Valuing Health State: An EQ-5D-5L Value Set for Ethiopians

Abraham G. Welie, MSc[1], * Gebremedhin Beedemariam Gebretekle, MSc[2], Elly Stolk, PhD, Clara Mukuria, PhD[3], Murray D. Krahn, MD, MSc,[5] Fikre Enquoselassie, PhD,[6] Teferi Gedif Fenta, PhD[2]

[1] School of Pharmacy, College of Health Science, Mekelle University, Mekelle, Ethiopia;
[2] School of Pharmacy, College of Health Science, Addis Ababa University, Addis Ababa, Ethiopia; [3]Executive Office, EuroQol Research Foundation, Rotterdam, The Netherlands; [4]Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, England, UK; [5]Toronto Health Economics and Technology Assessment Collaborative, University of Toronto, Toronto, ON, Canada; [6]School of Public Health, College of Health Science, Addis Ababa University, Addis Ababa, Ethiopia.

* Address correspondence to: Abraham G. Welie, MSc, School of Pharmacy, College of Health Science, Mekelle University, Mekelle, Ethiopia. Email: abrishg20@ gmail.com.

Highlights

There is a growing interest in Health Technology Assessment and economic evaluation in low-resource setting such as Ethiopia.

The study developed a value set for EQ-5D-5L using an Ethiopian general population sample with the EuroQol Group- Portable Valuation Technology (EQ-PVT) protocol administered in Amharic.

The new value set provides local users with societal preferences that are relevant.

The study also established the feasibility of using the less resource-intensive EQ-PVT which is relevant for future studies.

Summary

An Ethiopian value set for EQ-5D-5L was developed using an internationally accepted protocol with values ranging from -0.718 to 1.

Abstract

Objectives: There is a growing interest in Health Technology Assessment(HTA) and economic evaluations in developing countries like Ethiopia. The objective of this study was to derive an EQ-5D-5L value set from the Ethiopian general population in order to facilitate cost utility analysis.

Methods: A nationally representative sample(n=1,050) was recruited using stratified multi-stage quota sampling technique. Face-to-face, computer-assisted interviews using the EuroQol Portable Valuation Technology(EQ-PVT) protocol of composite time trade-off(*c*-TTO) and discrete choice experiments(DCE) were undertaken to elicit preference scores. EQ-PVT protocol feasibility was pilot tested in a sample of the population(n=110). A hybrid regression model combining *c*-TTO and DCE data was used to estimate the final value set.

Results: In the pilot study, acceptability of the tasks was good and there were no special concerns with undertaking *c*-TTO and DCE task. The coefficients generated from a hybrid model were logically consistent. The predicted values for the EQ-5D-5L ranged from -0.718 to 1. Level 5 anxiety/depression had the largest impact on utility decrement(-0.458) while level 5 self-care had the least impact(-0.222). The maximum predicted value beyond full health was 0.974 for the '11112' health state.

Conclusions: This is the first EQ-5D-5L valuation study in Africa using international valuation methods(*c*-TTO and DCE) and also the first using EQ-PVT protocol to derive a value set. We expect that the availability of this value set will facilitate HTA, health-related quality of life research and inform policy decision making in Ethiopia.

Keywords: Discrete choice experiment, EQ-5D-5L, Ethiopia, Health-state valuation, Time trade-off.

INTRODUCTION

Globally, countries with publicly financed health services struggle to provide universal health coverage. The process of deciding which healthcare technologies and interventions to invest in has thus become increasingly important. This problem is greater in developing countries as they have very limited healthcare budgets[1]. As a result, there is a growing interest in Health Technology Assessment (HTA) in developing countries[2–4]. An important aspect of HTA is economic evaluation, which involves comparison of two or more interventions in terms of costs and outcomes[2,5–7]. Although economic evaluation can be undertaken using natural outcomes such as life years saved, these are limited as they only capture one aspect of health e.g. mortality. Quality adjusted life years (QALYs) combine life years (mortality) with a quality adjustment that reflects the health related quality of life (morbidity) on a 0 (dead) to 1 (full health) utility scale[8]. QALYs are the recommended health outcome in economic evaluation in guidelines from the US, Canada, UK and many European countries[9–12].

Quality adjustment for QALYs relies on preferences for different health states derived using scaling methods such as the time trade off (TTO) and standard gamble (SG)[13]. Preference elicitation can be time consuming and costly if undertaken independently for each study. Therefore, indirect preference-based instruments such as EQ-5D have been developed. These measures include a descriptive system (e.g. EQ-5D has five dimensions with three or five levels of severity) as well as a tariff which can be applied to generate utility values. The tariff is based on a valuation in a representative sample of the community using methods such as TTO. For specific measurement applications, patients can complete the descriptive system; community-derived tariffs are applied to generate the relevant utility values[14].

A number of countries have generated country-specific value sets for the EQ-5D-5L[15,16,25,17–24]. Given the variation in culture, health and social care provision, preferences are likely to vary across different populations. It is therefore preferred that economic evaluations should use locally derived value sets[14]. In the context of developing countries, the presence of local value sets could facilitate economic evaluation research and can improve the quality of regulatory, coverage and reimbursement policy decisions in individual and public healthcare[4,26–28] e.g. a recent review in Central Europe found more HTA activities in countries with EQ-5D studies although the direction of causality is not certain[26].

The Ethiopian Federal Ministry of Health (EFMOH) has introduced a public health insurance system, and aims to support health care financing with principles of HTA[29]. The current practice of HTA for regulatory and formulary development is based on expert opinion and published HTA studies from other countries[6]. Although context-specific HTA is needed, local value sets are not yet available. The aim of this study was therefore to obtain an EQ-5D-5L value set from Ethiopian general population.

Methods

Sampling method and study population

Based on a multi-country pilot study, the recommended total sample size for EQ-5D-5L valuation studies is 1,000 participants per country. The total number of health states included for *c*-TTO is 86 and 196 pairs of health state valuation for DCE task[30]. A representative sample of 1,050 respondents was recruited based on multistage stratified quota sampling of geographic area/residence (urban/rural); gender (male/female); age group (18–24/24– 54/55-64/>65 years); and religion (Christian/Islam/Others).

Respondents interviewed in this study were living in Addis Ababa city and Butajira rural area, Southern Ethiopia. These two areas were selected as they have mixed populations in terms of ethnicity and culture due to migration from other regions[31,32]. In Addis Ababa city, supervisors identified the centre of each the ten sub-cities and then randomly identified the direction in which households would be identified from this point. One member of identified households fulfilling the inclusion criteria was interviewed. Interviews were then conducted in other households in that same direction until the required quota sample was obtained. The Demographic Health Surveillance Site (DSS) list for Butajira nine rural area (n=15,000 households) was used to identify households using their unique household number. The supervisors were responsible for selecting the households from the list based on the quotas that needed to be filled. Supervisors also used personal networks from previous DSS surveys to mediate contact between the interviewers and respondents.

The inclusion criteria were: 1) being 18+ years old; 2) able to understand the task (as judged by interviewer); and 3) able to give informed consent. Participants with presence of any illness or cognitive impairment (confirmed and/or interviewer judgement) that would interfere with the study task were excluded. All respondents provided written informed consent.

Outcome Measure

The EQ-5D-5L instrument is a generic, multi-attribute utility-based health status tool developed by the European quality of life (EuroQol) Group[33]. It consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression[34]. The EQ-5D initially included 3-level response options (no problems, some/moderate problems, extreme problems/unable to/confined to bed) under each dimension, but in 2009 was expanded to 5 levels (no problems, slight, moderate, severe, extreme/unable). The EQ-5D-5L therefore includes 3,125 health states (i.e., 5^5). The change has improved the instrument's sensitivity and

reduced ceiling effects[35]. The EQ-5D-5L has been translated to Amharic, the national language used in Ethiopia, using the standardized approach recommended by the EuroQol group[14].

Valuation techniques

This study was a population-based, interviewer-administered, face-to-face, cross-sectional survey followed a standardized valuation study protocol developed specifically for EQ-5D-5L value set studies[36]. The data was collected using laptops installed with the EuroQol Portable Valuation Technology (EQ-PVT) software using two elicitation techniques: composite time trade-off (*c*-TTO) and discrete choice experiments (DCE). A combination of both *c*-TTO and DCE method have been used in previous EQ-5D valuation studies[15,18,20,23,25].

c-TTO involves separate tasks for states considered better than dead and for states worse than dead. For states better than dead, respondents are asked their preference between living in poor health for 10 years and living in full health for a shorter period. The length of time lived in full health is varied until the respondent is indifferent between the two options. For states worse than dead, participants complete a lead-time TTO where respondents choose between 10 years of full health and the alternative is 20 years, which consists of 10 years spent in full health followed by 10 years in an impaired health state[37]. After completing the TTO tasks, participants provided feedback on the ranking of the states they valued by flagging any states that were in the wrong order.

DCE involves a set of dichotomous choices over two multi-attribute poor health states. The two methods generate different and complementary preference data. TTO elicits a value for each state with 1 and 0 defined as anchor points with worse than dead bounded at -1. The DCE approach generates binary data which allow for the derivation of a scale of non-anchored

relative values[16]. DCE is an approach that is increasingly used to assess preferences for health states because of the relative simplicity of the tasks[38].

Data collection and eliciting preferences methods

An EQ-VT version 2.1[39] was used to support computer aided data collection. EQ-PVT is a portable version of EuroQol Valuation Technology (EQ-VT) which allows preference data to be collected without requiring direct links to the EuroQol group software which is necessary for EQ-VT. It runs using similar algorithms to EQ-VT but these are executed via a program in Microsoft PowerPoint. Data are stored on the computer and can then be uploaded to secure sites when internet links are available.

The interviews were undertaken by 10 trained pharmacy Master's students from March to May, 2018. A three-day training course on the methodology and study procedures was provided to the interviewers. The training material from the EuroQol group was translated into Amharic. The training material included: (1) an introduction to related concepts such as health-related quality of life and the EQ-5D-5L as a generic questionnaire used to value health states; (2) an explanation of the EQ-PVT protocol and interviewer instructions; (3) practice in groups; and (4) pre-testing of the tools to ensure that interviewers understood the task.

During the data collection, all data collectors travelled in one group to the study sites. Regular supervision was done on each data collection day by the trained principal investigators (authors AG and GBG).

Each interview consisted of a paper based and a computer assisted task. The paper-based task includes basic background questions, self-reported health measurement using the 5 item EQ-5D-5L, and the companion visual analogue scale (VAS) of 0 (the worst health you can imagine) to 100 (the best health you can imagine).

The computer assisted interview task included a practice TTO task (valuation of a mild, moderate, and severe health state). In addition, based on the pilot participants were asked to provide an example of better and worse health state than being confined to a wheelchair by themselves and where this was not possible, wearing glasses or confined to bed were provided as examples respectively. This was followed by ten TTO tasks and seven DCE tasks. There were 86 TTO health states in ten blocks with each block containing a range of health states across the range of severity from mild to severe as well as the worst health state (55555). Health states are presented randomly within the block. There were seven DCE pairs in twenty-eight blocks which were also presented randomly. One block from TTO and DCE was randomly assigned to each participant.

Quality control (QC)

We used a cyclic quality control (QC) process employed using the EQ-PVT QC tool developed by the EuroQol group[16]. The EuroQol group's expert and the Ethiopia team organized Skype-based meetings every twenty interviews for the pilot study and every hundredth interview for the main study in order to discuss the QC reports with the EQ -PVT support team. The QC process consisted of an evaluation of protocol compliance of 40% threshold as cut-off points to assess interviewer effects[40]. The QC reports provided a number of statistics related to the quality of the data collected, differentiated by interviewer. Criteria used to evaluate the quality of the interviews included the following:

- Wheelchair time: when the duration of time an interviewer used to explain the 'wheelchair example' preceding the actual *c*-TTO tasks was less than 3 minutes.
- Wheelchair lead-time: when the interviewer did not explain the "worse than dead" element of the wheelchair or worse than wheelchair example.
- *c*-TTO duration: if completing the ten *c*-TTO tasks took less than 5 min.

- Inconsistency: 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 any of the four above-mentioned criteria were met, the interview was 'flagged' so that further discussions about the interview process could be undertaken. The distribution of data obtained from different interviewers was also reviewed and any anomalies such as spikes at critical points (1 0.5 0 -0.5 -1) or other differences were highlighted for further discussion. Spikes may indicate poor engagement at the interviewer level while distributions that are different may indicate interviewer-effects.

Pilot study

The feasibility of the valuation protocol for EQ-5D-5L valuation studies using the EQ-PVT in Ethiopia was pilot tested in a sample of the population recruited from Addis Ababa. Two Ethiopian principal investigators who were trained in the Netherlands completed 110 interviews in the pilot study from January to February, 2018 with QC support. The pilot sample was recruited using national strata on age, gender and religion. Interviewers paid attention to the understanding of the tasks, level of engagement, acceptability of thinking/talking about death, and acceptability of EQ-5D-5L health states. The concordance between *c*-TTO and DCE values was also investigated. The data of the pilot study were not part of the final dataset.

The QC process in the pilot identified issues related to the wheelchair example and presenting lead time TTO and these were discussed and highlighted when training the other interviewers. The acceptability of the tasks was good and there were no special concerns with thinking and talking about death. However, a frequently reported problem was that respondents found health states difficult to imagine. To resolve this issue, practice tasks were tailored to the local context by providing examples that were more meaningful i.e. wearing glasses and confined to bed. Preliminary analysis of the pilot data indicated that DCE data was fully consistent but there

were inconsistencies in TTO data but this was not considered problematic as sample sizes were small.

Data analysis

Statistical analysis was undertaken using the STATA 14.2 statistical package. Descriptive statistics were used to summarize respondents' characteristics and responses to the *c*-TTO and DCE tasks and the relative importance of each attribute is evaluated using logistic regression methods. We used a method developed by Ramos-Goñi *et al.,*[41] to model value sets. A variety of preference models were produced using both types of data (*c*-TTO and DCE) individually and together to provide complementary evidence on preferences by taking into account the nature of preference data that are 'bounded' (censored); heterogeneity of respondents' views in health utilities; and heteroskedasticity of the error terms. For *c*-TTO, several models were tested including ordinary least squares, generalised linear models, random coefficient models and tobit models taking into account the panel structure of the data. The dependent variable in the *c*-TTO part of the model was disutility (defined as 1 minus the *c*-TTO observed values) for a given health state.

DCE data were analysed using the likelihood function of a conditional logit distribution. In the DCE model, the dependent variable was a binary outcome 0/1 indicating the respondent's choice for each pair of EQ-5D-5L states. As the coefficients estimated from a conditional logit are expressed on a latent arbitrary utility scale, a rescaled parameter was used, which assumes that the *c*-TTO model coefficients are proportional to DCE model coefficients. The coefficients represented the utility decrements for the DCE rescaled model[41].

Continuous responses from *c*-TTO and dichotomous responses from DCE were combined in a single model using 'hyreg' command developed by Ramos Goñi *et al*[41] to undertake hybrid

models. This command allows the continuous and dichotomous responses to have different distributions (logistic and normal), and have different independent variables to model scaling terms. Since the variance of *c*-TTO data is not homogenous, a heteroskedasticity model was estimated in which *c*-TTO responses were censored at −1.

A main effects 20-parameter model consisting of 4 dummies for each EQ-5D-5L dimensions was explored using level 1 as the reference. Dummies were constructed to represent additional utility decrement of moving from one level to another. For instance, the mobility dimension had 4 dummies - MO2 to MO5. The coefficient associated with MO2 indicated the utility decrement of moving from no problems (level 1) to slight problems (level 2), MO3 the additional utility decrement of moving from slight (level 2) to moderate (level 3) problems, and so on. Therefore, the overall decrement of moving from no to severe problems could be calculated as the sum of the coefficients of MO2 to MO5. The same set of dummy variables were defined for each of the remaining dimensions: self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD).

Some TTO data were excluded including data from: a) participants who gave all 10 health states the same value; and (b) participants who gave the worst state, 55555, a value that was no lower than the value they gave to the mildest health state in their block.

Evaluation of model performance

Model performance was evaluated using logical consistency of parameters, goodness of fit and significance level. Estimated coefficients are said to be logically consistent if estimated coefficients of the parameters were positive; when the estimated coefficients magnitude values of logically worse health states are lower than those from logically better health states. Goodness of fit was assessed using the Akaike (AIC) and the Bayesian information criteria

(BIC). Finally, the level of significance was assessed based on *p*-value of the models. Models that met the performance criteria are reported.

Results

Respondent characteristics

Ten interviewers completed 1,050 interviews. No interviewers were excluded from the study; Figure 1 shows the distribution of observed TTO values by interviewer. A total of 1,041 responses for the TTO and 1,048 for the DCE task formed the sample for analysis. Nine participants were dropped TTO task from the main survey, because they were non-traders and two from the DCE task because of participant ID overlapping during the data collection. The characteristics of sample respondents were similar to the Ethiopian general population in terms of residence, age, gender, and religion (Table 1).

Self-reported health problems

As depicted in Table 2, the highest proportion of health problems were reported in the anxiety/depression dimension (43.29%) and least in the self-care dimension (6. 57%). The mean self-reported EQ VAS score was 87.27 (SD=13.63).

Modeling Results

Very few states were flagged in the TTO feedback module (0.5%) so these were retained in the analysis. There were 837 (8.04%) left-censored *c*-TTO observations, i.e. those in which the respondent gave the lowest possible value (-1) for a health state in the *c*-TTO task with evidence of a spike at 1 (Figure 1). An ordinary least squares (OLS) model was used for *c*-TTO observation data and all of the responses were logically consistent. The conditional logistic regression model was used to model the DCE responses. In the conditional logistic regression

model of DCE data had two inconsistent (negative) values in self-care and usual activity dimensions of level 3.

Table 3 shows estimation results from *c*-TTO, DCE, and hybrid model. The models set of coefficients were in relative agreement; that is, in the disutility amount the most important dimensions were anxiety/depression and pain/discomfort; and the least important was self-care. The hybrid model parameters were logically consistent and all parameters except one parameter (PD3) were statistically significant. It was in relative agreement of the 20 parameters with *c*-TTO, DCE and hybrid models. Anxiety/depression dimension influenced utility estimates the most, (disutility level five of 0.4578), and self-care influenced the least (disutility level five 0.2224) in the final value set of the hybrid model. Figure 2 shows the association between the three models. The scatterplots of the different models suggest the compatibility of TTO model and hybrid model and shows the effect of adding the DCE data to the *c*-TTO valuation in the hybrid model.

The observed value of utility ranged from -0.718 for state 55555 to 0.974 for states 11112 in the hybrid model which is preferred. Values for 11121 and 11122 were found to be 0.964 and 0.938, respectively. The mean observed value was negative for 10 out of 86 states that were included in the design.

In this sample, the mean (standard deviation) EQ-5D-5L values were 0.94 (0.10) based on the hybrid model. To calculate utility values, the sum of the utility decrements for the relevant level e.g. for level 3 this would be the sum of level 2 and level 3, health state 11113 is given by 1 – (0.0259+0.0589) = 0.915.

DISCUSSIONS

This study presents social preferences and an EQ-5D-5L value set from the Ethiopian general population. To obtain values attached to 3,125 EQ-5D-5L health states, 1,050 respondents were interviewed using the computer-assisted valuation protocol EQ-PVT with an extensive interviewer training and data inspection. The intensive QC ensured high data quality in terms of few inconsistencies, little clustering of values, and low interviewer effect. The socio-demographic characteristics of the respondents were similar to the general population with respect to geographic area/residence; gender; age group; and religion. This makes EQ-5D-5L suitable for health economic evaluations that will benefit the national health care financing with the principle of HTA.

Value sets have been generated based on *c*-TTO data and DCE data alone. The final value set reported here is derived from a hybrid of both *c*-TTO and DCE data because the two data provide different but complementary information about the views of the respondents. The hybrid model maximizes data usage, and it gives the highest validity on parameter estimation because *c*-TTO measures utilities trading-off poor health against time, and the DCE task asked respondents to trade-off between two poor health states. A hybrid model was also used by different studies[15,16,20,21,23,24] to estimate the value set of EQ-5D-5L. Results from the hybrid model had logically consistency of parameters. All the dimensions were statistically significant except one parameter (pain and discomfort level 3). This EQ-5D-5L value set considered the complementary data in a hybrid model using an innovative model[15,42] which takes into account that *c*-TTO data are left censored and the heteroskedasticity of the error terms to decrease biased parameter estimates.

Similar to other countries such as the Netherlands[25] and England[15], the preferences of the Ethiopian population suggest that anxiety/depression and pain/discomfort are the health problems that are most important. On the other hand, mobility is the most important dimensions

in Indonesia[20], Korea[22] Japan[21], Canada[19], and Uruguay[17]. While problems with self-care and usual activity dimensions are less important and dimensions which has the smallest impact to the utility decrement was varied by countries. In Ethiopia, the maximum predicted value beyond full health was 0.9741 for the '11112' state, while it was11112 for Indonesia[20] (utility = 0.921), 11211 and 12111 for England[15] (utility = 0.950), 11211 for China[18] (utility = 0.955).

There were 837 (8.04%) observed -1 value on the *c*-TTO observations: when respondent gave the lowest possible value (-1) for a health state in the *c*-TTO task which is the highest as compared to other published studies[15,16,18,20,25]. This study yields a utility value in the ranged of -0.718 for state 55555 to 0.974 for states 11112.Compared to other study used hybrid model, the worst health state had higher value than Indonesia (-0.865)[20] but lower than English population (-0.285)[15].

This study has several strengths. It is the first study in Ethiopia as well as in Africa to report a value set for the EQ-5D-5L. The data have been generated using an international standardized protocol developed by EuroQol group. The value set could also be potentially used by other African countries as although there are important cultural differences, these may be relatively smaller compared to other countries that have EQ-5D-5L tariffs. Furthermore, it provides evidence of the feasibility of valuing the EQ-5D-5L using the EQ-PVT which uses PowerPoint. This provides evidence that the more cost-effective EQ-PVT approach can be used in low-resource settings. EQ-PVT is also of advantage where reliable internet connections are not available such as in rural areas. The frequent QC meetings through Skype with EuroQol group team during the course of study strengthened the approach to high quality data.

Although the TTO data have good face validity, a potential limitation of the study is that there is evidence of clustering of values at 1 (non-trading for mild health states). This occurred despite assessment of performance of interviewers and may suggest that in this context,

participants were unwilling to trade-off time for what they consider to be mild states. There are also limitations in terms of differences in the distribution of background variables in the sample compared with the data provided by the National Bureau of Statistics. However, these differences are small, and limited to some residence and religion. The strategy of relying on supervisors to act as mediators between interviewers and the respondents based on personal networks may have introduced bias. A further investigation could be conducted to find out whether recruiting respondents via personal networks has an impact on data quality. A final limitation was that the data were only collected in Addis Ababa city and Butajira rural area of Gurage zone southern part of Ethiopia. This might raise questions about the representativeness of the study sample but as noted, these areas were selected for their diverse population interms of ethnicity and culture.

Conclusions

This study showed that it was feasible and culturally acceptable to estimate preferences for health states using the EQ-PVT software of EQ-5D-5L valuation. This study established an Ethiopian value set for EQ-5D-5L on the basis of *c*-TTO and DCE from the general population of Ethiopia. We expect that this work will serve as the foundation for applied health economic evaluation, lead to additional health preference studies, and to inform decision making in Ethiopia.

Conflict of interest: Elly Stolk and Clara Mukuria are members of the EuroQol Research Foundation. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

AG, GBG, TGF and MK conceived the idea; AG, TGF, GBG, MK and FE designed the study, acquired the funding, and provided detailed information regarding data collection processes in Ethiopia. ES and CM provided detailed information regarding data collection processes. AG, GBG, ES and CM performed data quality control. AG prepared the draft

manuscript. All authors reviewed the analysis, interpretation of the results, and the final manuscript.

REFERENCES

1. Chalkidudou K, Levine R, Dillon A. Helping poorer countries make locally informed health. *Br Med J*. 2010;(16):1-4. doi:10.1136/bmj.c3651
2. Attieh R, Gagnon MP. Implementation of local/hospital-based health technology assessment initiatives in low- and middle-income countries. *Int J Technol assessmrnt Heal care*. 2012;28(4):44451. doi:10.1017/S026646231200058X
3. Da´ vid Danko´. Health technology assessment in middle-income countries: recommendations for a balanced assessment system ´. *J Mark access Heal policy*. 2014;2(1):1-10.
4. Mueller D, Govender B, Basu D. Health technology assessment in South Africa – future promise Launch of the Southern African Human Genome Programme. *South African Med J*. 2011;101(5):285-286.
5. Towse A, Devlin N, Hawe E, Garrison L. The Evolution of HTA in Emerging Markets Health Care Systems : Analysis to Support a Policy Response. *Off Heal Econ*. 2011:1-119.
6. Babigumira JB, Jenny AM, Bartlein R, Stergachis A, Garrison LP. Health technology assessment in low- and middle-income countries : a landscape assessment. *J Pharm Heal Serv Res*. 2016;7(1):37-42. doi:10.1111/jphs.12120
7. Rajan A, Gutierrez-Ibarluezea I, Moharra M. Addressing issues in health technology assessment promotion : Motives , enablers and, barriers. *Int J Technol Assessmrnt Heal Care*. 2011;27(1):55-63. doi:10.1017/S0266462310001352
8. Weinstein MC, Torrance G, Mcguire A. QALYs : The Basics. *Value Heal*. 2009;12:5–9.
9. Brockis E, Marsden G, Cole A, Devlin N. A Review of NICE Methods Across Health Technology Assessment Programmes : Differences , Justifications and Implications. *Off Heal Econ*. 2016:1-33.
10. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines : results of a systematic review and expert consultation across eight European countries. *Eur J Heal Econ*. 2018;19(1):123-152. doi:10.1007/s10198-017-0871-0

11. Weinstein M, Kaiser HJ. Cost-per-QALY in the US and Britain : Damned if you Do and Damned if you Don ' t. Monographs. *Off Heal Econ*. 2011:1-16.

12. Guidelines for the Economic Evaluation of Health Technologies : Canada. 2017;4th ed.Ott(Mar).

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

14. van Reenen M, Janssen B. EQ-5D-5L User Guide basic information on how to use the EQ-5D-5L instrument. Rotterdam: *EuroQol Res Found*. 2015:1-25.

15. Devlin, N., Shah, K., Mulhern, B., Feng, Y., Tsuchiya, A. and van Hout BA. Valuing health - related quality of life : An EQ - 5D - 5L value set for England. *Health Econ*. 2014;27(1):7-22. doi:10.1002/hec.3564

16. Ramos-goñi JM, Craig BM, Oppe M, et al. Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach. *Value Heal*. 2018;21(5):596-604. doi:10.1016/j.jval.2017.10.023

17. Augustovski F, Irazola LRV, Morales M, Gibbons L, Manuel J, Ramos-gon JM. An EQ-5D-5L value set based on Uruguayan population preferences. *Qual Life Res*. 2016;25(2):323-333. doi:10.1007/s11136-015-1086-4

18. Luo N, Liu G, Li M, Guan H, Jin X, Rand-hendriksen K. Estimating an EQ-5D-5L Value Set for China. *Value Heal*. 2017;20(4):662-669. doi:10.1016/j.jval.2016.11.016

19. Xie F, Nick B, Stirling B, et al. A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada. *Med Care*. 2016;54(1):98-105.

20. Purba FD, Hunfeld JAM, Iskandarsyah A, Passchier J, Busschbach JJ V. The Indonesian EQ-5D-5L Value Set. *Pharmacoeconomics*. 2017;35(11):1153-1165. doi:10.1007/s40273-017-0538-9

21. Shiroiwa T, Ikeda S, Noto S, Igarashi A. Comparison of Value Set Based on DCE and / or TTO Data : Scoring for EQ-5D-5L Health States in Japan. *Value Heal*. 2016;19(5):648-654. doi:10.1016/j.jval.2016.03.1834

22. Kim S, Ahn J, Ock M, et al. The EQ-5D-5L valuation study in Korea. *Qual Life Res*. 2016;25(7):1845-1852. doi:10.1007/s11136-015-1205-2

23. Pattanaphesaj J, Thavorncharoensap M, Ramos- JM, Tongsiri S, Ingsrisawang L, Teerawattananon Y. The EQ-5D-5L Valuation Study in Thailand. *Expert Rev*

Pharmacoecon Outcomes Res. 2018;18(5):551-558. doi:10.1080/14737167.2018.1494574

24. Ludwig K, Schulenburg JG Von Der, Greiner W, Ludwig K. German Value Set for the EQ-5D-5L. *Pharmacoeconomics*. 2018;36(6):663-674. doi:10.1007/s40273-018-0615-8

25. Versteegh MM, Vermeulen KM, Evers SMAA, Wit GA De, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. *Value Heal*. 2016;19(4):343-352. doi:10.1016/j.jval.2016.01.003

26. Rencz F, Gulácsi L, Drummond M, Golicki D, Rupel VP, Simon J, Stolk EA, Brodszky V, Baji P, Závada J PG. EQ-5D in Central and Eastern Europe : 2000 – 2015. *Qual Life Res*. 2016;25(11):2693-2710. doi:10.1007/s11136-016-1375-6

27. Doaee SH, Olyaeemanesh A, Emami SH, Mobinizadeh M, Abooee P, Nejati M. Development and Implementation of Health Technology Assess- ment : A Policy Study. *Iran J Public Health*. 2013;42(1):50-54.

28. Oortwijn W, Mathijssen J, Banta D. The role of health technology assessment on pharmaceutical reimbursement in selected middle-income countries. *Health Policy (New York)*. 2010;95(2-3):174-184. doi:10.1016/j.healthpol.2009.12.008

29. Ali EE. Health Care Financing in Ethiopia : Implications on Access to Essential Medicines. *Value Heal Reg Issues*. 2014;4:37-40. doi:10.1016/j.vhri.2014.06.005

30. Oppe, M., & van Hout B. EuroQol Working Paper Series experimental design of the EQ-VT. *EuroQol Res Found*. 2017:1-17.

31. Berhane Y, Byass P. Physical Geography of the Butajira DSS Area. 2002:0-9.

32. Spaliviero M CF. *The State of Addis Ababa*.; 2017.

33. Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the Measurement of Health Utilities in Published Cost-Utility Analyses. *Value Heal*. 2006;9(4):213-218.

34. Brooks, Richard with the EG. EuroQol : the current state of play. *"Health policy."* 1996;37(1):53-72.

35. Herdman M, Gudex C, Lioyd C, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x

36. Oppe M, Devlin NJ, Hout B Van, Krabbe PFM, Charro F De. A Program of Methodological Research to Arrive at the New International EQ-5D-5L Valuation Protocol. *Value Heal*. 2014;17(4):445-453. doi:10.1016/j.jval.2014.04.002

37. Devlin NJ, Tsuchiya AKI, Buckingham KEN, Tilling C. A uniform time trade off method for states better and worse than dead : feasibility study of the ‘ lead time ’ approach. *Health Econ*. 2011;20(3):348-361. doi:10.1002/hec

38. Lancsar E, Louviere J. Conducting discrete choice experiment to Inform Healthcare Decision Making. *Pharmacoeeconomics*. 2008;26(8):661-677.

39. Stolk E, Ludwig K, Rand K, Hout B Van, Ramos-go 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*. 2018:1-8. doi:10.1016/j.jval.2018.05.010

40. Ramos-goñi JM, Oppe M, Slaap B, Busschbach JJ V, Stolk E. Quality Control Process for EQ-5D-5L Valuation Studies. *Value Heal*. 2017;20(3):466-473. doi:10.1016/j.jval.2016.10.012

41. Ramos-goñi JM, Craig B, Oppe M, van Hout B. Introducing a hybrid model that combines continuous and dichotomous responses in a single maximum likelihood function: the hyreg command. *Stata J*. 2016:Under review.

42. Ramos-gon JM, Serrano-aguilar P, Rivero-arias O, Cabase JM. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. 2014;00(00):1-8.

Table 1. Background characteristics of the sample and general population

Characteristics	Study sample (N = 1050), n (%)	Ethiopian general population (%)*
Residence		
Urban	300 (28.57)	19.92
Rural	750(71.43)	80.08
Gender		
Female	503(47.90)	50.15
Male	547(52.10)	49.85
Age (in years)		
18-24	391(37.24)	36.00
25-54	575(54.76)	52.32
55-64	49(4.67)	6.91
65+	35(3.33)	5.20
Marital status		
Married	559(53.24)	NA
Unmarried	444(42.23)	NA
Divorced	22(2.10)	NA
Widowed	25(2.40)	NA
Education		
No formal education	440(41.90)	NA
Primary school	173(16.48)	NA
Secondary school	265(25.33)	NA
Higher school	171(16.29)	NA
Employment status		
Employed or self-employed	693(66)	NA
Retired	30(2.86)	NA
Student	214(20.40)	NA
Looking after home or family	111(10.60)	NA
Other / none of the above	2(0.19)	NA
Experience of serious illness		
In self	142(13.52)	NA
In family	329(31.33)	NA
In self and in family	172(16.36)	NA
In caring for others	152(14.48)	NA
No	255(24.28)	NA
Religion		
Christian	679(64.67)	63
Muslim	368(35.05)	34
Others	3(0.29)	3.0

*Data obtained from the 2016 Ethiopian health survey

NA indicates not available

Table 2. Self-reported health using the EQ-5D-5L descriptive system and the EQ VAS

	EQ-5D-5L descriptive system with scores in %				
Parameters	**Mobility**	**Self-care**	**Usual activities**	**Pain/ discomfort**	**Anxiety/depression**
No problems	88.19	93.43	83.33	58.09	56.38
Slight problems	7.61	4.70	11.43	27.43	33.71
Moderate problems	4.00	1.61	4.66	12.00	8.28
Severe problems	0.09	0.19	0.57	2.28	1.33
Unable/extreme problems	0.09	0.00	0.00	0.19	0.28

	Mean	**SD**	**25th percentile**	**Median**	**75th percentile**
VAS score	87.26	13.64	75	90	95

Table 3. Estimation results for C-TTO model, DCE rescaled model, and hybrid model in incremental dummies.

Independent variables of the model	C-TTO OLS model			DCE conditional logistic model rescaled			Hybrid model censored C-TTO values at -1 (final value set)		
	Coef.	(SE)	*p*-value	Coef.	(SE)	*p*-value	Coef.	(SE)	*p*-value
Mobility (MO)									
MO2	0.0047	0.014	0.729	0.4780	0.061	0.000	0.0337	0.005	0.000
MO3	0.0166	0.015	0.262	0.1138	0.071	0.110	0.0307	0.009	0.000
MO4	0.1748	0.016	0.000	0.9810	0.070	0.000	0.1632	0.010	0.000
MO5	0.1038	0.016	0.000	0.7434	0.074	0.000	0.1322	0.010	0.000
Self-care (SC)									
SC2	0.0036	0.013	0.785	0.2044	0.067	0.002	0.0235	0.005	0.000
SC3	0.0494	0.016	0.002	-0.0024	0.074	0.974	0.0160	0.008	0.042
SC4	0.1189	0.015	0.000	0.6849	0.078	0.000	0.1024	0.009	0.000
SC5	0.0826	0.013	0.000	0.4234	0.073	0.000	0.0804	0.009	0.000
Usual-activities (UA)									
UA2	0.0188	0.014	0.176	0.3470	0.063	0.000	0.0323	0.005	0.000
UA3	0.0441	0.014	0.002	-0.0391	0.071	0.579	0.0160	0.008	0.042
UA4	0.1299	0.016	0.000	0.5818	0.071	0.000	0.1091	0.009	0.000
UA5	0.0936	0.015	0.000	0.6079	0.076	0.000	0.1147	0.010	0.000
Pain/discomfort (PD)									
PD2	0.0140	0.013	0.266	0.4499	0.067	0.000	0.0361	0.004	0.000
PD3	0.0161	0.017	0.331	0.1090	0.073	0.136	0.0155	0.008	0.061
PD4	0.2452	0.015	0.000	1.1358	0.077	0.000	0.2187	0.010	0.000
PD5	0.1421	0.016	0.000	0.5689	0.076	0.000	0.1361	0.011	0.000
Anxiety/depression (AD)									
AD2	0.0111	0.014	0.428	0.2718	0.070	0.000	0.0259	0.004	0.000
AD3	0.0381	0.015	0.012	0.3516	0.072	0.000	0.0589	0.008	0.000
AD4	0.2322	0.015	0.000	1.1803	0.079	0.000	0.2139	0.009	0.000
AD5	0.1414	0.013	0.000	0.8320	0.078	0.000	0.1591	0.010	0.000
AIC	10587.06			6498.30			14002.09		
BIC	10739.33			6650.17			14336.81		
Order of importance									
	AD			AD			AD		
	PD			MO			PD		
	MO			PD			MO		
	UA			UA			UA		
	SC			SC			SC		

Coef. – coefficient; SE – standard error

Items with a negative coefficient (in grey) represent inconsistent items
Order of importance based on sum of disutility which is the disutility associated with level 5

Figure 1. Distribution of observed TTO values of every interviewer and in total.

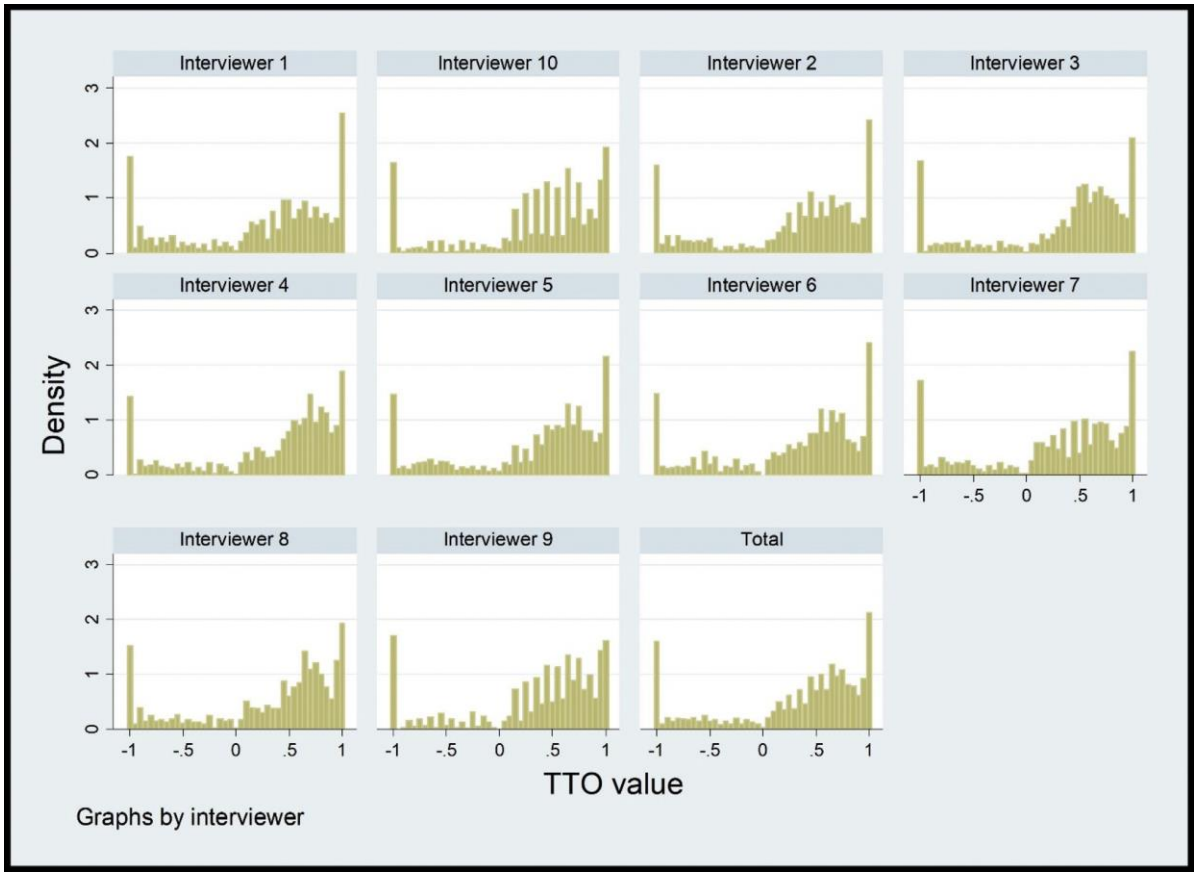

Figure 2. The scatterplots show the association between the three models of c-TTO and DCE rescaled predicted utilities, c-TTO and hybrid predicted utilities, DCE rescaled and hybrid predicted utilities, respectively.