski F, Rey-Ares L, Irazola V, Garay OU, Gianneo O, Fernández G, et al. An EQ-5D-5L value set based on Uruguayan population preferences. *Qual Life Res*. 2016;25:323–33.
43. Xie, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care*. 2016;54:98–105.
44. Versteegh M, Vermeulen MK, Evers MAAS, de Wit GA, Prenger R, Stolk AE. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19:343–52.
45. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goñi JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. 2016;34:993–1004.
46. Buchholz I, Janssen MF, Kohlmann T, Feng Y-S. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *Pharmacoeconomics*. 2018;36:645–61.
47. 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 Heal*. 2019;22:931–41.
48. Purba FD, Hunfeld JAM, Iskandarsyah A, Fitriana TS, Sadarjoen SS, Ramos-Goñi JM, et al. The Indonesian EQ-5D-5L value set. *Pharmacoeconomics*. 2017;35:1153–65.
49. Gandhi M, Rand K, Luo N. Valuation of health states considered to be worse than death—an analysis of composite time trade-off data from 5 EQ-5D-5L valuation studies. *Value Heal*. 2019;22:370–6.
50. Hofstede Insights. Country Comparison. <https://www.hofstede-insights.com/country-comparison/denmark,indonesia,the-usa/>. Accessed 8 Dec 2020.
51. Jakubczyk M, Golicki D, Niewada M. The impact of a belief in life after death on health-state preferences: true difference or artifact? *Qual Life Res*. 2016;25:2997–3008.
52. Danish Agency for Labour Market and Recruitment. Reform of the Disability Pension and Flexi-job Scheme (2012). <https://www.star.dk/en/recent-labour-market-policy-reforms/reform-of-the-disability-pension-and-flexi-job-scheme-2012/>. Accessed 8 Dec 2020.
53. Sundhedsstyrelsen. [In Danish] Rehabilitering på ældreområdet. Hvad fortæller danske undersøgelser os om kommunernes arbejde med rehabilitering i hjemmeplejen. Sundhedsstyrelsen; 2017. <https://www.sst.dk/da/Feeds/~media/887D8135638442D08F8C6BBA0FA4C130.ashx>. Accessed 8 Dec 2020.
54. Sundhedsdatastyrelsen. [In Danish] Udvalgte nøgletal for det regionale sundhedsvæsen 2009-2017. 2018.
55. 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 Health Reg Issues*. 2020;22:7–14.
56. Norman R, Cronin P, Viney R, King M, Street D, Ratcliffe J. International comparisons in valuing eq-5d health states: A review and analysis. *Value Health*. 2009;12:1194–200.
57. Danske Regioner. [In Danish] Mental Sundhed—Bedre behandling til mennesker med svær psykisk sygdom. 2018
58. Xie S, Wu J, He X, Chen G, Brazier JE. Do discrete choice experiments approaches perform better than time trade-off in eliciting health state utilities? Evidence from SF-6Dv2 in China. *Value Heal*. 2020;23:1391–9.
59. Lim S, Jonker MF, Oppe M, Donkers B, Stolk E. Severity-stratified discrete choice experiment designs for health state evaluations. *Pharmacoeconomics*. 2018;36:1377–89.
60. 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? *Med Decis Mak*. 2018;38:306–18.
61. Buckingham K, Devlin N. A theoretical framework for TTO valuations of health. *Health Econ*. 2006;15:1149–54.
62. Buckingham KJ, Devlin NJ. A note on the nature of utility in time and health and implications for cost utility analysis. *Soc Sci Med*. 2009;68:362–7.
63. Craig BM, Busschbach JJV. The episodic random utility model unifies time trade-off and discrete choice approaches in health state valuation. *Popul Health Metr*. 2009;7:1–10.
64. Devlin NJ. Theoretical foundations and challenges. ISPOR Milan 2015 Work W18 Util HTA Challenges Theory Pract Now Futur 2015. <https://www.slideshare.net/OHENews/theoretical-foundations-and-challenges>